Psychological Review
January 1995 Vol. 102, No. 1, 4-27

© 1995 by the American Psychological Association, Inc.
For personal use only--not for distribution.

Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes

Anthony G. Greenwald
Department of Psychology University of Washington
Mahzarin R. Banaji
Department of Psychology Yale University

ABSTRACT

Social behavior is ordinarily treated as being under conscious (if not always thoughtful) control. However, considerable evidence now supports the view that social behavior often operates in an implicit or unconscious fashion. The identifying feature of implicit cognition is that past experience influences judgment in a fashion not introspectively known by the actor. The present conclusion–that attitudes, self-esteem, and stereotypes have important implicit modes of operation–extends both the construct validity and predictive usefulness of these major theoretical constructs of social psychology. Methodologically, this review calls for increased use of indirect measures–which are imperative in studies of implicit cognition. The theorized ordinariness of implicit stereotyping is consistent with recent findings of discrimination by people who explicitly disavow prejudice. The finding that implicit cognitive effects are often reduced by focusing judges' attention on their judgment task provides a basis for evaluating applications (such as affirmative action) aimed at reducing such unintended discrimination.

Long before they became central to other areas of psychological theory, concepts of cognitive mediation dominated the analysis of social behavior. The constructs on which this article focuses achieved early prominence in social psychological theory with formulations that were partly (attitude) or entirely (stereotype) cognitive. By the 1930s, Allport (1935) had declared attitude to be social psychology's "most distinctive and indispensable concept" (p. 798), Thurstone (1931 ; Thurstone & Chave, 1929) had developed quantitatively sophisticated methods for attitude measurement, and Katz and Braly (1933 , 1935) had introduced a method that is still in use to investigate stereotypes. Self-esteem, an attitudinal construct to which this article gives separate treatment because of its prominence in recent research, also has a long-established history (e.g., James, 1890 ; see overview in Wylie, 1974 , 1979).

Through much of the period since the 1930s, most social psychologists have assumed that attitudes, and to a lesser extent stereotypes, operate in a conscious mode. This widespread assumption of conscious operation is most evident in the near-universal practice of operationalizing attitudes (including self-esteem) and stereotypes with direct (instructed self-report) measures. The pervasiveness of direct measurement for attitudes and stereotypes was documented by Greenwald (1990) and by Banaji and Greenwald (1994) and is further reviewed below. In contrast, this article describes an indirect, unconscious, or implicit mode of operation for attitudes and stereotypes. [1]

Implicit Social Cognition: Introduction and Overview

Implicit social cognition is offered as a broad theoretical category that integrates and reinterprets established research findings, guides searches for new empirical phenomena, prompts attention to presently underdeveloped research methods, and suggests applications in various practical settings. This section summarizes the goals of the review, starting from a definition of implicit social cognition.

Definition

The signature of implicit cognition is that traces of past experience affect some performance, even though the influential earlier experience is not remembered in the usual sense–that is, it is unavailable to self-report or introspection (cf. Graf & Schacter, 1985 ; Greenwald, 1990 ; Jacoby & Dallas, 1981 ; Jacoby, Lindsay, & Toth, 1992 ; Jacoby & Witherspoon, 1982 ; Kihlstrom, 1990 ; Roediger, Weldon, & Challis, 1989 ; Schacter, 1987). As an illustration of implicit cognition, consider a result that is readily obtained with the task of generating complete words in response to incomplete letter strings (word stems or word fragments). The completion responses are more likely to be words from a list to which subjects were casually exposed earlier in the experiment than to be equally likely words that were not presented. This effect of prior exposure occurs despite subjects' poor ability to recall or recognize words from the earlier list. The word-completion task provides an indirect measure of the effect of the prior experience. That is, even though the subject is not instructed to retrieve the earlier presented material and is presumably not trying to do so (and may well be incapable of such retrieval), the subject's responses indicate a residual effect. (For further reviews of indirect measurement in memory research, see Richardson-Klavehn & Bjork, 1988 ; Roediger, 1990 ; Roediger & Blaxton, 1987 ; Tulving & Schacter, 1990 . For further discussion of direct and indirect measures in other contexts that involve unconscious cognition, see Jacoby, Lindsay, & Toth, 1992 ; Reingold & Merikle, 1988 .)

A template for definitions of specific categories of implicit cognition is: *An implicit C is the introspectively unidentified (or inaccurately identified) trace of past experience that mediates R*. In this template, *C* is the label for a construct (such as attitude), and *R* names the category of responses (such as object-evaluative judgments) assumed to be influenced by that construct.

To the extent that implicit cognition differs from self-reportable (conscious or explicit) cognition, direct measures–that is, measures that presume accurate introspection–are necessarily inadequate for its study. Rather, investigations of implicit cognition require indirect measures, which neither inform the subject of what is being assessed nor request self-report concerning it. The usual justification for indirect measures in social psychological research is the empirical one of minimizing reactivity of research situations to avoid demand characteristics (Orne, 1962) and sources of self-presentational artifacts such as evaluation apprehension (Rosenberg, 1969). When used in this way to minimize reactivity, indirect measures are empirically desirable but not theoretically essential. By contrast, in studying implicit cognition, indirect measures are theoretically essential.

Theory: Relation to Other Treatments of Unconscious Aspects of Social Cognition

Much of the present review concerns unconscious cognitive involvement in (and especially interference with) deliberate judgments. This focus is still infrequent in social cognition literature, perhaps because the range of deliberate judgments that are affected by unconscious cognition has only recently become apparent, as a consequence of the explosion of interest in implicit memory. The existing work to which the present treatment is closest is that of Jacoby and colleagues (e.g., Jacoby & Dallas, 1981 ; Jacoby, Toth, Lindsay, & Debner, 1992 ; Jacoby & Witherspoon, 1982). Jacoby's group has pioneered theory and methods to identify implicit memory influences and has effectively established an important role of unconscious cognition in deliberate judgments. The phenomena investigated by Jacoby have often been ones in which the research subject mistakenly attributes ease of perception on reexposure to a stimulus (perceptual fluency) to some characteristic of the stimulus other than an unremembered recent encounter. The present analysis of implicit social cognition uses the same basic misattribution principle, while focusing on the social domain to locate influential prior experiences and affected target judgments.

The following paragraphs review the various ways in which unconscious cognition has already been integrated into social psychological theory. After considering the three categories that are focal to this review–attitudes, self-esteem, and stereotypes–the focus shifts to some related topics.

Attitudes.

Recent work has established that attitudes are activated outside of conscious attention, by showing both that activation occurs more rapidly than can be mediated by conscious activity (Bargh, Chaiken, Govender, & Pratto, 1992 ; Fazio, Sanbonmatsu, Powell, & Kardes, 1986) and that activation is initiated by (subliminal) stimuli, the presence of which is unreportable (Greenwald, Klinger, & Liu, 1989). The present analysis of

implicit attitudes extends work on automatic activation to explain how the attitude activated by one object can be (mis)attributed to another. An implicit attitude can be thought of as an existing attitude projected onto a novel object. The interpretation of several important existing findings as implicit attitude effects substantially expands the predictive and construct validity of social psychology's attitude construct. It also prompts the empirical search for further members of the potentially large class of implicit attitude effects. In the domain of attitude change, two recent theoretical analyses (Chaiken, 1987 ; Petty & Cacioppo, 1986) have distinguished relatively thoughtful (central or systematic) from relatively thoughtless (peripheral or heuristic) roles of cognition in persuasion. The implicit processes conceived in the present analysis are, in part, subsumed by the notions of peripheral or heuristic processing, but also involve processes operating even further from the range of conscious thought than conceived in these analyses.

Self-esteem.

The broad importance of self-esteem has been recognized in many works over the past century (e.g., Allport, 1937 ; Cooley, 1902/1964 ; Epstein, 1973 ; James, 1890 ; Rogers, 1951 ; Rosenberg, 1979 ; Sherif & Cantril, 1947). Recent reviews have further expanded the case for importance of the self-attitude (e.g., Beck, 1979 ; Greenwald & Pratkanis, 1984 ; Seligman, 1991 ; Steele, 1988 ; S. E. Taylor & Brown, 1988 ; Tesser, 1988). In the course of examining available evidence for the implicit operation of attitudes, evidence for implicit operation specifically of the attitude toward self was so prominent as to prompt treatment of implicit self-esteem as a distinct topic. The present review of implicit self-esteem phenomena demonstrates the pervasiveness of projections of the self attitude onto other objects, while also indicating the need for a class of measures that presently does not exist–measures of individual differences in implicit self-esteem.

Stereotypes.

Recent reviewers have very effectively documented the unconscious or automatic operation of stereotypes (Banaji & Greenwald, in press ; Bargh, 1994 ; Devine, 1989 ; Dovidio, Evans, & Tyler, 1986 ; Fiske, 1989 ; Geis, 1993 ; Gilbert & Hixon, 1991 ; Hamilton & Sherman, 1994 ; Perdue & Gurtman, 1990). This article's use of the "implicit" label for stereotypes serves primarily to emphasize the connection of the existing body of social cognition research on stereotypes to recent cognitive psychological research on implicit memory (cf. E. R. Smith & Branscombe, 1988). The present treatment of stereotypes also includes new findings that demonstrate this connection, using the methods of implicit memory research to reveal implicit gender stereotypes.

Effortless or automatic social cognition.

Single words have been shown to result in effortless activation of attitudes (discussed in the next paragraph), and behavior-describing sentences produce spontaneous trait inferences (Uleman, 1987 ; Winter & Uleman, 1984). These and other effortless activation effects often contribute importantly to, but are not synonymous with, implicit social cognition. As will be seen below, an implicit effect can occur when an actor (a) notices some aspect of an automatic effect caused by one stimulus and (b) mislabels it in a way that that influences the judgment of either that stimulus or some other stimulus. Discussions of effortless aspects of social cognition can be found in Bargh (1989) , Brewer (1988) , Epstein (1991) , Gilbert (1989) , and Lewicki (1986) .

Priming and context effects.

In common with implicit social cognition, priming and context effects involve the effect of prior events on the response to a current stimulus. However, whereas priming and context function as designations for operationally defined categories of effects, implicit cognition designates a theory-defined category of effects. Some established priming and context effects fall well within the boundaries of implicit social cognition. Others should not be grouped with implicit effects because of the subject's likely awareness of the effect of prior experience on behavior or judgments. Because the existing literatures on priming and context effects are very large, analysis to sort implicit effects from other effects within those literatures is beyond the scope of this article. Recent theoretical discussions of priming and context effects can be found in Higgins (1989) , Schwarz (1990) , Strack (1992) , Martin and Tesser (1992) , and Petty and Wegener (1993) .

Introspective access.

Nisbett and Wilson (1977a) argued persuasively that psychology's reliance on verbal self-report measures was unwarranted in the face of evidence showing the poverty of introspective access to the causes of behavior. Wegner and Vallacher (1977, 1981) also drew attention to influences on social behavior that escape introspective notice. They presented an "implicit psychology," drawing on and extending Bruner and Tagiuri's (1954) concept of implicit personality theory. However, because their work preceded the past decade's wave of research on implicit memory, Wegner and Vallacher did not share the present article's focus on empirical studies that use indirect measures. More recently, in several works, E. R. Smith (e.g., 1984 , 1990 , 1994) has emphasized the importance to social cognition of nonverbalized procedural knowledge. The present work shares Smith's emphasis on examining the introspectively inaccessible underpinnings of social cognition. In related work, Smith (see E. R. Smith & Branscombe, 1988) has also noted the relevance of implicit memory to social cognition. The research reviewed here sheds new light on the nature of causes that are hidden from introspection and provides some methods for observing them. In this respect, research in social cognition now appears to be taking an important further step along the path that was laid out by Nisbett and Wilson (1977a) .

Empirical Assessment Falsifiability.

Of any newly offered theoretical construct, it should be asked: How does the new construct differ from existing ones (or is it only a new label for an existing construct)? The preceding paragraphs show that implicit social cognition, although strongly rooted in existing constructs, offers a theoretical reorganization of phenomena that have previously been described in other ways and, in some cases, not previously identified as having an unconscious component. The relations to prior theorization are emphasized in this article by using established construct terms– *attitude* and *stereotype* –as labels for two major categories of implicit social cognition.

When a new construct indeed differs from existing ones, the new construct should provide a basis for either (a) predicting previously unobserved empirical phenomena or (b) guiding research to show a gain in the efficiency with which existing phenomena can be explained. Importantly, the construct should be well enough linked to research operations that its predictions, especially its predictions that differ from those afforded by existing constructs, can be disconfirmed.

This article's strategy is to identify parallels of method and findings between the domains of social cognition and implicit memory. The ease with which such parallels are discovered is the main evidence for value of the implicit social cognition constructs. This convergence-seeking strategy provides little opportunity for falsification of the general thesis that social cognition operates in implicit fashion. Rather, results that appear not to fit the thesis can be set aside as possibly involving inappropriate operationalizations. However, as parallels between the two domains of phenomena are increasingly established, there should be increased confidence in interpretations of research operations for social cognition constructs, and, consequently, results that fail to fit with theory will increasingly call theory into question.

Challenge to measurement method.

The present account supposes that individual differences in manifestations of implicit cognitive effects should be predicted by individual differences in the strength of theorized representations that underlie those effects. Measurement of those individual differences is beyond the means of present assessment technology; consequently, a large subset of the empirical implications of the present analysis are currently untestable. The need for new measurement methods and these methods' relation to available methods are discussed later in this article.

Application potential.

The empirical phenomena of implicit social cognition involve introspectively inaccessible effects of current stimulus or prior experience variations on judgments and decisions. As will be shown, these effects often result in subjects making judgments that they would regard as nonoptimal if made aware of the source of influence. Furthermore, these effects are likely to occur in situations that involve economically and socially important

decisions, such as hiring, educational admissions, and personnel evaluations. Consequently, a strong test of the empirical value of the analysis of implicit social cognition will be its ability to generate applications that can minimize these unwanted intrusions on judgment. This important application topic is considered briefly near the end of this article.

Generalizability.

It will be obvious that the great majority of evidence reviewed in this article comes from experimental studies done in late-20th-century North America. As a result, it is very likely that some of the specific properties of implicit attitudes, implicit self-esteem, and implicit stereotypes included in this review are culture bound and time bound. At the same time, there is no reason to believe that, as a general class of phenomena, implicit social cognition should be confined to modern North American culture. Furthermore, although the present treatment focuses on just the three categories of implicit attitudes, implicit self-esteem, and implicit stereotypes, implicit cognition is also expected to be involved broadly in other social phenomena.

The next three sections consider phenomena of implicit social cognition in the categories represented by construct designations of attitude, self-esteem, and stereotype. Following those comes a section that considers principles underlying potential applications that seek to reduce unintended implicit cognitive intrusions on judgment. Last, problems associated with measurement of implicit social cognition are discussed before general conclusions are stated.

Implicit Attitudes

Attitudes are favorable or unfavorable dispositions toward social objects, such as people, places, and policies. Attempts to establish the validity of the attitude construct have most often sought to demonstrate positive correlations between measured attitudes and the favorable—unfavorable aspect of observed behavior toward their objects. The frequently weak correlations observed in these attempts define the predictive validity problem for attitudes (documented especially by Wicker, 1969 ; see also Festinger, 1964 , and LaPiere, 1934). A notable accomplishment of modern research on attitudes has been the solution of this predictive validity problem. That is, conditions under which attitudes strongly correlate with behavior have now been well identified (especially by Ajzen & Fishbein, 1980 ; Fazio, 1986 , 1990b ; Fazio & Zanna, 1981 ; Fishbein & Ajzen, 1974 ; Zanna & Fazio, 1982). In particular, this research has established that attitudes have predictive validity in situations in which they are strongly activated and/or when the actor clearly perceives a link between attitude and behavior. Myers (1990) summarized these and related programs of research as showing that "our attitudes predict our actions... if, as we act, we are *conscious* of our attitudes" (Myers, 1990, p. 40 , emphasis added). Similarly, in the description of attitude—behavior relations in their recent treatise on the attitude construct, Eagly and Chaiken (1993, pp. 208—211) referred to the importance of attitudes "[coming] to mind" and the "perceived relevance" of attitude to action.

Although the modern synthesis achieved by the Fishbein—Ajzen (1974) and Fazio—Zanna (1981) research programs is now well established, it is difficult to avoid concluding that the attitude construct lost scope in the process. For those who can remember it, there might be justifiable nostalgia for an era in which Allport (1935) was able to proclaim that attitude was social psychology's "most indispensable concept." In an undisguised effort to restore this prominence, the present article seeks to preserve the modern synthesis while asserting that its opposite is also valid (cf. McGuire's [1973] Koan 7: "The opposite of a great truth is also true"); that is, attitudes of which the actor is *not* conscious at the moment of action (implicit attitudes) are also strongly predictive of behavior.

Ignored Consciousness in Conceptual Definitions of Attitude

The following list gives several definitions that have been influential in guiding scholarly and empirical treatments of attitudes, as indicated by their frequent citation in other works. Although the list may appear dated (the most recent entry is from 1962), it nevertheless remains current. Recent works (e.g., Eagly & Chaiken, 1993 ; Fazio, 1986 ; McGuire, 1985 ; Petty & Cacioppo, 1981 ; Zanna & Rempel, 1988) continue to draw on them and remain within their boundaries.

> Attitude is the affect for or against a psychological object. (Thurstone, 1931 p. 261,)
>
> An attitude is a mental and neural state of readiness, organized through experience, exerting a directive or dynamic influence upon the individual's response to all objects and situations with which it is related. (Allport, 1935, p. 810)
>
> Attitude is... an implicit, drive-producing response considered socially significant in the individual's society. (Doob, 1947, p. 136)
>
> An attitude is a predisposition to experience, to be motivated by, and to act toward, a class of objects in a predictable manner. (M. B. Smith, Bruner, & White, 1956, p. 33)
>
> [Attitudes] are predispositions to respond, but are distinguished from other such states of readiness in that they predispose toward an evaluative response. (Osgood, Suci, & Tannenbaum, 1957, p. 189)
>
> [An attitude is] a disposition to react favorably or unfavorably to a class of objects (Sarnoff, 1960, p. 261).
>
> Attitudes [are] enduring systems of positive or negative evaluations, emotional feelings, and pro or con action tendencies with respect to social objects. (Krech, Crutchfield, & Ballachey, 1962, p. 139)

The lack of mention of consciousness in this collection of attitude definitions accurately reflects a long scholarly tradition of nonconcern with the distinction between conscious and unconscious operation of attitudes. At the same time, nothing in this scholarly tradition actively opposes either the possibility or the importance of unconscious operation of attitudes.

Standing starkly in the above list as suggesting unconscious operation is Doob's (1947) definition, which labels attitude as an "implicit, drive-producing response." In spite of Doob's association with a behaviorist theory (Hull, 1943) that had no use for conceptions of either conscious or unconscious cognition, it is clear that Doob did conceive attitude as operating unconsciously (May & Doob, 1937, p. 13). In a recent communication to the present authors (October 27, 1992), Doob commented, "before World War II we all were impressed by psychoanalysis in addition to behaviorism," suggesting that, even though it may have gone unmentioned in many published treatments, the idea that attitudes operated unconsciously was quite acceptable in the 1940s and earlier. That conclusion is supported also by several passing references to the possibly unconscious nature of attitudes in Allport's (1935) review chapter.

Implied Consciousness in Operational Definitions of Attitude

The present authors conducted a census of studies published in all issues of *European Journal of Social Psychology, Journal of Personality and Social Psychology, Journal of Experimental Social Psychology,* and *Personality and Social Psychology Bulletin* for a single year (1989). The aim of this census was to compare the levels of use of direct and indirect measures of attitudes in published research literature. Studies that included attitude measures were categorized in terms of whether they measured attitude directly or indirectly. Direct measures included self-report procedures such as multi-item Thurstone, Likert, or semantic differential scales of the sort that are described in texts on attitude measurement (e.g., A. L. Edwards, 1957 ; Fishbein, 1967), as well as more informal single-item and multi-item self-report procedures. Indirect measures are identifiable chiefly by their lack of the defining feature of direct measures, that is, by their not alerting the subject to the identity of the object of the attitude being measured. Discussions and illustrations of indirect measures can be found in Hammond (1948) , Campbell (1950) , Jahoda, Deutsch, and Cook (1951) , Sherif and Hovland (1961, p. 93ff.) , Dawes and Smith (1985) , Pratkanis (1988) , and Dovidio and Fazio (1992) .

The authors' 1989 census included not only studies that obviously dealt with attitudes (as indicated by title or abstract), but also ones that more incidentally included measures of evaluation toward the self, others, or social

objects, for example in studies of person perception and in-group bias. Each of the 47 studies found to include an attitude measure was judged for presence of both direct and indirect attitude measures. All 47 (100%) used at least one direct measure of attitude, and only 6 (13%) used some form of indirect measure. An important class of attitudes discussed in this article is self-esteem. It is not necessary to conduct a census of publications to assert with confidence that self-esteem is generally assessed with direct measures (see Wylie, 1974 , 1979). In summary, the observed high level of reliance on direct measures of attitudes indicates a widespread (even if not widely stated) assumption that attitudes operate primarily in a conscious mode.

Definition of Implicit Attitude

Evidence concerning the strength of attitude—behavior relations has generally been regarded as the primary evidence bearing on predictive validity of the attitude construct. However, there are other categories of studies in which strong predictive effects that involve attitudes are routinely obtained. Examination of these now will indicate that some strong effects of attitude can and do occur when the actor is not attentionally focused on the attitude. These findings play a central role in justifying the concept of implicit attitude.

> *Implicit attitudes* are introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects. [2]

Relation to stimulus—response theory.

To those familiar with behavior theory of 30—50 years ago, the debt of this definition both to Doob's (1947) analysis of attitude as an implicit response and to subsequent mediationist stimulus—response theories will be obvious. The earliest mediationist stimulus—response formulations (mature statements of which appear in Hull, 1952 ; Spence, 1956) had proposed the existence of covert stimulus-producing responses (called "fractional anticipatory goal responses") as a means of explaining findings achieved within Tolman's (e.g., 1959) cognitive approach to learning theory. Dollard and Miller (1950) , Osgood (e.g., 1957) , and Mowrer (1960) most fully adapted these mediationist principles to the analysis of human social behavior. Campbell's (1963) treatise on the attitude construct preserved some of the implicit, mediationist character of the later behaviorist treatments, declaring that acquired dispositions such as attitudes "[retain] residues of experience of such a nature as to guide, bias, or otherwise influence later behavior" (p. 97). In retrospect, it is apparent that the conception of covert, stimulus-producing, mediating events did not take hold in social psychology when mediationist theory was dominant in learning-behavior theory. Nevertheless, that conception is a virtual equivalent of the view that reappears in the present definition of implicit attitude.

Direct versus indirect measures.

Experiments that demonstrate implicit attitudes rely on indirect measures to detect their operation. The distinction between direct and indirect measures depends only on the relation between what the subject is informed about the purpose of a measure and what the researcher chooses to infer from the subject's response. If the subject is advised that A is being measured, but the researcher uses the response to infer something about B, the direct measure of A is an indirect measure of B. In the case of nonreactive indirect measures, the subject is unaware that anything is being measured and, accordingly, there is no direct-measure interpretation of the subject's response. Only a small portion of the research considered in this article involves the use of nonreactive indirect measures.

An implicit attitude toward B may be indirectly indicated by a (direct) measure of evaluation of A, when A and B have some relation that predisposes the implicit influence. Possibly, the evaluative content of this implicit attitude may disagree with results from a direct measure of attitude toward B; such disagreement, referred to as a *dissociation* of implicit and explicit attitudes, is especially interesting and perhaps most dramatically indicates the value of the implicit attitude construct. Nevertheless, the occurrence of dissociation is not a necessary condition for identifying an attitude as implicit. More critical is that the effect of an attitudinal manipulation on an indirect measure may be reduced, eliminated, or reversed when subjects are made aware of the nature of the manipulation. Further treatment of this point appears below, in considering effects of attentional manipulations on expressions of implicit social cognition (i.e., on indirect measures).

Implicit Attitudes: Empirical Findings

Halo Effects

Thorndike (1920) named the halo effect, after noticing that personality ratings showed a tendency for positive characteristics to be associated with other positive characteristics more than they should be if experience is the only guide. Subsequently, the halo effect came to be regarded as the tendency for judgment of a novel attribute (A) of a person to be influenced by the value of an already known, but objectively irrelevant, attribute (B). In this case, the direct measure of evaluation of A implicitly expresses the attitude toward B. The attitude toward B is implicit, in present terms, when the subject does not identify the attitude toward B as the source of the evaluation of A.

In much halo effect research, physical attractiveness plays the role of the objectively irrelevant attribute that influences evaluative judgment on various other dimensions. For example, Dion, Berscheid, and Walster (1972) reported that attractive males and females are judged to be kinder, more interesting, more sociable, happier, stronger, of better character, and more likely to hold prestigious jobs. Similarly, Landy and Sigall (1974) found that essays attributed to a female student were judged by male students to be of higher quality when the stimulus materials included a photo that showed the author to be physically attractive, rather than unattractive. Downs and Lyons (1991) reported that defendant attractiveness was associated with judges levying smaller fines and lower bail levels in actual misdemeanor cases (although this relationship was weaker or absent in felony cases).

Physical attractiveness-based halo effects have been found to be greater for female than male targets (Bar-Tal & Saxe, 1976 ; Wallston & O'Leary, 1981). As Berscheid (1985) has pointed out, however, halo effect research is grossly imbalanced in its predominant use of male subjects and female targets–a strategy that may reflect both popular belief and evidence showing that physical attractiveness of an opposite-sex partner is more important for men than for women (Feingold, 1990). Alternately, the imbalance may indicate that attractiveness is a more potent implicit attitudinal cue component of females than males. The physical attractiveness-based halo effect has been replicated in subject populations of Black Americans (Cash & Duncan, 1984) and Japanese (Onodera & Miura, 1990), as well as across the life span (Adams & Crane, 1980). Eagly, Ashmore, Makhijani, and Longo (1991) concluded from their meta-analytic review that a component of the physical attractiveness halo effect can be explained by assuming that perceivers (consciously) expect physically attractive people to be socially adept; however, Eagly et al. also reported a sizeable halo effect for judgments on intellectual competence dimensions (mean effect size d = .46), for which there is no known basis in beliefs about abilities associated with physical attractiveness.

Nisbett and Wilson (1977b) demonstrated a reversal of the direction of the usual halo effect by presenting to subjects a male teacher who was coached to act in either an interpersonally warm or an interpersonally cold fashion. Subjects in the warm condition later judged the teacher to be more attractive in appearance, accent, and mannerisms than did subjects in the cold condition. Subjects denied that the teacher's likeableness influenced the other judgments. Quite the opposite, subjects in the cold condition incorrectly believed that the (unvaried) accent, mannerisms, and physical unattractiveness of the teacher had reduced the teacher's likeableness. The importance of this finding derives from the subject's inaccurate identification of the effect of one attribute on judgment of another, despite these being "attributes for which we generally assume we are capable of rendering independent assessments" (Nisbett & Wilson, 1977b, p. 250).

Despite the focus of much halo effect research on physical attractiveness as the cue, the range of potential halo effect cues is quite broad. For example, P. R. Wilson (1968) manipulated academic status by introducing a male stimulus person as a student, demonstrator, lecturer, senior lecturer, or professor, to different groups of observers at Cambridge University. The observers' subsequent memory estimates of the person's height were an increasing function of the status variable, with the student recalled to be shortest and the professor tallest. Another variant of the halo effect was demonstrated by Howard (1992) , who showed that evaluation of a product was positively influenced by first encountering it in an attractively gift-wrapped package rather than in a plain package. Still a different halo effect was reported by Frank and Gilovich (1988) , in showing that black color of sports uniforms is associated with both self-perceptions and observers' perceptions of aggressiveness. Another variant appeared in a controversial study by Peters and Ceci (1982) , who submitted manuscripts to journals that had previously published them, using fictitious authors' names and deliberately unimpressive

(fictitious) institutional affiliations. The subsequent rejection of most of these manuscripts has several possible explanations, but the majority of the published commentary that accompanied the Peters and Ceci article concluded that the institutional cue was a likely cause of negative evaluations.

As a general interpretation of halo effects, the present analysis supposes that the subject's learning that an unfamiliar target person possesses Attribute B tends to produce a diffuse positive or negative attitude (depending on the affective value of B) toward the target person; that attitude is then likely to generalize to any specific attribute (A) that the subject is asked to judge. The attitude toward B is said to operate implicitly when the subject does not notice that B is influencing the judgment of A. The halo effect is of great practical significance, being applied thousands of times a day just in the world of advertising. Advertisements typically set their products in contexts that contain attractive other objects, especially famous entertainers and physically attractive models. Because advertising audiences are aware that the advertiser is trying to influence attitude, there may be little likelihood of the audience misidentifying the source of positive reactions to the advertising message. But, as a reader of an earlier draft of this article reminded the authors, advertising audiences are frequently inattentive and, in that circumstance, may be susceptible to implicit effects of an extraneous attractiveness cue. A common situation in which halo effects are likely to be more generally potent is in managers' evaluations of workers; job-unrelated attractiveness may routinely affect such performance evaluations (e.g., Balzer & Sulsky, 1992 ; Cooper, 1981 ; Tsui & Barry, 1986).

Other Implicit Attitude Effects Mere exposure.

In a review that established the relationship between frequency of encounter and liking for a wide variety of stimuli in a wide variety of contexts, Zajonc (1968) named and established the mere exposure effect. For most of the time since, the mere exposure effect has been an enigma, a robust effect without a generally accepted explanation. However, several recent findings now indicate that the mere exposure effect is an implic it attitude effect. Bornstein and D'Agostino (1992) , pursuing a lead that came from Bornstein's (1989) meta-analytic review, established that mere exposure effects are strongest when conditions reduce subjects' memory for the effect-producing exposures. Under these circumstances, it appears that increased perceptual fluency of a repeatedly presented stimulus (that is, increased ease of its identification on re-exposure) is misattributed to liking, yielding a positive evaluation of the stimulus. [3] Some other results that indicate the same type of implicit effect are the effect of ease of comprehension on judgments of a statement's validity (Gilbert, 1991 [4] ; Sherman, Mackie, & Driscoll, 1990) and the effect of repeated encounters with a statement on its judged truth (Arkes, Boehm, & Xu, 1991 ; Begg, Armour, & Kerr, 1985 ; Hasher, Goldstein, & Toppino, 1977 ; Hawkins & Hoch, 1992).

Subliminal attitude conditioning.

Several researchers have sought to induce positive or negative attitudes to neutral stimuli by presenting a very briefly flashed pleasant or unpleasant stimulus just before presenting each clearly visible neutral stimulus. Subsequent affective judgments of the initially neutral stimuli provide a test of whether the affective value of the flashed stimulus transfers to the neutral stimulus. Although reliable procedures for obtaining this result are not yet well described, a few researchers have reported successful efforts (K. Edwards, 1990 ; Krosnick, Betz, Jussim, & Lynn, 1992 ; Niedenthal, 1990 ; Murphy & Zajonc, 1993). This procedure has occasionally been described as attitude conditioning; however, the "conditioning" label may be unwarranted because of both the small number of trials used and the reversal of the usual neutral-stimulus-first ordering of conditioning procedures. The result may better be understood as an implicit attitude effect, in which an attitude evoked by the first (briefly flashed) stimulus is mistakenly attributed to the second stimulus.

Instant attitudes?

An "instant attitude" can be defined as a near-immediate liking or dislike for a novel object on first encounter with it (see Fiske, 1982 , for a similar idea captured by the term *schema-triggered affect*). Although little studied in research, instant attitudes may be very common. For example, it is common–at least for the present authors–to experience a sense of almost immediate sympathy or antipathy for characters in drama, fiction, or news reports, and also with individual performers or teams in athletic competitions. In some cases, an explicit basis for the instant attitude may seem obvious to its possessor: there is apparent self—other similarity, or an

athletic team is from one's home town or university, or a fictional character is depicted in blatantly admirable or despicable terms. In other cases, the instant attitude may have no introspectively accessible basis and, therefore, likely merits identification as an implicit attitude.

Context effects in survey research.

Schwarz and Clore (1983) reported that quality-of-life judgments by telephone-interviewed subjects were greater for persons called on sunny days than for those who were called on rainy days. In a related study, Schwarz, Strack, and Mai (1991) asked groups of survey respondents about both their happiness with marriage and their happiness with life as a whole. When the marital question preceded the quality of life question, expressed life satisfaction increased for respondents who reported happy marriages, but decreased for those who reported unhappy marriages, as compared with the respondents who answered the life-as-a-whole question first. These effects may represent implicit influences of peripheral information on evaluative judgments, but may also be explainable as consequences of introspectively available knowledge. One means of supporting the interpretation of implicit influence is to show that drawing attention to the influential cue reduces the effect, which is precisely what was shown by Schwarz and Clore (1983) ; when subjects were asked to describe the weather early in the interview, the effect of weather on the later quality-of-life question was eliminated.

As just illustrated, a means of pressing the argument that the effects described in this section merit an "implicit" designation is to show that they weaken or disappear when subjects are made aware of the implicit attitudinal stimulus. The case for that conclusion is developed further in the section below, Attention as a Moderator of Implicit Cognition.

Implicit Self-Esteem

Studies done in the last few decades have established that a majority of almost any group of research subjects reports favorable judgments when asked to provide self-evaluations, including the self-evaluative responses to items on inventory measures of self-esteem. This very reliable result provides the basis for concluding that most people have a positive attitude toward self (Banaji & Prentice, 1994 ; Baumeister, 1982 ; Greenwald, 1980 ; Myers & Ridl, 1979 ; S. E. Taylor & Brown, 1988). Consequently, an expectable form of implicit attitude effect is that novel objects that are invested with an association to self should be positively evaluated. As in the case of the halo effect, it is assumed that the judgment of any previously unevaluated attribute (A) of the object is influenced by the (usually positive) attitude associated with self (Attribute B). The resulting positive evaluation of Attribute A is an implicit effect in the same sense that the halo effect is–it is assumed to occur without the subject's awareness of the influence and despite the objective irrelevance of the self-association to the subject's judgment

> *Implicit self-esteem* is the introspectively unidentified (or inaccurately identified) effect of the self-attitude on evaluation of self-associated and self-dissociated objects.

Three categories of effects lend themselves to interpretation in terms of implicit self-esteem. *Experimental* implicit self-esteem effects compare evaluations of novel stimuli that are arbitrarily associated with self or not; greater liking for self-associated objects is interpretable as an implicit self-esteem effect. *Naturally mediated* implicit self-esteem effects are similar, except that the association with self is made by means of object attributes that have a preexisting association to self. In a category that can be labeled *second-order* implicit self-esteem effects, the subject arrives at a judgment that has an inferential link to self-esteem; the self-esteem-consistent nature of this judgment can be regarded as an implicit manifestation of self-esteem.

Experimental Implicit Self-Esteem Effects Role playing in persuasion.

In experimental demonstrations of role-playing effects in persuasion, subjects are induced to present arguments supporting an arbitrarily assigned controversial position. Compared with control subjects who are exposed to similar arguments from an external source, role-playing subjects tend to credit greater validity to the novel arguments or to the position that those arguments support (e.g., Janis & King, 1954). This effect occurs both when subjects are induced to play a creative role in producing the arguments (e.g., Greenwald & Albert, 1968 ;

King & Janis, 1956) and when subjects merely judge the validity of novel issue-relevant arguments after accepting the role assignment (Greenwald, 1969). The interpretation of this result as an implicit self-esteem effect is based on assuming, first, that the subject's acceptance of the role-playing assignment creates a link of the assigned position to self (this is a psychological *unit formation,* in the sense of Heider, 1944) and, second, that positivity toward self generalizes to a positive evaluation of the assigned position.

Mere ownership.

In a procedure introduced by Feys (1991 ; see also Beggan, 1992), subjects first learn to discriminate four computer-displayed graphic icons (which represent the subject in a computerized game) from four others that represent the subject's opponent (the computer). When subjects subsequently judge all eight patterns for aesthetic attractiveness, the self-associated patterns receive higher ratings. The *mere ownership* label for this effect comes from an explanation suggested by Nuttin (1985 , 1987) for the name letter effect (discussed later in this section). A similar result that likely also merits an implicit self-esteem interpretation is the "instant endowment" effect described in several experiments by Kahneman, Knetsch, and Thaler (1990) . Those experiments established that the value attached to such objects as mugs, pens, and chocolate bars increased sharply as soon as the subject was given the object.

Minimal group effect.

In-group bias is the tendency to judge members of one's own group (in-group) more favorably than comparable persons who are members of another group (out-group). (This itself is a finding that could be included in the list, below, of naturally mediated implicit self-esteem effects.) Compelling demonstrations of a type first produced by Tajfel (1970) establish that in-group bias occurs even when care is taken to assure that in-group and out-group members are objectively similar in every respect other than group membership–for example, by allowing subjects to observe the random assignment of persons, themselves included, to groups within the experimental situation. In this form, in-group bias is identified as the minimal group effect. Recent discussions of in-group bias and minimal group effects can be found in Brewer (1979) , Gaertner, Mann, Dovidio, Murrell, and Pomare (1990) , Messick and Mackie (1989) , Mullen, Brown, and Smith (1992) , Ostrom and Sedikides (1992) , Tajfel and Turner (1986) , and Wilder and Shapiro (1984) . The minimal group effect indicates that shared group membership provides a sufficient link to self to permit implicit operation of self-esteem. Further evidence that minimal in-group cues may suffice as a basis for transfer of positive affect comes from a study by Perdue, Dovidio, Gurtman, and Tyler (1990) , who reported that in-group pronouns (*us* and *we*) served as affectively positive primes in a subliminal activation procedure.

Naturally Mediated Implicit Self-Esteem Effects Similarity-attraction.

Perceived similarity of opinions with another person has been demonstrated to be a powerful determinant of attraction (see Byrne, 1969 , for an overview of the series of studies that initially demonstrated this phenomenon). In typical similarity-attraction experiments, the subject makes judgments about a stranger (stimulus person) after having seen a sample of the stranger's responses to a series of opinion questions. Surreptitiously, the stranger's opinion responses have been constructed so as to define varying levels of similarity to the subject's own previously recorded opinion responses. The measure of attraction consists of summed responses to two judgment items, one indicating liking and the other a judgment of how enjoyable it would be to work together with the stranger. In a representative experiment by Byrne (1962) , 0—7 opinion responses by the stimulus person were manipulated to agree with those of the subject. Attraction was directly and strongly influenced by the number of these agreements. The interpretation of similarity-attraction findings as implicit self-esteem effects depends on the assumption that increasing numbers of opinions shared with the stranger constitute increasing strength of the link of the stranger to self. [5]

Unlike "experimental" implicit self-esteem effects, for the similarity-attraction effect, the link of the novel stimulus object (stranger) to self is mediated by attributes of this object that have a preexisting association to self. An interesting variant of the similarity-attraction effect was created by Finch and Cialdini (1989) , who led subjects to believe that they either shared or did not share a birthday (same day and month) with a notorious historical character, Rasputin. Subjects were more lenient in judging the deeds of Rasputin when the shared birthdate created a link to self. Subsequently, Prentice and Miller (1992) used Finch and Cialdini's shared

birthday technique to increase bonds between the subject and another participant in a prisoner's dilemma negotiating situation. Subjects who believed that they shared a birthdate with the other player cooperated significantly more than did subjects who were not provided this (false) information.

Cognitive responses to persuasion.

Novel persuasive arguments tend to be accepted as valid to the extent that their conclusions agree with one's existing opinions. This effect was established in the late 1960s (e.g., Cullen, 1968 ; Greenwald, 1968). A strong demonstration of the effect appeared in a study by Lord, Ross, and Lepper (1979) , in which students in favor of or opposed to capital punishment examined two sets of evidence, one set supporting capital punishment and one opposing. Students' judgments of the relative convincingness of the two sets of evidence were strongly shaped by their preexisting opinions. Each opinion group found that the evidence, on balance, favored their own position. This is due to the combinations of tendencies to be attracted to the arguments supporting one's own position and to be repelled by those supporting the opposing position. It is therefore apparent that the cognitive response effect is a close relative of the similarity-attraction effect. Instead of opinions being credited to a hypothetical stranger (the similarity-attraction effect), they are included in a persuasive communication. Instead of the effect being greater liking or disliking for the stranger with agreeable or disagreeable opinions, it is greater liking for the communication. The tendency to reject a communication that does not support one's own views is shown especially strongly in the form of the hostile media phenomenon (Vallone, Ross, & Lepper, 1985), in which each of two politically opposed groups interpreted the same (presumably unbiased) news media report of a terrorist attack as objectionable because of its lack of support for their own position.

Postdecisional spreading of choice alternatives.

In this procedure, subjects are asked to rate several members of a product category (such as music recordings), then to choose to receive one of two that were initially rated close to one another, and finally to rate all of the alternatives again. In comparison with control subjects who make no choice, choice subjects show an increase in rated attractiveness of the chosen alternative relative to the nonchosen one (Festinger & Walster, 1964 ; Steele, Spencer, & Lynch, 1992). This, again, is a pattern of increased positivity toward a stimulus that has a link with self. In this case the link to self is established by the subject's choosing to receive the object. [6] A powerful variant of this effect was demonstrated by Langer (1975) , who sold lottery tickets to office workers for $1 and later approached them to sell the ticket to another buyer. Langer varied whether, on the initial purchase, subjects had been allowed to choose their own ticket or were merely handed one. The dependent variable was the price the subject was willing to accept to sell the ticket. In the "choice" condition, the average price asked for the ticket was $8.67, compared with $1.96 in the "no-choice" condition. Here, a link of the item to self, however created, increased monetary value above the $1 purchase price (cf. the mere ownership effect, above), but the increase was much greater with the stronger link produced by active choice.

Liking for name letters.

When subjects are asked to choose a preferred letter from each of several pairs consisting of one letter from their name and one not (with subjects not being alerted to this aspect of the pairs' composition), they tend reliably to prefer letters from their name (Nuttin, 1985). Of several possible explanations for the effect, the most successful has been one based on ego attachment; that is, the preference for letters in one's name reflects a positive attitude to self (Hoorens, 1990 ; M. M. S. Johnson, 1986 ; Nuttin, 1985). Using a letter-attractiveness rating task to test this name letter effect, M. M. S. Johnson (1986) obtained the previously unpublished data shown in Figure 1 . These results reveal a statistically significant (but weak) name letter effect and a considerably stronger *initial letter effect,* that is, enhanced liking for letters that constitute the initials of one's name. Because initial letters are linked more strongly to self than are other name letters, M. M. S. Johnson's finding strengthens the interpretation of the name letter effect as a manifestation of implicit self-esteem. [7]

Second-Order Implicit Self-Esteem Effects Self-positivity in judgment.

There have been many research demonstrations of judgment biases that work to cast a positive light on the self. These findings of self-positivity biases are not reviewed here because earlier works (cited in the next paragraph) have documented several of them thoroughly. Self-positivity biases have special significance for the present

analysis because they do more than express self-esteem implicitly; they also provide support for self-esteem and are, therefore, second-order implicit self-esteem effects. [8]

The tendencies to accept responsibility for desired outcomes while finding external causes for undesired outcomes and, more generally, to construct judgments and revise memory in a fashion consistent with a positive self-image were identified by Greenwald (1980) as symptoms of one of a trio of cognitive biases (egocentricity, beneffectance, and cognitive conservatism) of the self. These biases, which share characteristics with the operation of a totalitarian society's propaganda apparatus, were described as adaptive, functioning to preserve the integrity of the self as a knowledge organization. S. E. Taylor and Brown (1988) more recently reviewed evidence for self-positive illusions, strongly documenting the case for the adaptive functions of these biases. And, on the basis of their analyses of cognitive processes of depressive patients, Beck (1979) , Scheier and Carver (1992) , and Seligman (1991) have indicated both the prominence and the adaptiveness of self-positivity in normal (nondepressive) cognitive functioning.

One example of self-positivity in judgments warrants mention here because of its special relevance to this article's focus (in a section to follow) on stereotypes and prejudices. This is Crosby's (1984) finding that members of disadvantaged groups, even though viewing other members of their own group as targets of discrimination, tend not to see themselves as having been so victimized. This phenomenon, which can be linked to just-world illusions (Lerner, 1980), has been confirmed in various samples in the United States (Crosby, Pufall, Snyder, O'Connell, & Whalen, 1989), Francophones in Quebec (Guimond & Dube-Simard, 1983), and Haitian and Indian immigrants in Canada (D. M. Taylor, Wright, Moghaddam, & Lalonde, 1990).

Implicit affiliation and rejection.

Festinger's (1954) theory of social comparison was based on a supposition that people compare themselves with others in order to gain an accurate assessment of their opinions and abilities (see also Suls & Wills, 1991). In more recent variations of the theory, especially Tesser's self-evaluation maintenance theory (Tesser, 1988 ; Tesser & Campbell, 1983) and Wills's (1981) analysis of downward comparison, the potential for social comparisons to occur in the service of self-enhancement has been emphasized. This modern view of social comparison describes self-esteem as influenced by the interrelationship of perceived relation of self to other (person or group) and evaluation of other. High self-esteem accompanies either association to positively valued others or dissociation from negatively valued others.

This general principle that affiliations and rejections are in the service of self-esteem is supported by the previously described (naturally mediated) implicit self-esteem phenomena of similarity-attraction, in-group bias, and biased cognitive response to persuasion, all of which involve effects of variations of perceived relation to another (person or communication) on attraction. It equally encompasses several well-established phenomena in which perceived relation to other is the dependent variable, rather than a manipulated independent variable. Several studies have varied the other's good or bad fortune and observed the consequences for both perceived association with and liking for the other. Basking in reflected glory (Cialdini et al., 1976) is perceiving a positive association with (and presumably admiring) a fortunate or successful other, whereas results characterized as showing downward comparison (Suls, 1991 ; Wills, 1981) involve distancing oneself from and being critical of a disadvantaged, unsuccessful, or unfortunate other.

Phenomena of implicit affiliation and rejection also reveal second-order implicit self-esteem, because either association with an attractive other or distancing from an unattractive other should bolster self-esteem. Support for this interpretation comes from studies in which tendencies to bask in reflected glory or to engage in self-enhancing downward comparison were increased by procedures, such as failing at an ego-involving task or learning that one possesses undesirable personality traits, that temporarily lowered subjects' self-esteem (e.g., Cialdini & Richardson, 1980 ; Crocker, Thompson, McGraw, & Ingerman, 1987). Further support comes from studies in which subjects are shown to be biased toward embracing, as self-descriptive, whichever of two opposing traits (extraversion or introversion) is presented as being predictive of personal success (Kunda & Sanitioso, 1989 ; Sanitioso, Kunda, & Fong, 1990).

Displaced self-esteem.

Displaced self-esteem (which has not previously been identified as a distinct class of findings) is a relative of well-known reciprocity and ingratiation effects that involve reciprocal exchange of gifts, favors, or praise (Jones, 1964 ; Regan, 1971 ; see Cialdini, 1993 , for an overview of reciprocity effects). Displaced self-esteem starts with an act of received praise, but thereafter deviates from reciprocal exchange phenomena in that the subsequent return praise is perceived as an independent judgment that has no relation to the received praise. Displaced self-esteem appears to have been central to a well-known finding by Aronson and Linder (1965) . In that experiment, a stimulus person who switched from criticizing to praising the subject was liked better than one who continuously praised the subject, even though the latter gave more total praise. Aronson and Linder proposed that the switch from criticism to praise prompted greater liking because it made the praiser seem more discriminating in interpersonal judgments. The present (related) interpretation is that the switch from criticism to praise provided a more potent boost to the subject's self-esteem than did the continuous praise, with the return liking then being a second-order implicit self-esteem effect; that is, it further supported the subject's self-esteem by boosting the credentials of the person who offered praise.

In as-yet-unpublished studies conducted at the University of Washington, a strong displaced self-esteem effect has been identified in a natural setting that routinely calls for two parties to take turns evaluating each other–the end-of-course student ratings of instructors (Greenwald, 1992b). In analyses both within and across course sections, higher grades of student by instructor have been strongly associated with higher ratings of instructor by student. The positive evaluation of an instructor by a student who expects to receive a high grade appears to provide support, implicitly, for the student's self-esteem.

Alternative Interpretations of Implicit Self-Esteem Effects

Each of the implicit self-esteem effects just reviewed is well established in multiple studies. Furthermore, several of these findings meet established criteria of being statistically strong effects (that is, they correspond to product—moment correlations of at least .50, or to treatment mean differences on a dependent measure of at least 0.8 standard deviations; Cohen, 1988). Also, these effects come from broadly diverse areas of research, including attitude change (role playing and cognitive response effects), decision making (postdecision spreading of alternatives), group process (minimal group effect), interpersonal attraction (similarity-attraction, implicit affiliation or rejection, and displaced self-esteem effects), personality (self-positivity biases), and social cognition (name or initial letter and mere ownership effects). Not surprisingly, this diverse collection of findings has a correspondingly diverse collection of existing theoretical interpretations, including reinforcement theory (similarity-attraction), cognitive balance or consistency theory (cognitive response and similarity-attraction), cognitive dissonance theory (role playing and postdecision spreading of choice alternatives), and social comparison theory (downward comparison).

Interpreting various findings in terms of implicit self-esteem does not require abandoning their other, existing theoretical interpretations. In some cases, the existing interpretations describe plausible mechanisms by which implicit self-esteem may operate. [9] However, the implicit self-esteem interpretation goes beyond other interpretations in predicting that the various effects should be moderated by self-esteem. That is, there should be stronger implicit manifestations of self-esteem for subjects who are higher in self-esteem. [10]

Unfortunately, the hypothesized moderating role of self-esteem cannot be tested confidently with existing self-esteem measures. The problem is that standard self-esteem inventories (e.g., Helmreich, Stapp, & Ervin, 1974 ; Rosenberg, 1965) use a direct measurement strategy that assesses an introspectively accessible representation of self-esteem. The hypothesized moderating role of self-esteem in implicit self-esteem effects is expected to be effectively testable only with the aid of a procedure for (indirectly) measuring implicit self-esteem. Such a measure does not now exist. [11] Below, the state of available technology for indirect measurement of implicit social—cognitive constructs is discussed.

The Nature of Evidence for Implicit Attitudes and Implicit Self-Esteem

In the absence of individual-difference measures for implicit attitudes, supporting evidence for that construct has focused on situations in which it can be assumed that subject samples are approximately homogeneous in the content of their implicit attitudes. For example, the interpretation of physical attractiveness halo effects as implicit attitude effects requires the assumption that physical attractiveness is positively evaluated by a large

majority of most research samples. Similarly, empirical demonstrations of implicit self-esteem require the assumption that positive self-regard is widespread within research samples. In this respect, research on implicit social cognition is on a similar footing with research on other forms of implicit cognition. In implicit memory research, for example, it is assumed that exposure to a given set of experimental materials produces an approximately uniform effect across subjects in establishing traces that can later influence performance on indirect measures.

Implicit Stereotypes

A stereotype is a socially shared set of beliefs about traits that are characteristic of members of a social category. Whereas an attitude implies a consistent evaluative response to its object, a stereotype may encompass beliefs with widely diverging evaluative implications. For example, the stereotype of members of a certain group (e.g., cheerleaders) may simultaneously include the traits of being physically attractive (positive) and unintelligent (negative). Stereotypes guide judgment and action to the extent that a person acts toward another as if the other possesses traits included in the stereotype.

As was the case for attitudes, scholarly definitions of stereotypes have generally not specified their conscious or unconscious operation. This can be seen, as it was for the attitude construct, by examining a list of influential

> A stereotype is a fixed impression, which conforms very little to the fact it pretends to represent, and results from our defining first and observing second. (Katz & Braly, 1935, p. 181)
>
> A stereotype is an exaggerated belief associated with a category. (Allport, 1954, p. 191)
>
> A categorical response, i.e., membership is sufficient to evoke the judgment that the stimulus person possesses all the attributes belonging to that category. (Secord, 1959, p. 309)
>
> An ethnic stereotype is a generalization made about an ethnic group, concerning a trait attribution, which is considered to be unjustified by an observer. (Brigham, 1971, p. 13)
>
> A set of beliefs about the personal attributes of a group of people. (Ashmore & Del Boca, 1981, p. 16)
>
> In stereotyping, the individual: (1) categorizes other individuals, usually on the basis of highly visible characteristics such as sex or race; (2) attributes a set of characteristics to all members of that category; and (3) attributes that set of characteristics to any individual member of that category. (Snyder, 1981, p. 415)
>
> A cognitive structure that contains the perceiver's knowledge, beliefs, and expectancies about some human group. (Hamilton & Trolier, 1986, p. 133)

As in the case of attitudes, it is useful to examine the extent to which stereotypes operate implicitly, outside of conscious cognition.

> *Implicit stereotypes* are the introspectively unidentified (or inaccurately identified) traces of past experience that mediate attributions of qualities to members of a social category.

Although research on stereotypes has often used direct or explicit measures (see Judd & Park, 1993), there are also substantial research programs on stereotypes that use indirect measures–ones in which a stigmatizing feature with which a stereotype is associated (e.g., weight, race, or gender) is peripheral to the respondent's judgment task (e.g., Darley & Gross, 1983) or in which the purpose of investigation is otherwise disguised (e.g., Hamilton & Gifford, 1976). Crosby, Bromley, and Saxe (1980) were able to locate enough research using indirect measures of prejudicial stereotypes to conclude that "anti-Black sentiments are much more prevalent among White Americans than the survey data [i.e., direct or explicit measures of stereotypes] lead one to expect." Although the use of indirect measures in these studies often reflects the researchers' intent to avoid

intrusion of unwanted demand or impression-management artifacts (which would plausibly suppress accurate expressions of conscious stereotypes), some of the research that is summarized just below was designed specifically to investigate unconscious operation of stereotypes. These studies suggest that stereotypes are often expressed implicitly in the behavior of persons who explicitly disavow the stereotype. The next two sections focus attention on race and gender stereotypes because these, having been much more heavily investigated than other stereotypes, have provided the most persuasive evidence for implicit stereotyping.

Implicit Race Stereotyping

Numerous findings have established automatic operation of stereotypes. Gaertner and McLaughlin (1983) presented subjects with pairs of letter strings, requesting a yes judgment if both were words, and no otherwise. Using speed of yes responses to measure strength of existing associations between the two words in a pair, they found that White subjects responded reliably faster to white-positive word pairs than to black-positive pairs (e.g., *white-smart* vs. *black-smart*). This difference did not emerge on judgments of negative traits (e.g., *white-lazy* vs. *black-lazy*). These results occurred similarly for subjects who scored high and for ones who scored low on a direct (i.e., standard self-report) measure of race prejudice. In a related study, Dovidio et al. (1986) used the procedure of presenting a prime (*black* or *white*) followed by a target (a positive or negative trait) and asking subjects to judge if the target trait could "ever be true" or was "always false" of the prime category. Again, subjects responded reliably faster to positive traits that followed the prime *white* than *black,* and in this study they also responded faster to negative traits that followed the prime *black* than *white.* These results were interpreted by Gaertner and Dovidio (1986) as evidence for *aversive racism,* which they defined as a conflict "between feelings and beliefs associated with a sincerely egalitarian value system and unacknowledged negative feelings and beliefs about Blacks" (p. 62).

Automatic operation of stereotypes provides the basis for implicit stereotyping. Devine (1989) reported that, after being subliminally exposed to a series of words, 80% of which were stereotypically associated with Black Americans (e.g., poor, jazz, slavery, Harlem, and busing), White subjects, in an ostensibly unrelated second task, judged a race-unspecified male target to be more hostile than did subjects for whom only 20% of the words had the stereotype association. Again, as was the case in the Gaertner and McLaughlin (1983) study, this result occurred equally for subjects who scored high and low on a direct measure of prejudice. Because Devine's subjects may have imagined a White target on the second task, it is possible that this result was due to some consequence of the priming procedure other than its activation of a Black stereotype (for example, it might have activated hostility). Nevertheless, Devine's study pioneered in identifying an implicit social cognition effect and in suggesting the role of automatic (unconscious) processes in stereotyping and prejudice. Gilbert and Hixon (1991) showed that a racial stereotype, presumably activated by including an Asian woman in a videotaped sequence seen by subjects, influenced subsequent word-fragment completions (a type of measure used frequently in implicit memory research). In the context of considering phenomena of implicit social cognition, this finding, along with Devine's, might be well described as revealing *implicit racism.*

Implicit Gender Stereotyping

Some of the methods used in studies of automatic components of ethnic and race stereotypes have begun to appear in studies of gender stereotypes (e.g., Jamieson & Zanna, 1989 ; Klinger & Beall, 1992 ; Paulhus, Martin, & Murphy, 1992). Also, a well-established line of research based on the finding that essays were judged more favorably when attributed to authors with male rather than female names (Goldberg, 1968) can be interpreted as indirectly assessing a gender stereotype that associates men with greater achievement than women. [12] The repeated finding of lower ratings of the same achievements when credited to women rather than to men might also be interpreted as reflecting greater attitude positivity toward men. However, because of several findings that show greater attitudinal positivity toward women than men (Eagly & Mladinic, 1989 ; Eagly, Mladinic, & Otto, 1991 ; also unpublished findings obtained by the present authors), the stereotype interpretation appears preferable.

Stereotypical Gender Bias in Fame Judgments

The stereotypic association of gender with achievement provided a basis for the present authors to adapt methods of implicit memory research to investigate implicit stereotypes. Jacoby, Kelley, Brown, and Jasechko

(1989 ; see also Jacoby & Kelley, 1987) demonstrated the operation of implicit memory in fame judgments. In their research, on Day 1 of a two-session experiment subjects read a list that contained names of both famous and nonfamous people. On Day 2, 24 hr later, the same subjects were presented with a list containing old (previously seen on Day 1) and new (unseen on Day 1) nonfamous names, interspersed with old and new famous names. Subjects judged each name on the Day 2 list in response to the question: Is this person famous? (to be answered yes or no). Jacoby et al. hypothesized that, although episodic memory (i.e., explicit, conscious recollection) for the nonfamous names would fade over the 24-hr delay, residual perceived familiarity due to the prior exposure could lead to erroneous judgments of fame. That is, subjects might misattribute the familiarity of a name to fame and, as a consequence, would judge more old (than new) nonfamous names as famous. As they predicted, Jacoby et al. (1989) found a higher false-alarm rate for old nonfamous names than for new ones. This effect, described by Jacoby et al. (1989) as making nonfamous names "become famous overnight," indicates a potent unconscious influence of memory.

Adapting the Jacoby et al. (1989) procedure, Banaji and Greenwald (in press) used equal numbers of male and female names (Jacoby et al., 1989 , had used a large majority of male names) and varied the gender of nonfamous names by attaching a female or male first name to a common last name (e.g., Peter Walker or Susan Walker). Data from each subject's judgments for each of the four within-subject conditions (old male, new male, old female, and new female) were reduced to a hit rate (proportion of famous names correctly judged famous) and a false-alarm rate (proportion of nonfamous names mistakenly judged famous). Signal-detection analysis (Green & Swets, 1966) permitted decomposition of the hit and false-alarm data into measures of sensitivity (*d* ') to the stimulus variable (name fame in this case), and threshold or criterion (β) for assigning the judgment. The analyses used these two derived measures, replacing β with its logarithm, which is better suited for statistical analyses because of the greater approximation to normality of its distribution.

The main findings are graphed in Figure 2 . As previously mentioned, hit rates (correct identifications of famous names) were greater for male than female names; false alarms were greater for old (previously seen) nonfamous names than for new ones; and the boost in false-alarm rates for old names was greater for male than female names. The signal-detection analysis indicated significantly greater sensitivity to the fame variation (*d* ') for new (i.e., not previously seen) male names than for any other category. However, the strongest results from the signal-detection analysis were on criterion (indexed by log beta). The criterion for fame judgments was lower (i.e., names were more readily judged famous) for old than new names and for male than female names. In addition to these two main effects on criterion, there was a significant interaction effect such that the familiarization procedure (presentation on Day 1) had a greater effect of lowering criterion for male than for female names.

These results clearly show that an ambiguous stimulus–a name that elicits a sense of familiarity–is more likely to produce a mistaken judgment of fame if it is male rather than female. This adaptation of an implicit memory experimental procedure provides clear evidence for implicit gender stereotypes that associate male gender, more than female gender, with achievement. Because names serve as flexible vehicles for communicating social categories (especially ethnic ones), and because judgments other than the fame judgment can be used as indirect measures of traits other than achievement, the implicit memory procedure is potentially adaptable to examine a broad variety of stereotypes (cf. Klinger & Beall, 1992). By using a large set of names, such procedures are protected against the threat that stereotyping effects are due to idiosyncratic properties of the particular names selected as representative of larger categories (see Kasof, 1993).

Stereotypical Gender Bias in Judgments of Dependence and Aggression

Higgins, Rholes, and Jones (1977 ; also Srull & Wyer, 1979) demonstrated that the presentation of trait-category information in one context can influence judgments of an ambiguously described target person in an unrelated context. Banaji, Hardin, and Rothman (1993) used a variant of Higgins et al.'s procedure to extend the concept of implicit stereotyping. On the basis of documented stereotypes that link males to the trait of aggressiveness and females to the trait of dependence (see Basow, 1986 ; Broverman, Vogel, Broverman, Clarkson, & Rosenkrantz, 1972), Banaji et al. (1993) asked whether activating one of these traits would differentially influence judgments about male and female target persons. The prediction was that the activated trait might influence judgments only for the category (male or female) for which the trait was stereotypically appropriate, thereby demonstrating implicit operation of the stereotype.

In Banaji et al.'s (1993) dependence experiment, subjects who were exposed to primes that described dependent behaviors judged a female target (Donna) as more dependent than did subjects who rated the same target after exposure to neutral primes. However, subjects exposed to exactly the same dependence primes judged a male target (Donald) as less dependent than did subjects who rated the target after exposure to neutral primes. In the aggression experiment, subjects who were exposed to primes that described aggressive behaviors judged the male target as more aggressive than subjects who rated the same target after exposure to neutral primes. When judging a female target, previous exposure to the same aggression primes did not affect judgment. A third experiment replicated this pattern of data for dependence ratings (i.e., a more extreme judgment of a female target after exposure to dependence than neutral primes, but no such difference for a male target) and also showed no relationship between such judgments and explicit memory for the primes.

These experiments demonstrated that a gender stereotype moderates, and may be a necessary precondition of, the Higgins et al. (1977) trait-priming effect. Like the previous demonstration of gender bias in fame judgments, this result involves implicit social cognition; it occurs without the subject being consciously aware of an influence of recent experience (name exposure and trait exposure, respectively). At the same time, these effects reveal gender stereotypes because they occur selectively when the information content of recent experience stereotypically fits the social category (gender) of the target of judgment. The lack of subject sex effects in both procedures indicates that the stereotypes they reflect are culturally shared among both men and women, rather than being stereotypes of an out-group by an in-group (cf. Jost & Banaji, 1994).

Attention as a Moderator of Implicit Cognition

Decisions that affect people–for example, personnel evaluations and admissions decisions–comprise a large and very important class of situations in which implicit cognition can intrude on deliberate judgment, with the result of producing unintended discrimination. The decision maker who intends to maintain a nonprejudicial course in these judgments may have little basis for knowing whether or how a specific cue is implicitly intruding on judgment. Consequently, decision makers could usefully be acquainted with general strategies for reducing implicit cognitive effects, even when they are ignorant of the precise form that those effects may take.

Existing research indicates an important role for attention in general strategies for reducing undesired implicit influences on judgment. The general principle is that *attentional focus attenuates weak automatic influences on judgment.* This principle can be observed in operation in settings in which the weak automatic influence produces both desired and undesired effects.

Effect of Attention on Weak Automatic Influences That Produce Desired Responses

In perception research, Wyatt and Campbell (1951 ; see also Bruner & Potter, 1964) demonstrated that asking subjects to generate hypotheses about the identity of a blurred picture reduced their ability to identify the picture as it was gradually brought into focus. T. D. Wilson and Schooler (1991 ; see also T. D. Wilson, Dunn, Kraft, & Lisle, 1989) reported that subjects who were asked to introspect on their reactions to several brands of jelly subsequently rank ordered those brands in a fashion more discrepant from expert rankings than did subjects not given the opportunity to introspect.

In research on memory, Schooler and Engstler-Schooler (1990) found that, after attempting to verbally describe the memory image of a previously seen face, subjects showed impaired ability to choose that face from an array that included several foils. Graf and Mandler (1984 ; see also Overson & Mandler, 1987) found that deliberate attempts to complete word stems with words recently seen (but poorly attended) resulted in fewer correct responses than a comparison condition in which subjects simply attempted to provide completions. Research on hypermnesia has demonstrated that subjects will sometimes show improved recall after relaxing their efforts to retrieve, compared with sustained retrieval efforts (Erdelyi & Kleinbard, 1978 ; Madigan, 1976). In other words, when memory traces are weak, active effort to retrieve (using direct measures) may interfere with retrieval compared with more relaxed efforts that approximate indirect measurement procedures.

For all of these results, it is plausible that subjects' attentional efforts disrupted the influence of weak cues–either perceptual cues or memory traces–that might otherwise have guided judgment. Marcel (1983) observed that subliminal semantic activation was more evident among subjects who took a relaxed approach to their

experimental task than among those who tried effortfully to extract information from stimuli that were degraded by backward masking. Mandler (in press) concluded that the "conscious/deliberate attempt to retrieve... subliminally presented material... seems to interfere with access." In other words, for detecting effects of weak stimuli, direct measures (which focus attention on task stimuli) can be less sensitive than indirect measures.

Effect of Attention on Automatic Influences That Produce Undesired Responses Distraction increases implicit effects.

In the social cognition domain, the following several findings have demonstrated that decreased attention, due to distraction or time pressure, results in increased implicit effects of cues that are peripheral to the subject's task. Kruglanski and Freund (1983) found that imposing time pressure on a judgment task (thereby reducing attentional resources available for the task) increased the level of ethnic stereotyping in subjects' judgments. Similar findings involving gender stereotyping were obtained by Jamieson and Zanna (1989) and Pratto and Bargh (1991) . Gilbert and Hixon (1991) used a distracting memory load to reduce attentional resources available for a word-completion task that was used as an indicator of ethnic stereotyping and found increased evidence of stereotyping on the indirect measure. (Although Gilbert and Hixon reported, as just described, that stereotype expression was increased by distraction, they also reported that stereotype activation was reduced by distraction.) Paulhus and Levitt (1987) placed affect-arousing words adjacent to trait words that were being judged as self-descriptive or not and found that this procedure increased self-positive responding (endorsement of positive traits or rejection of negative ones). One interpretation of Paulhus and Levitt's finding is that the affect-arousing words diverted attentional resources from the judgment task, allowing implicit self-esteem to be a stronger influence on judgment than in the control (nonarousing adjacent words) condition.

Attention to source of implicit effect reduces the effect.

Several studies have demonstrated that when attention is focused on the source of an implicit effect that interferes with judgment, that interference is reduced or eliminated (and sometimes even reversed). For example, consistent with the expectation of a physical attractiveness halo effect, it has been found that subjects award lighter sentences to attractive criminal defendants (Efran, 1974 ; Sigall & Ostrove, 1975). However, when the crime was related to attractiveness (a swindle), Sigall and Ostrove observed a reversal that was presumably related to subjects' increased focus on the defendant's attractiveness and consideration of its relation to the crime, that is, physically attractive defendants received harsher sentences. Schwarz and Clore's (1983) finding that quality-of-life judgments were affected by weather was eliminated when subjects' attention was focused on the irrelevant cue by first asking them to describe the weather. These studies show that a cue that redirects attention to the source of a possible implicit effect can produce a reduction or reversal of that implicit effect. Research by Fiske and Neuberg (1990 ; Fiske, 1993 ; Neuberg, 1989) has shown similarly that increased attention (in their case, induced by giving subjects a goal of forming accurate impressions) increases the use of individuating information, rather than category-based stereotypes, in judging stimulus persons.

Recall of implicit cue decreases implicit effect.

In the memory domain, implicit cognitive effects have been found either to be weakened or reversed for subjects who could recall the stimuli that ordinarily produce those effects. Jacoby et al. (1989) showed that the false fame effect was reduced when the initial list of nonfamous names was well enough attended so that subjects would recognize nonfamous names as having been encountered earlier in the experiment. E. R. Smith, Stewart, and Buttram (1992) found that facilitative effects of repeatedly judging the same stimulus were larger when the stimulus was not recognized as having been judged previously (7 days earlier) than when it was recognized. In a review of much research showing that mere exposure to visual objects increases their judged attractiveness, Bornstein (1992) observed that this mere exposure effect was increased by procedures that decreased memorability of prior exposures. That observation was subsequently bolstered by an experimental test in which memorability of exposures was reduced by using very brief presentations (Bornstein & D'Agostino, 1992). Lombardi, Higgins, and Bargh (1987 ; see also Strack, Schwarz, Bless, Kübler, & Wänke, 1993) reported that effects of a priming procedure on trait inferences from an ambiguous description were reversed among subjects who could recall some of the priming stimuli. Relatedly, Martin, Seta, and Crelia (1990) found that this reversal of priming for a recallable priming manipulation was itself undone when subjects had a distracting task

concurrent with the test of the priming effect, indicating that the reversal observed by Lombardi et al. (1987) depended on the availability of attentional resources.

Exceptions.

It is clear from existing literature that attending to or recalling a cue does not invariably reduce its cognitive impact (see Martin & Achee, 1992 ; Petty & Wegener, 1993 ; Schwarz & Bless, 1992 ; Strack, 1992). In an unpublished replication of Johnson's initial letter effect (see Figure 1) performed by the present authors, subjects were asked to provide their names in advance, and it was mentioned that the letters they were about to rate included ones that were in their names. The authors expected that drawing attention to this cue would reduce its effect, but it did not; rather, the initial letter effect was undiminished. In retrospect, it was plausible that the effect was not altered because subjects had no reason to suppress this particular implicit self-esteem effect. Presumably, one should even expect augmentation of an implicit effect if the direction of the effect is consistent with a conscious motive. For example, consider a hypothetical variant on the initial letter effect experiment in which students are asked to rate attractiveness of various colors. If, in advance of the rating, the experimenter points out that the university colors are among those being rated, those colors might be rated as more attractive than in a control condition in which attention is not drawn to the association.

Conclusion.

The findings summarized in this section share the point that conscious attentional effort can weaken the influence of a current or previous cue on performance. However, the method of weakening likely depends on whether, at the time of the performance measure, the cue is or is not clearly cognized. When attentional effort is directed to a weak stimulus or memory trace, the reduced effectiveness of that cue is likely due to conscious strategies overriding and interfering with automatic cognitive effects. This interpretation is very reminiscent of a principle derived from learning-behavior theories of the 1940s and 1950s–that increases in motivation or drive amplify stronger habits relative to weaker ones (e.g., Spence, 1956). On the other hand, when the cue in question is cognized clearly, reduction of its implicit effect likely occurs because (and only to the extent that) the judge can anticipate and compensate for the event's possible influence. In this case, reversals of potential implicit effects may occur when the judge overcompensates for the influence effect, perhaps because the judge overestimates its magnitude or is overly zealous in seeking to avoid any appearance of having been influenced.

Appraisal of Strategies for Avoiding Unintended Discrimination

Over the last several decades, concern about the prevalence of social discrimination in employment, education, and other public settings has led to the proposal and implementation of antidiscrimination strategies that fall into three categories: *Blinding* denies a decision maker access to potentially biasing information. Exactly the opposite of blinding, *consciousness raising* encourages the decision maker to have heightened awareness of potential cues that could elicit discrimination. The third strategy, *affirmative action,* differs from the other two in having a deliberate compensatory component: An attribute that is known to be responsible for adverse discrimination is treated instead as if it were a positive qualification for the decision in question. The treatment here makes no attempt to acknowledge the substantial existing research on these three strategies, instead considering their possible effectiveness in situations in which discrimination might be based on implicit attitudes (for example, based on halo effects or prejudices) or implicit stereotypes.

Blinding.

In principle, blinding appears to qualify as a foolproof method of avoiding unintended discrimination. If a stigmatizing attribute is unavailable to a decision maker, it would seem that it could not possibly influence judgment. The value of blinding appears to be confirmed by experimental tests that permit clean manipulation of presence versus absence of a potentially stigmatizing attribute, independent of all other stimulus variation (see, e.g., the review by Eagly, Makhijani, & Klonsky, 1992). However, because natural possession versus nonpossession of almost any socially stigmatizing attribute tends to be correlated with possession of other characteristics that are not (and often cannot be) removed by blinding, effective blinding of the sort producible in experiments is often not achievable in practice.

Consciousness raising.

Research on the role of attention in weakening the effects of implicit cognition (reviewed above) supports consciousness raising as a strategy for avoiding unintended discrimination. That is, when a decision maker is aware of the source and nature of a bias in judgment, that bias may effectively be anticipated and avoided. Consciousness raising may also have some value in attenuating implicit bias when the source of implicit bias is not properly identified, as suggested by findings that attentional effort reduces effects of weak cues.

Affirmative action.

The controversiality of affirmative action is well captured by its frequently being described as "reverse discrimination." To discriminate in favor of a disadvantaged group may seem a defect even to those who desire only not to discriminate against them. At the same time, research on implicit social cognition provides several bases for concluding that unintended discrimination can be avoided only by deliberately applying compensatory strategies. Affirmative action strategies may therefore be deemed suitable not only as compensation to a stigmatized group for past explicit discrimination by others who intended to discriminate against them, but also as compensation for past, present, and likely future *implicit* discrimination by persons who have no intent to discriminate. That is, in addition to (or instead of) their interpretation as "reverse discrimination," affirmative action strategies might be understood as strategies for *reversal* of discrimination.

The contrast of affirmative action with blinding can be considered in the context of a case reported by Allmendinger and Hackman (1993). American symphony orchestras have a long tradition of predominant male membership, and women instrumentalists have historically not found equal opportunity in obtaining experience that can allow them to compete effectively with men. Being alert to possible discrimination, orchestras now routinely have candidates for vacant positions perform from behind a partition, removing all cues other than the sound of the performance. Unfortunately, if the performance reveals (as it should) any benefits accrued from (differential) past experience, then men will maintain relative success, even in hypothetical cases of equal aptitude and equal ability to benefit from experience. Disadvantages that are inherited from past discrimination are not undone by blinding.

Measuring Individual Differences in Implicit Social Cognition

Implicit social-cognitive effects have been demonstrated most clearly in experimental studies in which a group of subjects is uniformly exposed to cues that influence their subsequent responses on indirect measures; the implicit effect is sought in comparisons between averaged performances of groups exposed to different cues. Although such experimental designs effectively demonstrate some basic properties of implicit social cognition, they do not allow assessment of individual differences. The many existing individual-difference measures of attitudes, self-esteem, and stereotypes primarily assess introspectively accessible self-knowledge (and, sometimes, deliberately managed self-presentations). To measure individual differences in introspectively inaccessible implicit social cognition, sensitive indirect measures are needed. Two categories of such indirect measures have received substantial development effort.

Judgment Latency Measures

Uses of judgment latencies to provide indirect measures of social cognition have been based on the contributions of Donders (1868/1969), Sternberg (1969), and Posner (1978). Representative uses appear in the work of Bargh (1982), Devine (1989), Dovidio and Fazio (1992), Niedenthal (1992), Perdue and Gurtman (1990), and E. R. Smith et al. (1992). Because judgment latencies tend to show substantial within-person variability, obtaining measures with adequate reliability requires averaging the subject's response latencies to large numbers of similar stimuli. Consequently, latency-based indirect measures have been used chiefly for hypothesis tests that compare measures averaged over groups of subjects. Recent treatments of the use of latency measures to assess individual differences in automatic operation of attitudes can be found in works by Bargh et al. (1992), Roskos-Ewoldsen and Fazio (1992), and Fazio (1990a, 1993).

Projective Measures

A very different approach to indirect measurement, based on research of Murray (1943) and McClelland, Atkinson, Clark, and Lowell (1953) , asks a respondent to generate stories in response to ambiguous photographs or drawings, or to generate descriptions of what is seen in abstract stimuli (for example, inkblots). The use of projective measures for implicit motives, and their comparison with measures of explicit motives, was the subject of an article by McClelland, Koestner, and Weinberger (1989) , who concluded that projective and direct measures assess different constructs. A similar conclusion was reached in a meta-analytic review of achievement motivation measures by Spangler (1992) , who also reported that projective measures of achievement motivation had greater predictive validity than did parallel questionnaire (direct) measures.

Other Indirect Measures

There has been a continuing series of well-justified calls for development of indirect measures for use in social psychological research (e.g., Campbell, 1950 ; Hammond, 1948 ; Jahoda et al., 1951 ; Webb, Campbell, Schwartz, Sechrest, & Grove, 1981). Those calls have not gone unheeded, and there is consequently a substantial history of social psychological uses of indirect measures, especially in the domain of prejudice research. For example, indirect measures used by Word, Zanna, and Cooper (1974) showed that White interviewers maintained greater physical distance, demonstrated less eye contact, and administered shorter interviews when interacting with Black (as opposed to White) interviewees. Milgram, Mann, and Harter (1965) showed that stamped envelopes left in various public places ("lost letters") were more likely to be mailed (by the passersby who found them) the more favorable were public attitudes toward the addressee organizations. Gaertner and Bickman (1971) developed an effective telephone version of the Milgram et al. lost-letter technique, showing that it plausibly assessed prejudice toward Blacks. Porter, Geis, and Jennings-Walstedt (1983) showed that greater use of head-of-table seating as an indicator of leadership for male than female stimulus persons provided an indirect measure of gender stereotyping. J. D. Johnson, Jackson, and Gatto (in press) similarly reported an indirect measure of race stereotyping in the form of greater influence by damaging inadmissible evidence for Black than White defendants in a simulated trial. Campbell, Kruskal, and Wallace (1966) observed spontaneous seating aggregation by race in a large classroom, suggesting that spatial proximity can serve as an indirect indicator of racial attitudes. Greenwald and Schuh (in press) developed a citation analog of the Campbell et al. (1966) aggregation measure, showing that scientific citations aggregate along lines of author ethnicity (Jewish and non-Jewish).

In summary of existing efforts at indirect measurement of implicit social cognition: Research on latency decomposition, projective tests, and miscellaneous other procedures indicate that indirect measurement of individual differences in implicit social cognition is possible. At the same time, such measurement has not yet been achieved in the efficient form needed to make research investigation of individual differences in implicit social cognition a routine undertaking. When such measures do become available, there should follow the rapid development of a new industry of research on implicit cognitive aspects of personality and social behavior.

Conclusion

Much social cognition occurs in an implicit mode. This conclusion comes from a reinterpretation of many findings that indicate the importance of implicit operation of attitudes, and of the self-esteem attitude in particular, and also from existing and new evidence for the implicit operation of stereotypes. By adding this conception of the implicit mode to existing knowledge of the explicit mode of operation of social psychology's basic constructs, the scope of those constructs is extended substantially. In addition, many possibilities for application in decision-making settings are suggested by interpreting social judgment in terms of an interaction of implicit and explicit social cognition.

Implicit social cognition overlaps with several concepts that were significant in works of previous generations of psychologists. Psychoanalytic theory's concept of cathexis contained some of the sense of implicit attitude, and its concept of ego defense similarly captured at least part of the present notion of implicit self-esteem. Partly under the influence of psychoanalytic theory, in the 1930s and 1940s, attitudes were regarded as capable of unconscious operation. The authoritarian personality concept (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950) extended the psychoanalytic approach to include social phenomena of prejudice and stereotyping. At a time when the influence of psychoanalytic theory in academic psychology was declining, its conceptions of unconscious phenomena that related to implicit social cognition were being imported into

behavior theory (Dollard & Miller, 1950 ; Doob, 1947 ; Osgood, 1957). The New Look in Perception of the 1950s focused on several phenomena that are interpretable as implicit social cognition. The developing cognitive approach to these phenomena can be seen in Bruner's (1957) introduction of the concept of *perceptual readiness*. Still later, the New Look approach was tied together with psychoanalytic theoretical influences in a cognitive-psychological account by Erdelyi (1974 , 1985) .

Importantly, the psychoanalytic, behaviorist, and cognitive treatments just mentioned all lacked an essential ingredient, that is, they lacked reliable laboratory models of their focal phenomena that could support efficient testing and development of theory. The missing ingredient is now available, as cognitive psychologists have succeeded in producing several varieties of unconscious cognition reliably in the laboratory (see overviews by Greenwald, 1992a ; Kihlstrom, 1987), and investigations of implicit social cognition are well underway (see Bornstein & Pittman, 1992 ; Uleman & Bargh, 1989). The methods of research on implicit memory, in particular, are applicable to the implicit attitude, self-esteem, and stereotype phenomena reviewed in this article. Perhaps the most significant remaining challenge is to adapt these methods for efficient assessment of individual differences in implicit social cognition.

References

Adams, G. R. & Crane, P. (1980). An assessment of parents' and teachers' expectations of preschool children's social preferences for attractive or unattractive children and adults. *Child Development, 51*, 224-231.

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J. & Sanford, R. N. (1950). *The authoritarian personality*. (New York: Harper)

Ajzen, I. & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. (Englewood Cliffs, NJ: Prentice-Hall)

Allmendinger, J. M. & Hackman, J. R. (1993). *The more, the better? On the inclusion of women in professional organizations*. (Report No. 5, Cross-National Study of Symphony Orchestras, Harvard University)

Allport, G. W. (1935). Attitudes.(In C. Murchison (Ed.), *A handbook of social psychology* (pp. 798—844). Worcester, MA: Clark University Press.)

Allport, G. W. (1937). *Personality: A psychological interpretation*. (New York: Holt, Rinehart., & Winston)

Allport, G. W. (1954). *The nature of prejudice*. (Cambridge, MA: Addison-Wesley)

Arkes, J. R., Boehm, L. E. & Xu, G. (1991). Determinants of judged validity. *Journal of Experimental Social Psychology, 27*, 576-605.

Aronson, E. (1968). Dissonance theory: Progress and problems.(In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg., & P. H. Tannenbuam (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 5—27). Chicago: Rand McNally.)

Aronson, E. (1992). The return of the repressed: Dissonance theory makes a comeback. *Psychological Inquiry, 3*, 303-311.

Aronson, E. & Linder, D. E. (1965). Gain and loss of esteem as determinants of interpersonal attractiveness. *Journal of Experimental Social Psychology, 1*, 156-171.

Aronson, E. & Mettee, D. (1968). Dishonest behavior as a function of differential levels of induced self-esteem. *Journal of Personality and Social Psychology, 9*, 121-127.

Ashmore, R. D. & Del Boca, F. K. (1981). Conceptual approaches to stereotypes and stereotyping.(In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 1—36). Hillsdale, NJ: Erlbaum.)

Balzer, W. K. & Sulsky, L. M. (1992). Halo and performance appraisal research: A critical examination. *Journal of Applied Psychology, 77*, 975-985.

Banaji, M. R. & Greenwald, A. G. (1994). Implicit stereotyping and unconscious prejudice.(In M. P. Zanna & J. M. Olson (Eds.), *The psychology of prejudice, The Ontario Symposium* (Vol. 7, pp. 55—76). Hillsdale, NJ: Erlbaum.)

Banaji, M. R. & Greenwald, A. G. (in press). Implicit stereotyping in judgments of fame. *Journal of Personality and Social Psychology, ,*

Banaji, M. R., Hardin, C. & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology, 65*, 272-281.

Banaji, M. R. & Prentice, D. A. (1994). The self in social contexts. *Annual Review of Psychology, 45*, 297-332.

Bargh, J. A. (1982). Attention and automaticity in the processing of self-relevant information. *Journal of*

Personality and Social Psychology, 43, 425-436.
Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition.(In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3—51). New York: Guilford Press.)
Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition.(In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., & Vol. 1, pp. 1—40). Hillsdale, NJ: Erlbaum.)
Bargh, J. A., Chaiken, S., Govender, R. & Pratto, F. (1992). The generality of the automatic activation effect. *Journal of Personality and Social Psychology, 62*, 893-912.
Bar-Tal, D. & Saxe, L. (1976). Physical attractiveness and its relationship to sex-role stereotyping. *Sex Roles, 2*, 123-148.
Basow, S. (1986). *Gender stereotypes: Traditions and alternatives.* (Monterey, CA: Brooks/Cole)
Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91*, 3-26.
Beck, A. T. (1979). *Cognitive therapy of depression.* (New York: Guilford Press)
Begg, I., Armour, V. & Kerr, T. (1985). On believing what we remember. *Canadian Journal of Behavioural Science, 17*, 199-214.
Beggan, J. K. (1992). On the social nature of nonsocial perception: The mere ownership effect. *Journal of Personality and Social Psychology, 62*, 229-237.
Berscheid, E. (1985). Interpersonal attraction.(In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (Vol. 2, pp. 413—484). New York: Random House.)
Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968—1987. *Psychological Bulletin, 106*, 265-289.
Bornstein, R. F. (1992). Subliminal mere exposure effects.(In R. F. Bornstein & T. S. Pittman (Eds.), *Perception without awareness: Cognitive, clinical, and social perspectives* (pp. 191—210). New York: Guilford Press.)
Bornstein, R. F. & D'Agostino, P. A. (1992). Stimulus recognition and the mere exposure effect. *Journal of Personality and Social Psychology, 63*, 545-552.
Bornstein, R. F. & Pittman, T. S. (Eds.). (1992). *Perception without awareness: Cognitive, clinical, and social perspectives.* (New York: Guilford Press)
Brewer, M. (1979). Ingroup bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86*, 307-324.
Brewer, M. (1988). A dual process model of impression formation.(In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1—36). Hillsdale, NJ: Erlbaum.)
Brigham, J. C. (1971). Ethnic stereotypes. *Psychological Bulletin, 76*, 15-33.
Broverman, I., Vogel, S. R., Broverman, D. M., Clarkson, F. & Rosenkrantz, P. S. (1972). Sex role stereotypes: A current appraisal. *Journal of Sex Roles, 28*, 59-78.
Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64*, 123-152.
Bruner, J. S. & Potter, M. C. (1964). Interference in visual recognition. *Science, 144*, 424-425.
Bruner, J. S. & Tagiuri, R. (1954). The perception of people.(In G. Lindzey (Ed.), *Handbook of social psychology* (Vol. 2, pp. 634—654). Reading, MA: Addison-Wesley.)
Byrne, D. (1962). Response to attitude similarity-dissimilarity as a function of affiliation need. *Journal of Personality, 30*, 164-177.
Byrne, D. (1969). Attitudes and attraction.(In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 4, pp. 35—89). New York: Academic Press.)
Byrne, D., Clore, G. L. & Smeaton, G. (1986). The attraction hypothesis: Do similar attitudes affect anything? *Journal of Personality and Social Psychology, 51*, 1167-1170.
Campbell, D. T. (1950). The indirect assessment of social attitudes. *Psychological Bulletin, 47*, 15-38.
Campbell, D. T. (1963). Social attitudes and other acquired behavioral dispositions.(In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 6, pp. 94—172). New York: McGraw-Hill.)
Campbell, D. T., Kruskal, W. & Wallace, W. (1966). Seating aggregation as an index of attitude. *Sociometry, 29*, 1-15.
Cash, T. F. & Duncan, N. C. (1984). Physical attractiveness stereotyping among Black American college students. *The Journal of Social Psychology, 122*, 71-77.
Chaiken, S. (1987). The heuristic model of persuasion.(In M. P. Zanna, J. M. Olson., & C. P. Herman (Eds.), *Social influence: The Ontario Symposium* (Vol. 5, pp. 3—39). Hillsdale, NJ: Erlbaum.)
Cialdini, R. B. (1993). *Influence: Science and practice* (3rd ed.(). New York: Harper Collins)
Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S. & Sloan, L. R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology, 34*, 366-375.

Cialdini, R. B. & Richardson, K. D. (1980). Two indirect tactics of image management: Basking and blasting. *Journal of Personality and Social Psychology, 39*, 406-415.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.(). Hillsdale, NJ: Erlbaum)
Cooley, C. H. (1964). *Human nature and the social order.* (New York: Scribner's. (Original work published 1902)
Cooper, W. H. (1981). Ubiquitous halo. *Psychological Bulletin, 90*, 218-244.
Crocker, J., Thompson, L. L., McGraw, K. M. & Ingerman, C. (1987). Downward comparison, prejudice, and evaluations of others: Effects of self-esteem and threat. *Journal of Personality and Social Psychology, 52*, 907-916.
Crosby, F. (1984). The denial of personal discrimination. *American Behavioral Scientist, 27*, 371-386.
Crosby, F., Bromley, S. & Saxe, L. (1980). Recent unobtrusive studies of Black and White discrimination and prejudice: A literature review. *Psychological Bulletin, 87*, 546-563.
Crosby, F. J., Pufall, A., Snyder, R. C., O'Connell, M. & Whalen, P. (1989). The denial of personal disadvantage among you, me, and all the other ostriches.(In M. Crawford & M. Gentry (Eds.), *Gender and thought: Psychological perspectives* (pp. 79—99). New York: Springer-Verlag.)
Cullen, D. M. (1968). *Attitude measurement by cognitive sampling.* (Unpublished doctoral dissertation, Ohio State University)
Darley, J. M. & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology, 44*, 20-33.
Dawes, R. M. & Smith, T. L. (1985). Attitude and opinion measurement.(In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 1, pp. 509—566). New York: Random House.)
Deutsch, M., Krauss, R. & Rosenau, N. (1962). Dissonance or defensiveness? *Journal of Personality, 30*, 16—28.(Devine P. G. 1989 Journal of Personality and Social Psychology 56 5 18 Dion K. Berscheid E. Walster E. 1972 Journal of Personality and Social Psychology 24 207 213 Dollard J. Miller N. E. 1950 New York: McGraw-Hill.)
Donders, F. C. (1969). Over de snelheid van psychische processen.(Onderzoekingen gedaan in het Psyiologish Laboratorium der Utrechtsche Hoogeschool; 1868—1869 [On the speed of mental processes]. (W. G. Koster, Ed. and Trans.). *Acta Psychologica, 30*, 412—431. (Original work published 1868))
Doob, L. (1947). The behavior of attitudes. *Psychological Review, 54*, 135-156.
Dovidio, J. F., Evans, N. E. & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology, 22*, 22-37.
Dovidio, J. F. & Fazio, R. H. (1992). New technologies for the direct and indirect assessment of attitudes.(In J. Tanur (Ed.), *Questions about questions: Meaning, memory, expression, and social interactions in surveys* (pp. 204—237). New York: Russell Sage Foundation.)
Downs, A. C. & Lyons, P. M. (1991). Natural observations of the links between attractiveness and initial legal judgments. *Personality and Social Psychology Bulletin, 17*, 541-547.
Eagly, A. H., Ashmore, R. D., Makhijani, M. G. & Longo, L. C. (1991). What is beautiful is good, but .(A meta-analytic review of research on the physical attractiveness stereotype.) *Psychological Bulletin, 110*, 109-128.
Eagly, A. H. & Chaiken, S. (1993). *The psychology of attitudes.* (Fort Worth, TX: Harcourt Brace)
Eagly, A. H., Makhijani, M. G. & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin, 111*, 3-22.
Eagly, A. H. & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin, 15*, 543-558.
Eagly, A. H., Mladinic, A. & Otto, S. (1991). Are women evaluated more favorably than men? An analysis of attitudes, beliefs, and emotions. *Psychology of Women Quarterly, 15*, 203-216.
Edwards, A. L. (1957). *Techniques of attitude scale construction.* (New York: Appleton-Century-Crofts)
Edwards, K. (1990). The interplay of affect and cognition in attitude formation and change. *Journal of Personality and Social Psychology, 59*, 212-216.
Efran, J. S. (1974). The effect of physical appearance on the judgment of guilt, interpersonal attraction, and severity of recommended punishment in a simulated jury trial. *Journal of Research in Personality, 8*, 45-54.
Epstein, S. (1973). The self-concept revisited: Or a theory of a theory. *American Psychologist, 28*, 404-416.
Epstein, S. (1991). Cognitive-experiential self-theory: An integrative theory of personality.(In R. Curtis (Ed.), *The relational self: Convergences in psychoanalysis and social psychology* (pp. 111—137). New York: Guilford Press.)
Erdelyi, M. H. (1974). A new look at the New Look: Perceptual defense and vigilance. *Psychological Review, 81*, 1-25.

Erdelyi, M. H. (1985). *Psychoanalysis: Freud's cognitive psychology.* (New York: W. H. Freeman)
Erdelyi, M. H. & Kleinbard, J. (1978). Has Ebbinghaus decayed with time? The growth of recall (hypermnesia) over days. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 275-289.
Fazio, R. H. (1986). How do attitudes guide behavior? In R.(M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 204—243). New York: Guilford Press.)
Fazio, R. H. (1990a). A practical guide to the use of response latency in social psychological research.(In C. Hendrick & M. S. Clark (Eds.), *Research methods in personality and social psychology* (pp. 74—97). Newbury Park, CA: Sage.)
Fazio, R. H. (1990b). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework.(In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75—110). San Diego: Academic Press.)
Fazio, R. H. (1993). Variability in the likelihood of automatic attitude activation: Data re-analysis and commentary on the paper by Bargh, Chaiken, Govender, and Pratto. *Journal of Personality and Social Psychology, 64*, 753-758.
Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C. & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229-238.
Fazio, R. H. & Zanna, M. P. (1981). Direct experience and attitude-behavior consistency.(In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 161—202). New York: Academic Press.)
Feingold, A. (1990). Gender differences in effects of physical attractiveness on romantic attraction: A comparison across five research paradigms. *Journal of Personality and Social Psychology, 59*, 981-993.
Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117-140.
Festinger, L. (1964). Behavioral support for opinion change. *Public Opinion Quarterly, 28*, 404-417.
Festinger, L. & Walster, E. (1964). The post-decision process.(In L. Festinger (Ed.), *Conflict, decision, and dissonance* (pp. 100—112). Stanford, CA: Stanford University Press.)
Feys, J. (1991). Briefly induced belongingness to self and preference. *European Journal of Social Psychology, 21*, 547-552.
Finch, J. F. & Cialdini, R. B. (1989). Another indirect tactic of (self-) image management: Boosting. *Personality and Social Psychology Bulletin, 15*, 222-232.
Fishbein, M. (1967). A consideration of beliefs and their role in attitude measurement.(In M. Fishbein (Ed.), *Readings in attitude theory and measurement* (pp. 257—266). New York: Wiley.)
Fishbein, M. & Ajzen, I. (1974). Attitudes toward objects as predictors of single and multiple behavior criteria. *Psychological Review, 81*, 59-74.
Fiske, S. T. (1982). Schema-triggered affect: Applications to social perception.(In M. S. Clark & S. T. Fiske (Eds.), *Affect and cognition: The seventeenth annual Carnegie Symposium on Cognition* (pp. 55—76). Hillsdale, NJ: Erlbaum.)
Fiske, S. T. (1989). Examining the role of intent: Toward understanding its role in stereotyping and prejudice.(In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 253—283). New York: Guilford Press.)
Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist, 48*, 621-628.
Fiske, S. T. & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation.(In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1—74). San Diego, CA: Academic Press.)
Frank, M. G. & Gilovich, T. (1988). The dark side of self- and social perception: Black uniforms and aggression in professional sports. *Journal of Personality and Social Psychology, 54*, 74-85.
Gaertner, S. L. & Bickman, L. (1971). Effects of race on the elicitation of helping behavior: A focus on "liberals." In P. Katz (Ed.), *Toward the elimination of racism* (pp. 183—211). New York: Pergamon Press.)
Gaertner, S. L. & Dovidio, J. F. (1986). The aversive form of racism.(In S. L. Gaertner & J. F. Dovidio (Eds.), *Prejudice, discrimination, and racism* (pp. 61—89). New York: Academic Press.)
Gaertner, S. L., Mann, J. A., Dovidio, J. F., Murrell, A. J. & Pomare, M. (1990). How does cooperation reduce intergroup bias? *Journal of Personality and Social Psychology, 59*, 692—704.(Gaertner S. L. McLaughlin J. P. 1983 Social Psychology Quarterly 46 23 30 Geis F. 1993 In A. E. Beall & R. J. Sternberg (Eds.), *The psychology of gender* (pp. 9—54). New York: Guilford Press.)
Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process.(In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189—211). New York: Guilford Press.)
Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46*, 107-119.
Gilbert, D. T. & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs.

Journal of Personality and Social Psychology, 60, 509-517.
Goldberg, P. (1968). Are women prejudiced against women? *Transaction, 5*, 28—30.(Graf P. Mandler G. 1984 Journal of Verbal Learning and Verbal Behavior 23 553 568 Graf P. Schacter D. 1985 Journal of Experimental Psychology: Learning, Memory, and Cognition 11 501 518 Green D. M. Swets J. A. 1966 New York: Wiley.)
Greenwald, A. G. (1968). Cognitive learning, cognitive response to persuasion, and attitude change.(In A. G. Greenwald, T. C. Brock., & T. M. Ostrom (Eds.), *Psychological foundations of attitudes* (pp. 147—170). New York: Academic Press.)
Greenwald, A. G. (1969). The open-mindedness of the counterattitudinal role player. *Journal of Experimental Social Psychology, 5*, 375-388.
Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist, 35*, 603-618.
Greenwald, A. G. (1990). What cognitive representations underlie social attitudes? *Bulletin of the Psychonomic Society, 28*, 254—260.(Greenwald A. G. 1992a American Psychologist 47 766 779 Greenwald A. G. 1992b Unpublished manuscript, University of Washington.)
Greenwald, A. G. & Albert, R. D. (1968). Acceptance and recall of improvised arguments. *Journal of Personality and Social Psychology, 8*, 31-34.
Greenwald, A. G., Klinger, M. R. & Liu, T. J. (1989). Unconscious processing of dichoptically masked words. *Memory and Cognition, 17*, 35-47.
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.)
Greenwald, A. G. & Ronis, D. L. (1978). Twenty years of cognitive dissonance: Case study of the evolution of a theory. *Psychological Review, 85*, 53-57.
Greenwald, A. G. & Schuh, E. S. (in press). An ethnic bias in scientific citations. *European Journal of Social Psychology, ,*
Guimond, S. & Dube-Simard, L. (1983). Relative deprivation theory and the Quebec nationalist movement: The cognitive-emotion distinction and the personal-group deprivation issue. *Journal of Personality and Social Psychology, 44*, 526-535.
Hamilton, D. L. & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotype judgments. *Journal of Experimental Social Psychology, 12*, 392-407.
Hamilton, D. L. & Sherman, J. W. (1994). Stereotypes.(In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., & Vol. 2, pp. 1—68). Hillsdale, NJ: Erlbaum.)
Hamilton, D. L. & Trolier, T. K. (1986). Stereotypes and stereotyping: An overview of the cognitive approach.(In S. L. Gaertner & J. F. Dovidio (Eds.), *Prejudice, discrimination, and racism* (pp. 127—157). New York: Academic Press.)
Hammond, K. (1948). Measuring attitudes by error-choice: An indirect method. *Journal of Abnormal and Social Psychology, 43*, 38-48.
Hasher, L., Goldstein, D. & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior, 16*, 107-112.
Hawkins, S. A. & Hoch, S. J. (1992). Low-involvement learning: Memory without evaluation. *Journal of Consumer Research, 19*, 212-225.
Heider, F. (1944). Social perception and phenomenal causality. *Psychological Review, 51*, 358-374.
Heider, F. (1958). *Interpersonal relations.* (New York: Wiley)
Helmreich, R., Stapp, J. & Ervin, C. (1974). The Texas Social Behavior Inventory, an objective measure of self-esteem. *JSAS Catalog of Selected Documents in Psychology, 4*, 79
Higgins, E. T. (1989). Knowledge accessibility and activation: Subjectivity and suffering from unconscious sources.(In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 75—123). New York: Guilford Press.)
Higgins, E. T., Rholes, W. S. & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology, 13*, 141-154.
Hoorens, V. (1990). Nuttin's affective selfparticles hypothesis and the name letter effect: A review. *Psychologica Belgica, 30*, 23-48.
Howard, D. J. (1992). Gift-wrapping effects on product attitudes: A mood-biasing explanation. *Journal of Consumer Psychology, 3*, 197-224.
Hull, C. L. (1943). *Principles of behavior.* (New York: Appleton-Century-Crofts)
Hull, C. L. (1952). *A behavior system.* (New Haven, CT: Yale University Press)
Jacoby, L. L. & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*, 306-340.

Jacoby, L. L. & Kelley, C. M. (1987). Unconscious influences of memory for a prior event. *Personality and Social Psychology Bulletin, 13*, 314-336.
Jacoby, L. L., Kelley, C. M., Brown, J. & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56*, 326-338.
Jacoby, L. L., Lindsay, D. S. & Toth, J. P. (1992). Unconscious influences revealed: Attention, awareness, and control. *American Psychologist, 47*, 802-809.
Jacoby, L. L., Toth, J. P., Lindsay, D. S. & Debner, J. A. (1992). Lectures for a layperson: Methods for revealing unconscious processes.(In R. F. Bornstein & T. S. Pittman (Eds.), *Perception without awareness* (pp. 81—120). New York: Guilford Press.)
Jacoby, L. L. & Witherspoon, D. (1982). Remembering without awareness. *Canadian Journal of Psychology, 36*, 300-324.
Jahoda, M., Deutsch, M. & Cook, S. (1951). *Research methods in social relations: With special reference to prejudice.* (*Part 1: Basic processes* . New York: Dryden Press.)
James, W. (1890). *The principles of psychology.* (New York: Holt)
Jamieson, D. A. & Zanna, M. P. (1989). Need for structure in attitude formation and expression.(In A. R. Pratkanis, S. J. Breckler., & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 383—406). Hillsdale, NJ: Erlbaum.)
Janis, I. L. & King, B. (1954). The influence of role-playing on opinion change. *Journal of Abnormal and Social Psychology, 49*, 211-218.
Johnson, J. D., Jackson, L. A. & Gatto, L. (in press). Justice is still not colorblind: Differential racial effects of exposure to inadmissible evidence. *Personality and Social Psychology Bulletin, ,*
Johnson, M. M. S. (1986). *The initial letter effect: Ego-attachment or mere exposure?* Unpublished doctoral dissertation, Ohio State University.(Jones E. E. 1964 New York: Appleton-Century-Crofts.)
Jost, J. T. & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology, 33*, 1-27.
Judd, C. M. & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review, 100*, 109-128.
Kahneman, D., Knetsch, J. L. & Thaler, R. H. (1990). Experimental test of the endowment effect and the coase theorem. *Journal of Political Economy, 98*, 1325-1347.
Kasof, J. (1993). Sex bias in the naming of stimulus persons. *Psychological Bulletin, 113*, 140-163.
Katz, D. & Braly, K. (1933). Racial stereotypes of 100 college students. *Journal of Abnormal and Social Psychology, 28*, 280-290.
Katz, D. & Braly, K. (1935). Racial prejudice and racial stereotypes. *Journal of Abnormal and Social Psychology, 30*, 175-193.
Kihlstrom, J. F. (1987). The cognitive unconscious. *Science, 237*, 1445-1452.
Kihlstrom, J. F. (1990). The psychological unconscious.(In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 445—464). New York: Guilford Press.)
King, B. T. & Janis, I. L. (1956). Comparison of the effectiveness of improvised versus non-improvised role playing in producing opinion changes. *Human Relations, 9*, 177-186.
Klinger, M. R. & Beall, P. M. (1992, May). *Conscious and unconscious effects of stereotype activation.* (Paper presented at the annual meetings of the Midwestern Psychological Association, Chicago)
Krech, D., Crutchfield, R. S. & Ballachey, E. L. (1962). *Individual in society.* (New York: McGraw-Hill)
Krosnick, J. A., Betz, A. L., Jussim, L. J. & Lynn, A. R. (1992). Subliminal conditioning of attitudes. *Personality and Social Psychology Bulletin, 18*, 152-162.
Kruglanski, A. W. & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology, 19*, 448-468.
Kunda, Z. & Sanitioso, R. (1989). Motivated changes in the self-concept. *Journal of Experimental Social Psychology, 25*, 272-285.
Landy, D. & Sigall, H. (1974). Beauty is talent: Task evaluation as a function of the performer's physical attractiveness. *Journal of Personality and Social Psychology, 29*, 299-304.
Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32*, 311-329.
LaPiere, R. (1934). Attitudes versus actions. *Social Forces, 13*, 230-237.
Lerner, M. (1980). *The belief in a just world: A fundamental delusion.* (New York: Plenum Press)
Lewicki, P. (1986). *Nonconscious social information processing.* (New York: Academic Press)
Lombardi, W. J., Higgins, E. T. & Bargh, J. A. (1987). The role of consciousness in priming effects on

categorization: Assimilation versus contrast as a function of awareness of the priming task. *Personality and Social Psychology Bulletin, 13*, 411-429.
Lord, C. G., Ross, L. & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*, 2098-2109.
Madigan, S. (1976). Reminiscence and item recovery in free recall. *Memory & Cognition, 4*, 233-236.
Mandler, G. (in press). On remembering without really trying: Hypermnesia, incubation, and mind-popping.(In C. Umiltàg & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and unconscious information processing.* Cambridge, MA: MIT Press.)
Marcel, A. (1983). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology, 15*, 238-300.
Martin, L. L. & Achee, J. W. (1992). Beyond accessibility: The role of processing objectives in judgment.(In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 195—216). Hillsdale, NJ: Erlbaum.)
Martin, L. L., Seta, J. J. & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. *Journal of Personality and Social Psychology, 59*, 27-37.
Martin, L. L. & Tesser, A. (Eds.). (1992). *The construction of social judgments.* (Hillsdale, NJ: Erlbaum)
May, M. & Doob, L. W. (1937). *Competition and cooperation.* (New York: Social Science Research Council)
McClelland, D. C., Atkinson, J. W., Clark, R. A. & Lowell, E. L. (1953). *The achievement motive.* (New York: Appleton-Century-Crofts)
McClelland, D. C., Koestner, R. & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96*, 690—702.(McGuire W. J. 1973 Journal of Personality and Social Psychology 26 446 456 McGuire W. J. 1985 In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 2, pp. 233—346). New York: Random House.)
Messick, D. M. & Mackie, D. M. (1989). Intergroup relations. *Annual Review of Psychology, 40*, 45-81.
Milgram, S., Mann, L. & Harter, S. (1965). The lost-letter technique: A tool for social science research. *Public Opinion Quarterly, 29*, 437-438.
Mowrer, O. H. (1960). *Learning theory and symbolic processes.* (New York: Wiley)
Mullen, B., Brown, R. J. & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Psychology, 22*, 103-122.
Murphy, S. T. & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64*, 723-729.
Murray, H. A. (1943). *Thematic apperception test manual.* (Cambridge, MA: Harvard University Press)
Myers, D. G. (1990). *Social psychology* (3rd ed.(). New York: McGraw-Hill)
Myers, D. G. & Ridl, J. R. (1979, August). A better than average insight into pride.(*Psychology Today*, pp. 89, 95—96, 98.)
Neuberg, S. L. (1989). The goal of forming accurate impressions during social interactions: Attenuating the impact of negative expectancies. *Journal of Personality and Social Psychology, 56*, 374-386.
Niedenthal, P. M. (1990). Implicit perception of affective information. *Journal of Experimental Social Psychology, 26*, 505-527.
Niedenthal, P. M. (1992). Affect and social perception: On the psychological validity of rose-colored glasses.(In R. F. Bornstein & T. S. Pittman (Ed.), *Perceptionwithout awareness: Cognitive, clinical, and social perspectives* (pp. 211—235). New York: Guilford Press.)
Nisbett, R. E. & Wilson, T. D. (1977a). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.
Nisbett, R. E. & Wilson, T. D. (1977b). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35*, 250-256.
Nuttin, J. M. (1985). Narcissism beyond Gestalt and awareness: The name letter effect. *European Journal of Social Psychology, 15*, 353-361.
Nuttin, J. M. (1987). Affective consequences of mere ownership: The name letter effect in twelve European languages. *European Journal of Social Psychology, 17*, 381-402.
Onodera, T. & Miura, M. (1990). Physical attractiveness and its halo effects on a partner: "Radiating beauty" in Japan also? *Japanese Psychological Research, 32*, 148-153.
Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17*, 776-783.
Osgood, C. E. (1957). A behavioristic analysis of perception and language as cognitive phenomena.(In *Contemporary approaches to cognition* (pp. 75—118). Cambridge, MA: Harvard University Press.)
Osgood, C. E., Suci, G. J. & Tannenbaum, P. H. (1957). *The measurement of meaning.* (Urbana, IL: University

of Illinois Press)
Ostrom, T. M. & Sedikides, C. (1992). Out-group homogeneity effects in natural and minimal groups. *Psychological Bulletin, 112*, 536-552.
Overson, C. & Mandler, G. (1987). Indirect word priming in connected semantic and phonological contexts. *Bulletin of the Psychonomic Society, 25*, 229-232.
Paulhus, D. L. & Levitt, K. (1987). Desirable responding triggered by affect: Automatic egotism? *Journal of Personality and Social Psychology, 52*, 245-259.
Paulhus, D. L., Martin, L. C. & Murphy, G. K. (1992). Some effects of arousal on sex stereotyping. *Personality and Social Psychology Bulletin, 18*, 325-330.
Perdue, C. W., Dovidio, J. F., Gurtman, M. B. & Tyler, R. B. (1990). Us and them: Social categorization and the process of intergroup bias. *Journal of Personality and Social Psychology, 59*, 475-486.
Perdue, C. W. & Gurtman, M. B. (1990). Evidence for the automaticity of ageism. *Journal of Experimental Social Psychology, 26*, 199-216.
Peters, D. P. & Ceci, S. J. (1982). Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral and Brain Sciences, 5*, 187-255.
Petty, R. E. & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches.* (Dubuque, IA: Brown)
Petty, R. E. & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to persuasion.* (New York: Springer-Verlag)
Petty, R. E. & Wegener, D. T. (1993). Flexible correction processes in social judgment: Correcting for context-induced contrast. *Journal of Experimental Social Psychology, 29*, 137-165.
Porter, N., Geis, F. L. & Jennings-Walstedt, J. (1983). Are women invisible as leaders? *Sex Roles, 9*, 1035-1049.
Posner, M. I. (1978). *Chronometric explorations of mind.* (Hillsdale, NJ: Erlbaum)
Pratkanis, A. R. (1988). The attitude heuristic and selective fact identification. *British Journal of Social Psychology, 27*, 257-263.
Pratto, F. & Bargh, J. A. (1991). Stereotyping based on apparently individuating information: Trait and global components of sex stereotypes under attention overload. *Journal of Experimental Social Psychology, 27*, 26-47.
Prentice, D. & Miller, D. T. (1992). *The psychology of ingroup attachment.* (Paper presented at the conference on The Self and The Collective, Princeton University. Princeton, NJ)
Regan, D. T. (1971). Effects of a favor and liking on compliance. *Journal of Experimental Social Psychology, 7*, 627-639.
Reingold, E. M. & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception and Psychophysics, 44*, 563-575.
Richardson-Klavehn, A. & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39*, 474-543.
Roediger, H. L. (1990). Implicit memory: Retention without awareness. *American Psychologist, 45*, 1043-1056.
Roediger, H. L. & Blaxton, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word fragment completion. *Memory and Cognition, 15*, 379-388.
Roediger, H. L., Weldon, M. S. & Challis, B. H. (1989). Explaining dissociations between measures of retention: A processing account.(In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness* (pp. 3—42). Hillsdale, NJ: Erlbaum.)
Rogers, C. (1951). Client-centered therapy.() *Boston: Houghton-Mifflin, ,*
Rosenbaum, M. (1986). The repulsion hypothesis: On the nondevelopment of relationships. *Journal of Personality and Social Psychology, 51*, 1156-1166.
Rosenberg, M. J. (1965). When dissonance fails: On eliminating evaluation apprehension from attitude measurement. *Journal of Personality and Social Psychology, 1*, 28-42.
Rosenberg, M. J. (1969). The conditions and consequences of evaluation apprehension.(In R. Rosenthal & R. L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 279—349). New York: Academic Press.)
Rosenberg, M. (1979). *Conceiving the self.* (New York: Basic Books)
Roskos-Ewoldsen, D. R. & Fazio, R. H. (1992). On the orienting value of attitudes: Attitude accessibility as a determinant of an object's attraction of visual attention. *Journal of Personality and Social Psychology, 63*, 198-211.
Sanitioso, R., Kunda, Z. & Fong, G. T. (1990). Motivated recruitment of autobiographical memories. *Journal of Personality and Social Psychology, 59*, 229-241.
Sarnoff, I. (1960). Psychoanalytic theory and social attitudes. *Public Opinion Quarterly, 24*, 251-279.
Schacter, D. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, 12, 432-444.
Scheier, M. F. & Carver, C. S. (1992). Effects of optimism on psychological and physical well-being: Theoretical overview and empirical update. *Cognitive Therapy and Research, 16*, 201-228.
Schooler, J. W. & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology, 22*, 36-71.
Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states.(In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 527—561). New York: Guilford Press.)
Schwarz, N. & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment.(In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217—245). Hillsdale, NJ: Erlbaum.)
Schwarz, N. & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*, 513-523.
Schwarz, N., Strack, F. & Mai, H. P. (1991). Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *Public Opinion Quarterly, 55*, 3-23.
Secord, P. (1959). Stereotyping and favorableness in the perception of Negro faces. *Journal of Abnormal and Social Psychology, 59*, 309-315.
Seligman, M. E. P. (1991). *Learned optimism.* (New York: Knopf)
Sherif, M. & Cantril, H. (1947). *The psychology of ego-involvements.* (New York: Wiley)
Sherif, M. & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change.* (New Haven, CT: Yale University Press)
Sherman, S. J., Mackie, D. M. & Driscoll, D. M. (1990). Priming and the differential uses of dimensions in evaluation. *Personality and Social Psychology Bulletin, 16*, 405-418.
Sigall, H. & Ostrove, N. (1975). Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on juridic judgment. *Journal of Personality and Social Psychology, 31*, 410-414.
Smith, E. R. (1984). Model of social inference processes. *Psychological Review, 91*, 392-413.
Smith, E. R. (1990). Content and process specificity in the effects of prior experiences.(In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 3, pp. 1—60). Hillsdale, NJ: Erlbaum.)
Smith, E. R. (1994). Procedural knowledge and processing strategies in social cognition.(In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., & Vol. 1, pp. 99—151). Hillsdale, NJ: Erlbaum.)
Smith, E. R. & Branscombe, N. R. (1988). Category accessibility as implicit memory. *Journal of Experimental Social Psychology, 24*, 490-504.
Smith, E. R., Stewart, T. L. & Buttram, R. T. (1992). Inferring traits from a behavior has long-term highly specific effects. *Journal of Personality and Social Psychology, 62*, 753-759.
Smith, M. B., Bruner, J. S. & White, R. W. (1956). *Opinions and personality.* (New York: Wiley)
Snyder, M. (1981). On the self-perpetuating nature of social stereotypes.(In D. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 183—212). Hillsdale, NJ: Erlbaum.)
Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin, 112*, 140-154.
Spence, K. W. (1956). *Behavior theory and conditioning.* (New Haven, CT: Yale University Press)
Srull, T. K. & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37*, 1660-1672.
Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self.(In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261—302). New York: Academic Press.)
Steele, C. M., Spencer, S. J. & Lynch, M. (1992). *Self-image resilience and dissonance: The role of affirmational resources.* (Unpublished manuscript, Stanford University)
Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57*, 421-457.
Strack, F. (1992). The different routes to social judgments: Experiential versus informational strategies.(In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 249—275). Hillsdale, NJ: Erlbaum.)
Strack, F., Schwarz, N., Bless, H., Kübler, A. & Wänke, M. (1993). Awareness of the influence as a determinant of assimilation versus contrast. *European Journal of Social Psychology, 23*, 53-62.
Suls, J. (1991). Similarity and self-esteem in downward comparison.(In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 51—78). Hillsdale, NJ: Erlbaum.)
Suls, J. & Wills, T. A. (1991). *Social comparison: Contemporary theory and research.* (Hillsdale, NJ: Erlbaum)
Swim, J., Borgida, E., Maruyama, G. & Myers, D. G. (1989). Joan McKay versus John McKay: Do gender

stereotypes bias evaluation? *Psychological Bulletin, 105*, 409-429.
Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223*, 96-102.
Tajfel, H. & Turner, J. C. (1986). The social identity theory of intergroup behavior.(In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7—24). Chicago: Nelson-Hall.)
Taylor, D. M., Wright, S. C., Moghaddam, F. M. & Lalonde, R. N. (1990). The personal/group discrimination discrepancy: Perceiving my group, but not myself, to be the target of discrimination. *Personality and Social Psychology Bulletin, 16*, 254-262.
Taylor, S. E. & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.
Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior.(In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181—227). New York: Academic Press.)
Tesser, A. & Campbell, J. (1983). Self-definition and self-evaluation maintenance.(In J. Suls & A. G. Greenwald (Eds.), *Psychological perspectives on the self* (Vol. 2, pp. 1—32). Hillsdale, NJ: Erlbaum.)
Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4*, 25-29.
Thurstone, L. L. (1931). The measurement of attitudes. *Journal of Abnormal and Social Psychology, 26*, 249-269.
Thurstone, L. L. & Chave, E. J. (1929). *The measurement of attitude.* (Chicago: University of Chicago Press)
Tolman, E. C. (1959). Principles of purposive behavior.(In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 2, pp. 92—157). New York: McGraw-Hill.)
Tsui, A. S. & Barry, B. (1986). Interpersonal affect and rating errors. *Academy of Management Journal, 29*, 586-599.
Tulving, E. & Schacter, D. (1990). Priming and human memory systems. *Science, 267*, 301-306.
Uleman, J. S. (1987). Consciousness and control: The case of spontaneous trait inferences. *Personality and Social Psychology Bulletin, 13*, 337-354.
Uleman, J. S. & Bargh, J. A. (Eds.). (1989). *Unintended thought.* (New York: Guilford Press)
Vallone, R. P., Ross, L. & Lepper, M. R. (1985). The hostile media phenomenon: Biased perception and perceptions of media bias in coverage of the "Beirut Massacre." *Journal of Personality and Social Psychology, 49*, 577-585.
Wallston, B. & O'Leary, V. (1981). Sex makes a difference: Differential perceptions of men and women. *Review of Personality and Social Psychology, 2*, 9-41.
Webb, E. J., Campbell, D. T., Schwartz, R. D., Sechrest, L. B. & Grove, J. B. (1981). *Nonreactive measures in the social sciences.* (Boston: Houghton-Mifflin)
Wegner, D. M. & Vallacher, R. R. (1977). *Implicit psychology.* (New York: Oxford)
Wegner, D. M. & Vallacher, R. R. (1981). Common-sense psychology.(In J. P. Forgas (Ed.), *Social cognition: Perspectives on everyday understanding* (pp. 225—246). London: Academic Press.)
Wicker, A. W. (1969). Attitudes vs. actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues, 25*, 41-78.
Wilder, D. A. & Shapiro, P. N. (1984). Role of out-group cues in determining social identity. *Journal of Social and Personality Psychology, 47*, 342-348.
Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin, 90*, 245-271.
Wilson, P. R. (1968). Perceptual distortion of height as a function of ascribed academic status. *Journal of Social Psychology, 74*, 97-102.
Wilson, T. D., Dunn, D. S., Kraft, D. & Lisle, D. J. (1989). Introspection, attitude change, attitude-behavior consistency: The disruptive effects of explaining why we feel the way we do.(In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 22, 287—343). New York: Academic Press.)
Wilson, T. D. & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60*, 181-192.
Winter, L. & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology, 47*, 237-252.
Wood, J. V., Taylor, K. L., Michela, J. L. & Gaus, V. (1993). *Measuring self-esteem indirectly: Personality correlates and responses to threats to self-esteem.* (Unpublished manuscript, University of Waterloo, Ontario, Canada)
Word, C. O., Zanna, M. P. & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology, 10*, 109-120.
Wyatt, D. F. & Campbell, D. T. (1951). On the liability of stereotype or hypothesis. *Journal of Abnormal and*

Social Psychology, 46, 496-500.
Wylie, R. C. (1974). *The self-concept* ((Vol. 1). Lincoln: University of Nebraska Press)
Wylie, R. C. (1979). *The self-concept* ((Vol. 2). Lincoln: University of Nebraska Press)
Zajonc, R. B. (1968). Attitudinal effects of mere exposure.(*Journal of Personality and Social Psychology, 9* (No. 2, Pt. 2).)
Zanna, M. P. & Fazio, R. H. (1982). The attitude-behavior relation: Moving toward a third generation of research.(In M. P. Zanna, E. T. Higgins., & C. P. Herman (Eds.), *Consistency in social behavior: The Ontario Symposium* (Vol. 2, pp. 283—301). Hillsdale, NJ: Erlbaum.)
Zanna, M. P. & Rempel, J. K. (1988). Attitudes: A new look at an old concept.(In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 315—334). Cambridge, England: Cambridge University Press.)

1

The terms *implicit—explicit* capture a set of overlapping distinctions that are sometimes labeled as *unaware—aware, unconscious—conscious, intuitive—analytic, direct—indirect, procedural—declarative,* and *automatic—controlled.* These dichotomies vary in the amount and nature of implied theoretical interpretation. This article uses the *implicit—explicit* pair because of that dichotomy's prominence in recent memory research, coupled with the present intention to connect research on attitudes, self-esteem, and stereotypes to memory research.

2

Some notes on this definition: First, it was generated from the earlier stated template for definitions of implicit cognition. Second, alternate definitions of the attitude response class might be substituted for the stated one ("favorable or unfavorable feeling, thought, or action toward social objects"). Third, the qualification "[introspectively] inaccurately identified" includes an important class of cases in which a prior experience is identifiable, but its influence on an evaluative response is not. For example, a student may be aware of having been graded highly in a course, but not suspect that this experience influences responses to the course's end-of-term course evaluation survey.

3

The mere exposure effect does not share a major feature of the other implicit attitude effects reviewed here, namely, evaluation of one object displaced onto another. Even though recent research has greatly advanced understanding of the mere exposure effect, the basis for its evaluative character remains obscure.

4

Gilbert (1991) does not interpret the effect of comprehension on judged validity to be due to a misattribution of familiarity to validity but, rather, to a direct effect of comprehension. Establishment of an empirical test that can distinguish between these interpretations may be difficult.

5

There has recently been some debate over the relative contribution of similarity versus dissimilarity cues to similarity-attraction findings (Byrne, Clore, & Smeaton, 1986 ; Rosenbaum, 1986). The opposing views in this debate fit equally with an interpretation in terms of implicit self-esteem, but differ in their assumptions about the relative importance of dissimilarity-induced negative attitudes and similarity-induced positive attitudes.

6

Postdecisional spreading of choice alternatives is well known as a prediction of cognitive dissonance theory. Interpretations of cognitive dissonance reduction as often occurring in the service of maintaining a positive self-image (e.g., Aronson, 1968 , 1992), or in the service of self-affirmation (Steele, 1988), or in the service of maintaining self-esteem (Deutsch, Krauss, & Rosenau, 1962 ; Greenwald & Ronis, 1978) potentially place the entire class of dissonance-reducing cognitive changes into the category of implicit self-esteem effects.

7

It might be supposed that greater frequency of exposure to initial letters than other name letters can explain why the name letter effect is strongest for initials. However, this interpretation is implausible both because empirical exposure frequency effects asymptote at much lower frequencies than the huge frequencies characteristic of letters of the alphabet, and any frequency advantage for initials over other name letters must be quite small when computed as a percentage difference in total exposures to the letters.

8

Some of the naturally mediated implicit self-esteem effects described previously also likely serve to support self-esteem, but even less directly than do those considered in this section. For example, by helping to assure that one's acquaintances will express agreeable views, the similarity-attraction effect protects against exposure to opinion disagreements that might undermine confidence in the wisdom of one's opinions.

9

As one example, Heider's (1958) balance theory analyzes the triad composed of (a) a perceiver and (b) a positively evaluated other who is associatively linked to (c) a previously neutral object. Substituting self into the role of other yields the implicit self-esteem effect (liking for a self-associated object) as the result of Heider's theorized tendency toward cognitive balance (with all links positive) in this triad.

10

This prediction of the moderating role of self-esteem is in some cases quite compatible with existing theories. For example, it is readily translated into the terms of both balance theory and cognitive dissonance theory. Nevertheless, with only occasional exceptions (e.g., Aronson & Mettee, 1968 ; Crocker et al., 1987 ; Steele et al., 1992 ; Suls, 1991), those theories have not guided searches for evidence to evaluate possible relationships of their predicted effects to individual differences in self-esteem.

11

This unavailability of measures of implicit self-esteem may be rectified by efforts currently underway (Wood, Taylor, Michela, & Gaus, 1993).

12

In a meta-analytic review of this research, Swim, Borgida, Maruyama, and Myers (1989) concluded that evidence for this indirect form of gender stereotyping was, on average, weak. However, examination of the

portions of their meta-analysis that corresponded most closely to Goldberg's (1968) original procedures does support the conclusion that Goldberg's finding is well established (cf. Banaji & Greenwald, 1994).

Preparation of this report as well as conduct of some of the research reported in it were supported by National Science Foundation Grants DBC-9205890 and DBC-9120987 and by National Institute of Mental Health Grant MH-41328. We thank Ieck Ajzen, John Bargh, R. Bhaskar, Irene Blair, Robert Bornstein, Marilynn Brewer, Robert Crowder, Leonard Doob, Russell Faizo, Klaus Fiedler, Deborrah Frable, Daniel Gilbert, Jack Glaser, Richard Hackman, Curtis Hardin, Roger Hughes, John Jost, Larry Jacoby, John Kihlstrom Mark Klinger, David Myers, Delroy Paulhus, Richard Petty, Alex Rothman, Peter Salovely, Eric Schuh, Norbert Schwarz, Constantine Sedikides, Jerzy Trzebinski, James Uleman, Wendi Walsh, Timothy Wilson, Joanne Wood, and two anonymous reviewers for comments and for permission to use the data presented in Figure 1 .
This article is dedicated to the memory of Tom Ostrom, a dear colleague who continues to have a profound influence on both of the authors.
Correspondence may be addressed to Anthony G. Greenwald, Department of Psychology, NI-25, University of Washington, Seattle, Washington, 98195.
Electronic mail may be sent to agg@u.washington.edu
Received: July 20, 1993
Revised: December 7, 1993
Accepted: April 12, 1994

Figure 1. An implicit self-esteem effect (the initial letter effect). Subjects judged letters in their names to be more attractive than those not in their names, with most of this effect being explained by liking for first and last initial letters. (Data provided by 597 subjects, with means based on a minimum of 13 observations; from M. M. S. Johnson, 1986 .)

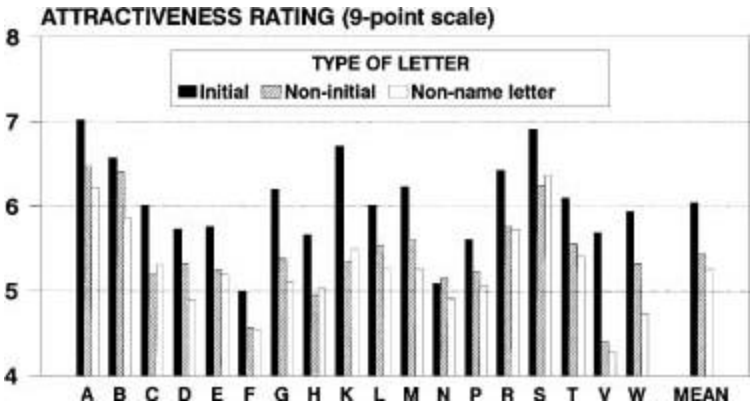
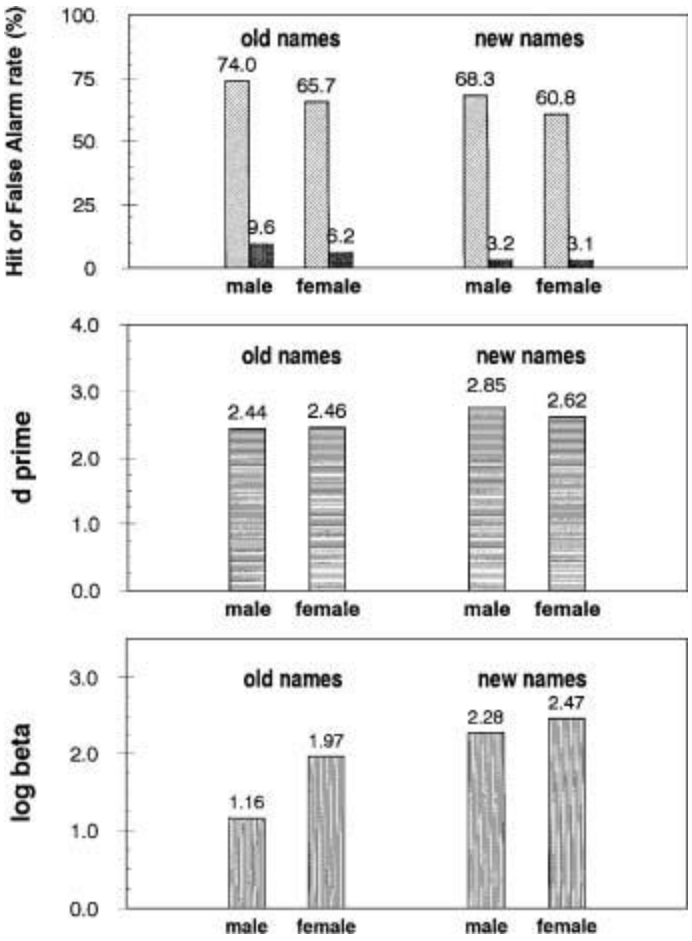


Figure 2. An implicit stereotyping effect. (Data from Banaji & Greenwald, in press , averaged over four experiments, total $N = 186$; procedure described in text.) Signal-detection analysis indicated that subjects were somewhat more sensitive (d') to male than female fame for "new" names, which were presented for the first time in Session 2. However, the most substantial result was that subjects had a lower criterion (β) for judging fame for male than female names, especially for "old" names, which had been increased in familiarity by being presented in Session 1. (Hit and false-alarm rates are, respectively, the tall and the short bars in the top panel.)


Hit or False Alarm rate (%)
100
75
50
25
0
old names
new names
74.0
9.6
65.7
6.2
68.3
3.2
60.8
3.1
male
female
male
female
d prime
4.0
3.0
2.0
1.0
0.0
old names
new names
2.44
2.46
2.85
2.62
male
female
male
female
log beta
3.0
2.0
1.0
0.0
old names
new names
1.16
1.97
2.28
2.47
male
female
male
female