DeSteno, V, & B + study 1

(10) L/H-relationships, OR H/8-self-esteem, pg 626

(11) avg (2), pg 631

(12) self (3), pg 630

(13) self (3), pg 630

(14) Not (0), pg 631

(15) self-rep eno (3), pg 631

(16) algor (3), pg 630 ← ⊙ error: sard log before ✓

(17) words (0), pg 630

(18) one (1), pg 630

(19) 1
(20) 2
(21) 3
(22) bef (1)
(23) bef (1) ←
(24) same (0) –
(25) same (0) –

pg 631

(26) sd-5, pg 630
(27) sd-5, pg 630
(28) con-3.5, p 631
(29) spec-1, p. 630-1
(30) spec-1, p. 631
(31) comp-1, p 630

⊙ error: file says 2("att-1") ✓

(32) Not (0), pg 626
(33) dual (2), pg 630

Journal of Experimental Social Psychology
www.elsevier.com/locate/jesp

ed on the tendency to explain Black stereotype-
ypic Explanatory Bias or SEB) is predictive of be-
3 predicted White males' choice to ask stereotypic
interview. In Experiment 2 the type of explanation
amined. Results showed that White participants who
re rated more positively and those who made external
se results point to the potential of implicit stereotyping as

Thompson, a

09-1109, USA
A
Australia

typing

on

(e.g., Greenwald, McGhee, & Schwartz, 1998). Some-
what less research attention has focused on implicit
stereotype measures (e.g., Wittenbrink, Judd, & Park-
1997), and implicit stereotypes/stereotyping for
unintended influence of stereotypes on information
processing (cf. Brewer, 1996). In part, this focus on
prejudice rather than stereotyping probably
emerged because prejudice has traditionally been
thought to be more consequential than stereotyping for
behavioral outcomes such as discrimination (Brigham,
1971; Stangor, Sullivan, & Ford, 1991).
To the extent that measures of implicit prejudice and
stereotyping assess important processes relevant to in-
tergroup attitudes and perceptions (von Hippel, Se-
kaquaptewa, & Vargas, 1995, 1997), it seems reasonable
to expect them to relate to intergroup behavior. Yet
such demonstrations are rare. In one study, White
participants who implicitly favored Whites over
Americans were rated by observers as having
interactions with a White than a Black
nell & Leibold, 2001; see also

nal of Personality and Social Psychology
06, Vol. 91, No. 4, 626–641

0022-3514/06/$12.00 DOI: 10.1037/0022-3514.91.4.626

Jealousy and the Threatened Self: Getting to the Heart of the Green-Eyed Monster

David DeSteno, Piercarlo Valdesolo, and Monica Y. Bartlett
Northeastern University

Several theories specifying the causes of jealousy have been put forth in the past few decades. Firm support for any proposed theory, however, has been limited by the difficulties inherent in inducing jealousy and examining any proposed mediating mechanisms in real time. In support of a theory of jealousy centering on threats to the self-system, 2 experiments are presented that address these past limitations and argue for a model based on context-induced variability in self-evaluation. Experiment 1 presents a method for evoking jealousy through the use of highly orchestrated social encounters and demonstrates that threatened self-esteem functions as a principal mediator of jealousy. In addition to replicating these findings, Experiment 2 provides direct evidence for jealousy as a cause of aggression. The ability of the proposed theory of jealousy to integrate other extant findings in the literature is also discussed.

Keywords: jealousy, self-esteem, aggression, emotion

Jealousy, it seems, is a fundamental aspect of human social life. For as far back in time or as widely across civilizations as one can peer, the green-eyed monster has reared its head. From Gilgamesh's romps retold in the first millennium B.C.E, to Othello's throes portrayed in the middle part of the last millennium, to modern day soap operas and drama series, fascination with the jealousy motif has not waned among artists and audiences alike. From cultures representing geographically and socially disparate milieus, research documents the pervasiveness of jealousy among men and women from childhood to old age (e.g., Bryson, 1991; Buunk, Angleitner, Oubaid, & Buss, 1996; Geary, Rumsey, Bow-Thomas, & Hoard, 1995; Hupka et al., 1985; Masciuch & Kienapple, 1993). Jealousy's ubiquity is so well accepted that even Freud (1922/1955) himself suggested that its absence, not its presence (at least within normal levels), is a sign of pathology.

From a functional perspective, jealousy stands as an exemplary candidate for a fundamental social emotion. Emotions, like many psychological phenomena, are theorized to exist because they serve some adaptive purpose. That is, although their specific components and sequelae may operate on many different levels (e.g., neurochemical, interpersonal, cultural), emotions are designed to increase the success with which an organism meets specific challenges by shunting cognition and behavior toward certain outcomes (Frijda, 2000; Keltner & Gross, 1999; Keltner & Haidt, 1999; Lazarus, 1991; LeDoux & Phelps, 2000; Öhman & Wiens, 2003). The cognitive and physiological changes associated with fear and anxiety, for example, prepare an organism to detect and/or escape from an impending danger more efficiently (LeDoux & Phelps, 2000; Öhman, 2002). It is important to note, however, that organisms whose existence is characterized by high degrees of collective or social living confront not only challenges involving the successful navigation of the physical environment but also those involving the social one (e.g., social exchange, coalition building, social bonding, and relationship maintenance; Bartlett & DeSteno, 2006; Cosmides & Tooby, 2000; Darwin, 1872/1998; Keltner & Busswell, 1997; Keltner & Haidt, 1999; Lewis, 2000). The importance of such challenges suggests the need for specific emotional responses that are intrinsically tied to sociality.

Jealousy: Form and Function

For humans, adaptive functioning is intrinsically tied to social interactions through which myriad needs are met (e.g., protection, resource acquisition, reproduction). Accordingly, engagement in interpersonal relationships stands as a fundamental predictor of human physical and psychological health (Baumeister & Leary, 1995; Berscheid & Reis, 1998; Cacioppo et al., 2002) and is fostered by the seemingly universal motive to belong to social groups and be a member of interpersonal relationships (Baumeister & Leary, 1995). Indeed, involvement in social relationships is of such central value to adaptive functioning that it has been documented to increase psychological well-being (Diener, 1984; Myers & Diener, 1995), resistance to cardiovascular disease (Berkman, Vaccarino, & Seeman, 1993), resistance to cancer (Glanz & Lerman, 1992), and immune system function (Booth & Pennebaker, 2000; Kennedy, Kiecolt-Glaser, & Glaser, 1990; Kiecolt-Glaser, 1999).

Given the benefits provided by relationships, competition for them frequently arises (Salovey, 1991). Consequently, the exis-

David DeSteno, Piercarlo Valdesolo, and Monica Y. Bartlett, Department of Psychology, Northeastern University.

Monica Bartlett is now at the Department of Psychology, Gonzaga University.

This research was supported by National Institute of Mental Health Grant MH068240. We thank Nilanjana Dasgupta and members of the Boston Emotion Research Lab for insights and comments regarding this work.

Correspondence concerning this article should be addressed to David DeSteno, Department of Psychology, Northeastern University, Boston, MA 02115. E-mail: d.desteno@neu.edu

evaluations that are based on intellectual, athletic, or other abilities; cf. Crocker & Wolfe, 2001). Additionally, explicit measures of self-esteem usually require a more deliberate consideration of the self. Such measures, because of their greater controllability, are more amenable to strategic attempts meant to obscure threats to self-esteem (Greenwald & Banaji, 1995; Koole et al., 2001).

The basic structure of the two studies is quite similar and involved the formation and subsequent threatening of a valued relationship through the interaction of participants with two confederates: one playing the role of partner and one the role of rival. The jealousy manipulation involved whether the partner indicated interest in working with the rival and ended his or her working relationship with the participant. Following this manipulation, participants completed measures of self-esteem and jealousy. In addition, the second study examined the links between jealousy and direct aggression aimed at partners and rivals.

Study 1

The primary goals of this study were to demonstrate that jealousy can be evoked in a laboratory setting and to investigate whether jealousy is mediated by threats to self-esteem. As noted earlier, the occurrence of jealousy is not limited to romantic relationships; it occurs in relationships of all types involving a valued partner. Accordingly, we expected that after participants formed a novel and pleasant relationship with a work partner, threats to that relationship posed by a rival should produce jealousy. Moreover, we expected that jealousy intensity would vary as a direct function of decreases in self-esteem. The intensity of any resulting jealousy can be expected to be relatively mild as the relationship is quite new. Nonetheless, jealousy should occur whenever there is a threat to even a budding relationship of potential value and, thereby, provide an opportunity to examine the functioning of this emotion in real time.

Method

Participants

Forty-six female undergraduates at Northeastern University participated in this experiment in partial fulfillment of a course requirement.[3] Participants were randomly assigned to either the jealousy or the control condition.

Manipulations and Measures

Jealousy manipulation. In order to induce jealousy in vivo, a complex triadic interaction involving the participant was staged through specific actions by two confederates playing the respective roles of the partner and the rival. The details of the induction are noted in the procedure description in the following section as they are integrated with the unfolding of the experimental paradigm. In brief, a confederate playing the role of the partner forms an enjoyable working relationship with each participant. At a later point in the experimental session, the bonds of this relationship are threatened and broken because of either the usurpation of the relationship by a confederate playing the rival (i.e., the jealousy condition) or fate (i.e., the control condition). In all conditions, the partner was male and the rival was female.

Implicit self-esteem. Implicit self-esteem (ISE) was assessed with an implicit association test (IAT) based closely on that developed by Greenwald and Farnham (2000). This measure has been shown to possess good reliability and predictive validity with respect to both self-report and behavioral measures (Bosson, Swann, & Pennebaker, 2000; Greenwald & Farnham, 2000). For example, ISE measures that use the IAT have been demonstrated to predict defensive behavior in response to threats to self-esteem when used to assess narcissism in consort with explicit self-esteem measures (Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003; McGregor & Marigold, 2003). This measure has also been shown to predict clinical status with respect to depression and susceptibility of depressed individuals to contextual changes in self-evaluation and mood (Gemar, Segal, Sagrati, & Kennedy, 2000).

In this task, the self-versus-other category was represented by 10 self-relevant versus 10 nonself-relevant items. The evaluative attribute was represented by 10 pleasant (e.g., joy, peace) and 10 unpleasant (e.g., agony, vomit) words (see the Appendix for the complete stimulus set). Stimuli were presented by using DirectRT software (Jarvis, 2004) on PC-type desktop computers (Intel Pentium III, 550 MHz processors) equipped with CRT color monitors.

At the start of the ISE task, each participant provided the self-relevant information items (e.g., last name, student ID) in response to prompts by the computer (see the Appendix for the complete set of prompts). Of importance, these items did not possess any intrinsic positive or negative qualities; any valenced associations would only arise through their association with the self. In order to disallow any sense of personal association with the nonself-relevant stimuli, a set of 10 items matching the form of the self-relevant items was provided for all participants (see the Appendix for the complete list). The assumption of lack of any self-association was checked both through the comparison of generated items and debriefing.

After providing this information, participants completed an IAT that assessed self-esteem. Participants were instructed to categorize four types of stimuli (self-relevant vs. other-relevant information, pleasant vs. unpleasant words) by using two designated response keys. Errors were always noted by the appearance of the word *error* on the screen, after which participants had to press the appropriate key to continue to the next trial. Response latencies for error trials were recorded as the time from stimulus onset to the time of correct categorization (Greenwald, Nosek, & Banaji, 2003). In the first block (20 trials), participants categorized items as belonging to the self or other category. In the second block (20 trials), participants categorized words as pleasant or unpleasant. In the third block (20 practice trials followed by 40 critical trials), participants completed a combined categorization task by classifying informational items as self or other and words as pleasant or unpleasant by using the two keys (for a randomly selected half of the participants, pleasant was paired with self and unpleasant with other; for the other half, this pairing was reversed). In the fourth block (20 trials), participants had to categorize pleasant versus unpleasant words by using the opposite keys to those used in the earlier blocks. Finally, in the fifth block (20 practice trials followed by 40 critical trials), participants again completed a combined categorization task by classifying information items as self or other and words as pleasant or unpleasant by using the two keys. In this block, all participants categorized self–nonself and pleasant–unpleasant stimuli in a manner that was opposite to the stimulus pairing combination used in the third block.

To the extent that participants held a positive evaluation of themselves, they should have been faster at associating self-related words with pleasant stimuli and slower at associating self-related words with unpleasant stimuli (Dasgupta, Greenwald, McGhee, & Banaji, 2000; Greenwald & Farnham, 2000; Greenwald et al., 1998, 2003). Scoring of the ISE measure was done in accordance with the *D* algorithm developed by Greenwald et al. (2003). Each participant's *D* was computed by subtracting the mean response time for Block 3 from Block 5 and dividing the resulting quantity by the pooled standard deviation of the two blocks. The *D* measure may be conceptually understood as an index of individual differences in the degree to which

[3] The sample was limited to women due to gender constraints in the participant pool.

responses for the Self + Bad trials were slower than those for the Self + Good trials adjusted for individual differences in the variability of response times. Higher *D* values indicate higher self-esteem as indexed by increased difficulty in completing the Self + Unpleasant as compared with the Self + Pleasant trials.

The *D* metric has been shown to be free from contamination effects due to stimuli ordering and to group differences in task-switching ease (Mierke & Klauer, 2003). Therefore, any resulting differences between the experimental conditions that use this metric cannot be attributed to the effects of simple distraction arising from the use of the jealousy manipulation. That is, differences in *D* scores did not occur because the jealousy manipulation simply occupied cognitive resources in the jealousy group (e.g., rumination) and, thereby, made it more difficult for individuals to respond to the changing stimulus pairings inherent in the IAT.

Explicit self-esteem. Explicit self-esteem was assessed by using the State Self-Esteem Scale (Heatherton & Polivy, 1991).

Jealousy. Jealousy was assessed by using a feeling state questionnaire in which participants indicated the degree to which each of 10 adjectives described their current state. The questionnaire consisted of both positive and negative items, embedded in which were four items that specifically targeted jealousy: jealous, angry, betrayed, and hurt (Cronbach's $\alpha = .81$). Parrot and Smith (1993) have demonstrated that these feeling descriptors capture the multifaceted experience of jealousy in a way that is distinct from other related negative emotions (e.g., envy). Participants' jealousy scores reflect the mean score on these four items.

Procedure

Participants were run individually for all sessions. Upon arrival at the lab, the participant (S) was greeted by the experimenter and asked to sit in a chair in front of a cubicle containing a PC. The room contained five such cubicles with an accordion wall that partially expanded so as to separate two cubicles from the other three. Immediately after S entered the room, a confederate playing the role of the partner (P) arrived and was similarly greeted. The experimenter then informed them that two other participants were also scheduled to arrive and that they would therefore wait a few minutes before beginning the session. At this point, P introduced himself to S and began use of predetermined conversational probes that were designed to initiate a sense of familiarity and liking. After 3 min, the experimenter returned and noted that the experiment would begin without the other participants.

The experimenter informed S and P that the study in which they would take part was designed to examine differences in task performance levels as a function of working alone or in pairs. Moreover, as some of the tasks would be conducted on computers, the experiment would also involve S and P taking two hand–eye coordination tests that would allow the experimenter to adjust scores for individual differences in hand–eye acuity for computer use. After the first such test, S and P would be free to choose to work together or alone on the first problem-solving task.

At this point, S and P were instructed to turn to their computers to complete the first hand–eye coordination task. In actuality, this task was an IAT taken from Greenwald et al. (1998) that assesses positive attitudes toward flowers versus insects. Its only purpose was to familiarize participants with the IAT so that it would require less instruction to complete the ISE measure after the introduction of the critical manipulation.

When S and P had finished this IAT, the experimenter returned to the room and provided instructions for the first problem-solving task (in actuality, participants would only complete one such task). This was a word unscrambling task. P and S were handed sheets of paper that contained letter matrices at the top of each. The task was to find as many words as possible that were contained in each matrix. After reminding them that as there were only 2 of them they could choose to work together or alone, the experimenter left the room. P then turned to S and asked if she would like to work together.[4]

The problem-solving task served only as a vehicle to foster the formation of a pleasant working relationship. During the next 5 min, P's task was to ensure that S enjoyed working with him. He did this through repeated smiling and the use of a set of verbal responses. For example, he would provide encouragement (e.g., "let's see if we can figure out this one") and validation (e.g., "that's a good one" and "I'm glad we're doing this together") to the participant. After 5 min had passed, a knock was heard at the door and the experimenter appeared from a side room to answer it. The confederate playing the rival (R) then entered the room and apologized to the experimenter for being late. The experimenter informed R that she would complete the earlier hand–eye coordination task at the end of the experiment, handed her a clipboard containing the materials for the word scramble task, gave brief instructions for it, and left the room. R then grabbed a chair and sat next to S and P. For the next 3 min, the three individuals worked together. However, R was instructed to devote most of her attention and interactions (i.e., validations and encouragements) to P.

At this point, the critical manipulation occurred. In the jealousy condition, P suddenly noted that he thought the experimenter said they could only work alone or in pairs. After expressing concern that this could be a problem, he went into the next room and asked the experimenter. The experimenter and P returned to the room at which point the experimenter noted that they could only work in pairs or alone before turning to leave. P then turned toward R and asked if she would like to continue as his partner. R agreed and the two moved to the other side of the room (i.e., behind the partially expanded accordion wall) and continued working within earshot of S for 1 min. In the control condition, P suddenly noted that he had an appointment at the campus medical center that he had forgotten. He then went into the next room to tell the experimenter who could be heard excusing him with the caveat that he return later to finish the study. In this way, the enjoyable working relationship was severed in both conditions. However, in one it was due to the presence of a rival and in the other to consequences of fate.

At this point in both conditions, the experimenter then returned to the room and instructed the individuals to turn toward their individual PCs and to follow the instructions provided. Participants then completed the implicit and explicit measures of self-esteem, the jealousy scale, and a questionnaire concerning demographic information. The confederate(s) always left the experimental room before the participant had finished. Upon completion of the study, participants were extensively debriefed and given a small gift of candy for their participation.

Results and Discussion

In accord with expectations, the termination of a relationship due to a partner leaving to work with a rival as opposed to leaving for a scheduling conflict was successful in evoking jealousy. Participants reported higher levels of jealousy in the jealousy condition ($M = 1.65$, $SD = 0.89$) than in the control condition ($M = 1.21$, $SD = 0.30$), $t(44) = 2.21$, $p = .03$. It is instructive to note that the variation in reported jealousy is quite large in the jealousy condition relative to the control condition.[5] Thus, even though the mean level of self-reported jealousy in the jealousy condition falls in the mild to moderate range, it masks a high degree of variability. Such differences in variability are to be expected given the lack of jealousy in the control condition and individual differences in self-presentational concerns related to the stigmatizing nature associated with admitting to jealous feelings

[4] In all cases except one, this proposal was accepted. Data from the participant who chose to work alone were discarded from all analyses.

[5] A *t* test assuming unequal variances for the two groups also showed a significant difference in jealousy ($t = 2.29$, $p = .03$).

Rydell & Mc 1 | Study 4

(10) 11/L - rel?
(11) attitude = avg (2), distance = single (1) pg 1004 plus..
(12) ~~att~~ (1), pg 998
(13) att (1), pg 998
(14) Not (0) for desired control ~~attitude~~, obs (1) for distance, pg 1004
(15) s-r ~~f~~ judge (4) for desired control ~~att~~, obs (5) for distance pg 1004
(16) not rep (3), pg 999
(17) pics (1), pg 998 — error: put words (0) before
(18) one (1), pg 998
(19) 2nd (2) I, p1004 ~~E, I, B~~
(20) 1st (1) E, p1004
(21) 3rd (3) B, p1004
(22) bet (1)
(23) bet (1)
(24) same (0)
(25) same (0)
(26) sd - (1) pg 1004, 998
(27) sd - (1) pg 1004, 998
(28) desired dist - 9, seat = 3 pg 1004
(29) spec = 3, pg 1004 & 998
(30) spec = 3, pg 1004 & 998
(31) comp - 2, pg 998, 1004
(32) Non-rec (0), pg 995
(33) dual (2)

Understanding Implicit and Explicit Attitude Change: A Systems of Reasoning Analysis

Robert J. Rydell
University of California, Santa Barbara

Allen R. McConnell
Miami University

There is considerable controversy about how to conceptualize implicit and explicit attitudes, reflecting substantial speculation about the mechanisms involved in implicit and explicit attitude formation and change. To investigate this issue, the current work examines the processes by which new attitudes are formed and changed and how these attitudes predict behavior. Five experiments support a systems of reasoning approach to implicit and explicit attitude change. Specifically, explicit attitudes were shaped in a manner consistent with fast-changing processes, were affected by explicit processing goals, and uniquely predicted more deliberate behavioral intentions. Conversely, implicit attitudes reflected an associative system characterized by a slower process of repeated pairings between an attitude object and related evaluations, were unaffected by explicit processing goals, uniquely predicted spontaneous behaviors, and were exclusively affected by associative information about the attitude object that was not available for higher order cognition.

Keywords: implicit attitudes, explicit attitudes, attitude change

The study of attitudes—evaluations of the self, individuals, groups, and other objects—has a long and rich history in social psychology (Eagly & Chaiken, 1993). In recent years, the focus of attitude research has shifted from understanding explicit attitudes (i.e., attitudes that people can report and for which activation can be consciously controlled) to examining implicit attitudes (i.e., attitudes for which people do not initially have conscious access and for which activation cannot be controlled).[1] Past research has shown that relying on implicit rather than explicit measures of attitudes can circumvent self-presentational motives (e.g., Dunton & Fazio, 1997) and can often uniquely predict spontaneous behaviors (e.g., McConnell & Leibold, 2001); however, less is known about the processes underlying how implicit and explicit attitudes form and operate. The current work posited that there are important differences between them, especially in how they change. Specifically, we propose that explicit attitudes form and change through the use of fast-learning, rule-based reasoning, whereas implicit attitudes form and change through the use of slow-learning, associative reasoning (Sloman, 1996).

Heretofore, implicit attitude change and explicit attitude change have been studied in relative isolation. Indeed, research on explicit attitude change has been one of the most productive areas of study in social psychology (Eagly & Chaiken, 1993; Petty & Wegener, 1998). Although some researchers have found that implicit attitudes are relatively difficult to change with conventional attitude change manipulations (e.g., Gawronski & Strack, 2004; Gregg, Seibt, & Banaji, 2006; Petty, Tormala, Briñol, & Jarvis, 2006), other research has demonstrated that implicit attitudes can change relatively quickly in response to contextual stimuli or social roles (e.g., Barden, Maddux, Petty, & Brewer, 2004; Dasgupta & Greenwald, 2001; Wittenbrink, Judd, & Park, 2001). But despite these demonstrations, the theory underlying implicit attitude change is relatively underdeveloped (see Devine, 2001; Fazio & Olson, 2003; Wilson, Lindsey, & Schooler, 2000), and experimental paradigms that can systematically examine the concurrent formation and change of implicit and explicit attitudes

Robert J. Rydell, Department of Psychology, University of California, Santa Barbara; Allen R. McConnell, Department of Psychology, Miami University.

This research was supported by National Institute of Mental Health Grant MH068279 and National Science Foundation Grant BCS 0516931. Portions of this work were submitted by Robert J. Rydell in partial fulfillment of the requirements for a doctoral degree from Miami University. We thank Doris Bergen, Heather Claypool, Dave Hamilton, Kurt Hugenberg, Diane Mackie, Jeff Sherman, and Laura Strain for their extremely helpful comments and guidance in this research.

Correspondence concerning this article should be addressed to Robert J. Rydell, Department of Psychology, University of California, Santa Barbara, Santa Barbara, CA 93106-9660. E-mail: rydell@psych.ucsb.edu

[1] Although there is disagreement about the use of the terms *implicit attitudes* and *explicit attitudes* in the literature (e.g., Fazio & Olson, 2003), we agree with Strack and Deutsch (2004) who note that "explicit and implicit measures are defined by the cognitive operations that they capture. In this sense, explicit measures tap into people's knowledge or beliefs, implicit measures tap into their associative structures" (p. 239; see also, Wilson et al., 2000). Because we contrast and compare implicit and explicit measures, we use the terms *implicit attitudes* and *explicit attitudes* throughout this article.

Journal of Personality and Social Psychology, 2006, Vol. 91, No. 6, 995–1008
 0022-3514/06/$12.00 DOI: 10.1037/0022-3514.91.6.995

toward him. Second, the introduction of the counterattitudinal information provided a window in which explicit (which relies on the fast-learning system) but not implicit (which relies on the slow-learning system) attitudes should change in the face of new target-relevant information. Thus, we have an opportunity to assess and understand how this new information affects implicit attitudes and explicit attitudes differently, shedding light on the processes involved in their change.

We expected to observe that people would quickly change their explicit attitudes in the face of counterattitudinal information, especially when the initial learning was very consistent (Kerpelman & Himmelfarb, 1971) and when the counterattitudinal information was negative (Fiske, 1980). However, we did not expect implicit attitudes to change as quickly in response to a modest amount of counterattitudinal information nor did we expect to observe a valence asymmetry for implicit attitudes.

Method

Participants. A sample of 170 undergraduates at Miami University participated in return for research credit in their introductory psychology courses. They were randomly assigned to a 2 (valence of learned attitude: positive vs. negative) × 2 (level of reinforcement: 100%, 75%) × 2 (counterattitudinal condition: control vs. counterattitudinal conditioning) between-subjects factorial.

Learning task. The current work used a modified version of the attitude learning paradigm developed by Kerpelman and Himmelfarb (1971). Specifically, participants were presented with a target person's behaviors that were either relatively positive or negative in valence, and participants judged whether each behavior was characteristic or uncharacteristic of him. As part of a between-subjects manipulation, participants were given different levels of reinforcement in their responses, leading them to form different attitudes toward him.

First, participants completed the learning task on a computer, in which they were told that they would be receiving information about a person named *Bob*. In the initial learning trials, participants read 100 behaviors performed by *Bob* while a picture of *Bob* was presented on the computer monitor directly above each behavior.[3] After reading each behavior, participants indicated whether they believed that the behavior was characteristic or uncharacteristic of *Bob* by pressing the *C* key (characteristic) or the *U* key (uncharacteristic). After they responded, participants were given feedback about whether the behavior was characteristic of *Bob* for 5 s. Specifically, feedback consisted of the word *correct* (in blue text) or *incorrect* (in red text) positioned in the center of the computer monitor and, at the same time, the behavior was restated "correctly," on the basis of the assigned reinforcement condition, at the bottom of the computer monitor (e.g., "Helping the neighborhood children is characteristic of Bob." or "Helping the neighborhood children is uncharacteristic of Bob."). In the initial 100 learning trials, the feedback given portrayed *Bob* as positive or as negative in 100% or in 75% of the behaviors (with 25 of the trials in the 75% reinforcement condition being counterattitudinal). The ordering of the behaviors and feedback were randomly determined (in accordance with the experimental condition) for each participant.

Following these 100 trials, participants in the control condition received 20 neutral trials (i.e., the behavior performed by *Bob* was neither positive nor negative; e.g., "Bob waited at the street corner."). However, participants in the counterattitudinal condition (20 CA) received counterattitudinal feedback about *Bob* on 20 trials (i.e., the behaviors that were described as characteristic or uncharacteristic of *Bob* were opposite of the valence presented during the initial learning trials). Finally, participants completed implicit and explicit attitude measures.[4]

Explicit attitude measure. To assess explicit attitudes, participants judged how likable *Bob* was on a scale ranging from 1 (*very unlikable*) to 9 (*very likable*). In addition, they completed five semantic differential scales, each using a 9-point scale to describe *Bob*: good–bad, pleasant–mean, agreeable–disagreeable, caring–uncaring, and kind–cruel. Further, participants provided their evaluation of *Bob* on a feeling thermometer that ranged in temperature from 0° to 100°. The response for each explicit measure was standardized and an overall mean was computed (in all experiments to be reported, αs > .90). Then the standardized scores in the negative valance condition were reverse scored so that greater scores on this measure indicated that explicit attitudes were more extreme in the direction of initial learning.

Implicit attitude measure. The Implicit Associations Test (IAT; Greenwald, McGhee, & Schwartz, 1998) was used to assess implicit attitudes toward *Bob*. The IAT had 26 stimuli: 1 picture of *Bob*, 5 different pictures of White men who were not *Bob*, 10 positive adjectives (e.g., *wonderful*), and 10 negative adjectives (e.g., *disgusting*). All stimuli were presented in the center of the monitor, and the adjectives were always presented in lowercase letters.

This IAT task was a modified version of the task used by McConnell and Leibold (2001), featuring seven blocks with 20 trials per block. Participants were informed that the task involved making category judgments for a variety of stimuli (photos or words) presented on a computer monitor by using one of two responses (the *D* or *K* keys on the keyboard). During each block, category label reminders were displayed on the left and right sides of the display (assignment of particular labels to the *D* and *K* keys was counterbalanced across participants and produced no effects). Participants were instructed to complete that task quickly while also minimizing errors, and they were told to keep their index fingers on the *D* and *K* keys throughout the experiment to minimize delays in responding. There was a 250-ms intertrial interval.

In Block 1, participants judged photos of *Bob* or *not Bob* and in Block 2 they judged whether the adjectives were "negative" or "positive." In Blocks 3 and 4 (Combination 1), participants judged whether the stimuli were "Bob or negative" or "not Bob or positive." In Block 5, participants performed the same judgment task as Block 2 except the assignment of response keys assigned to the two valence categories was reversed. Finally, in Blocks 6 and 7 (Combination 2), participants judged whether the stimuli were "Bob or positive" or "not Bob or negative." As in past IAT research, half of the participants performed Combination 1 in Blocks 3–4 and Combination 2 in Blocks 6–7, whereas the rest performed Combination 2 in Blocks 3–4 and Combination 1 in Blocks 6–7 (this counterbalancing manipulation produced no effects).

In order to assess implicit attitudes toward *Bob*, we subtracted the mean response latencies of Combination 2 from the mean response latencies of Combination 1 (regardless of the order they were completed).[5] Again, the

[3] Photographs of one of 5 different White males were randomly presented as *Bob*. These 5 White males were judged as equal in attractiveness and the target used did not affect the results in any of the experiments. The positive and the negative behaviors used in the current work were borrowed from those developed by McConnell, Sherman, and Hamilton (1994a).

[4] In all of the experiments, half of the participants completed the implicit measure first and the other half completed the explicit measure first. This order variable produced no effects in any of the studies and thus is not discussed further.

[5] Following Greenwald et al. (1998), all trials in the critical blocks were retained, responses faster than 300 ms were recoded as 300 ms, and trials slower than 3,000 ms were recoded as 3,000 ms. After any such adjustments were made, each latency was then log transformed to reduce positive skew inherent in response latency data (Fazio, 1990). Alternative scoring techniques for the IAT (e.g., Greenwald, Nosek, & Banaji, 2003) produced the same results in all studies reported. Analyses were performed on the log-transformed values, but means are reported as standardized scores.

standardized scores in the negative valance condition were reverse scored so that greater scores on this measure indicated that implicit attitudes were more extreme in the direction of initial learning.

Results

Explicit attitudes. To examine whether explicit attitudes changed in response to small amounts of counterattitudinal information and were more likely to show attitude change with greater initial reinforcement, a 2 (valence of learned attitude) × 2 (level of reinforcement) × 2 (counterattitudinal condition) analysis of variance (ANOVA) was conducted on explicit attitude extremity (see Figure 1). First, there were significant main effects of level of reinforcement, $F(1, 162) = 39.22$, $p < .001$, and of counterattitudinal condition, $F(1, 162) = 89.90$, $p < .001$. As one would expect, the main effect of level of reinforcement showed that explicit attitudes were more extreme in the direction of initial learning in the 100% reinforcement condition ($M = 0.94$, $SD = 0.62$) than in the 75% reinforcement condition ($M = 0.54$, $SD = 0.45$). Similarly, the main effect of counterattitudinal condition revealed that explicit attitudes were more extreme in the direction of initial learning in the control condition ($M = 1.04$, $SD = 0.52$) than in the 20 CA condition ($M = 0.44$, $SD = 0.47$). More important, the anticipated two-way interaction between level of reinforcement and counterattitudinal condition was significant, $F(1, 162) = 19.06$, $p < .001$. To examine this interaction, the simple effect of counterattitudinal condition was examined for each level of reinforcement. In the 75% reinforcement condition, there was a simple effect of counterattitudinal condition, $F(1, 162) = 13.54$, $p < .001$, showing that participants in the control condition had more extreme explicit attitudes toward *Bob* ($M = 0.71$, $SD = 0.42$); than participants in the 20 CA condition ($M = 0.38$, $SD = 0.41$). In the 100% reinforcement condition, there was an even stronger effect of counterattitudinal condition, $F(1, 162) = 97.03$, $p < .001$, indicating that although participants in the control condition had especially extreme explicit attitudes (in the direction of initial conditioning) toward *Bob* ($M = 1.37$, $SD = 0.37$), counterattitudinal information led to far less extreme attitudes toward *Bob* ($M = 0.51$, $SD = 0.52$). Thus, the interaction reflects the much larger effect of counterattitudinal condition on explicit attitude extremity in the 100% reinforcement condition than in the 75% reinforcement condition (replicating Kerpelman & Himmelfarb, 1971). Consistent with negative asymmetries, the two-way interaction between counterattitudinal condition and valence of learned attitude was also significant, $F(1, 162) = 16.64$, $p < .001$. In the positive learned attitudes condition, those in the control condition had far more extreme explicit attitudes ($M = 1.17$, $SD = 0.59$) than those in the 20 CA condition ($M = 0.28$, $SD = 0.47$), $F(1, 162) = 76.70$, $p < .001$. In the negative learned attitudes condition, this effect was significant but weaker, with those in the control condition having more extreme explicit attitudes ($M = 0.91$, $SD = 0.39$) than those in the 20 CA condition ($M = 0.61$, $SD = 0.42$), $F(1, 162) = 8.12$, $p < .005$. In other words, negative counterattitudinal information had a greater impact on attitude extremity than did positive counterattitudinal information (e.g., Fiske, 1980; Skowronski & Carlston, 1987). No other effects were significant.

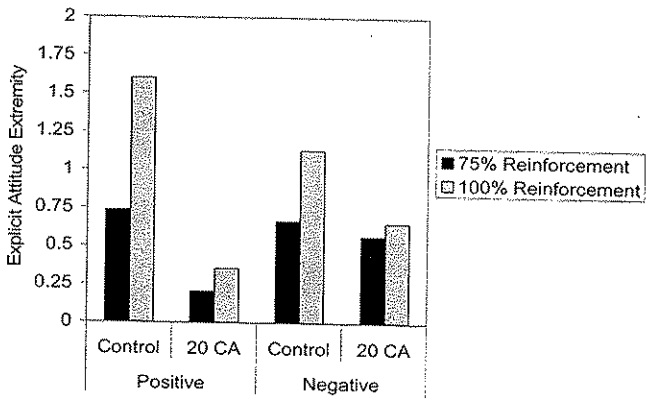

Figure 1. Explicit attitude extremity as a function of reinforcement and counterattitudinal condition (20 CA) in Experiment 1. Values for the negative initial learning condition have been reverse scored to reflect attitude extremity.

Implicit attitudes. As with the explicit attitude data, a 2 (valence of learned attitude) × 2 (level of reinforcement) × 2 (counterattitudinal condition) ANOVA was conducted on implicit attitude extremity (see Figure 2). In stark contrast to the explicit attitudes, the interaction of reinforcement and counterattitudinal condition and the interaction of valence of learned attitude and counterattitudinal condition were not significant for implicit attitudes ($Fs < 1$). In fact, the only effect to obtain for implicit attitudes was an effect showing the that grand mean was significantly different than zero, $F(1, 166) = 55.12$, $p < .001$ ($M = 0.50$, $SD = 0.87$). This shows that participants formed implicit attitudes about *Bob* in accordance with the valence of their initial learning but that subsequent counterattitudinal information had no impact on them. It is important that this effect was not statistically moderated by any of the experimental manipulations, showing no evidence of changes in attitude extremity or negative asymmetries for implicit attitudes.[6]

Discussion

A systems of reasoning conceptualization of attitude change was supported in this experiment because explicit attitudes were changed dramatically by the introduction of counterattitudinal information, whereas implicit attitudes were unaltered by this same information. This suggests that explicit attitudes are the product of a fast-learning system, whereas implicit attitudes reflect a slow-learning system. In this study, participants did form implicit attitudes about *Bob*, but,

[6] When implicit and explicit attitude measures were simply standardized (i.e., the standardized attitudes in the negative valence of learned attitude condition were not reverse scored) and submitted to a 2 (valence of learned attitude) × 2 (level of reinforcement) × 2 (counterattitudinal condition) × 2 (standardized attitude measure: implicit vs. explicit, a repeated measure) mixed-model factorial ANOVA, the expected four-way interaction was significant, $F(1, 156) = 3.98$, $p < .05$, reflecting differential responses to counterattitudinal feedback for explicit attitudes and implicit attitudes. In all subsequent experiments, similar omnibus analyses were conducted by using the standardized attitude measure as a within-subjects factor, and the highest order interaction obtained in each experiment ($Fs > 3.88$, $ps < .03$). These analyses reveal that examining implicit and explicit attitudes separately throughout the article is justified inferentially. In the current work, we present the data as examining attitude extremity by reverse scoring the negative learning condition attitude measures in order to simplify the presentation of how implicit and explicit attitudes are differentially affected by our manipulations.

tudes predict more deliberative, intentional behavior (e.g., Jellison et al., 2004; McConnell & Leibold, 2001). We were interested in whether the attitudes toward *Bob* created in the current experiments could predict behavior in the same manner as past research. Specifically, would explicit attitudes toward *Bob* only predict deliberate judgments about him but not predict more subtle forms of behavior (i.e., seating distance)? Similarly, would implicit attitudes toward *Bob* only predict subtle behaviors but not explicit judgments about him? Experiment 4 tests these predictions, anticipating unique predictive value for implicit and explicit attitudes.

These findings could be important for at least three additional reasons. First, past research has shown such double dissociations on the basis of measures of group prejudice (e.g., Dovidio et al., 2002; Jellison et al., 2004); however, this would be the first time that such effects have been shown for a different type of attitude object (i.e., a target person). Second, this previous work has shown these outcomes for groups with preexisting attitudes, whereas this would be the first study to demonstrate such dissociation effects on the basis of attitudes engineered in a controlled laboratory setting. For example, it is possible that cultural prescriptions might shape both implicit prejudice and subtle forms of social behavior toward social group members, providing the appearance of an attitude–behavior relation when, in fact, other factors may produce both. By engineering attitudes in the laboratory without any other target-relevant knowledge, it is far more likely that behavior reflects the influence of attitudes directly. Finally, if we show that implicit attitudes have unique predictive utility for subtle behavior in this study, then the findings would argue against concerns that our implicit measure has poor sensitivity. One might argue that slow changes on our implicit measure may reflect a relatively weak measure (i.e., it is simply less responsive to change than our explicit measures) rather than a slow-learning system. By establishing that our implicit (but not explicit) attitude measure can uniquely predict theoretically derived types of behavior, we could provide evidence inconsistent with a position that our implicit attitude measure is simply a poor measure.

Method

Participants. A sample of 29 undergraduates at Miami University participated in return for research credit in their introductory psychology courses. Participants were randomly assigned to receive no counterattitudinal information about *Bob* (control) or to receive 20 counterattitudinal pieces of information about *Bob* (20 CA).

Procedure. All materials, methods, and measures paralleled Experiment 1, with these exceptions. First, only the positive-valence condition was used, and only the 100% reinforcement condition was used. The two experimental conditions (control and 20 CA) were selected to maximize the discrepancy between implicit and explicit attitudes. In Experiment 1 there was a drastic change in explicit attitudes between the control and the 20 CA conditions, however there was no difference in implicit attitudes between them. Additionally, as in Experiment 3, because there was no negative initial learning condition to reverse score, greater standardized measures of attitudes reflected more positive attitudes toward *Bob*.

In addition to the attitude measures, participants completed explicit judgments of desire for social contact with *Bob*. Specifically, participants rated the extent to which they would want to have *Bob* as a neighbor, friend, classmate, roommate, and family member, each on 100-point scales ($\alpha = .92$). Greater scores on this measure indicated that they wanted more social contact with *Bob*.

After completing the attitude measures and the explicit social contact judgments, participants were told that they would "have a 2-min get acquainted session with Bob." They were escorted to a different room in which two chairs were set 221 cm apart. One chair had a book bag and a book next to it (where *Bob* was supposedly sitting), the other chair (for the participant) was on wheels and set against the wall of the room. The experimenter told each participant, "It looks like *Bob* has stepped out for a moment. Take that seat against the wall and move it so that you can have a face-to-face conversation with Bob." Participants took the seat and moved it into a position to converse with *Bob*. Afterward, they were told that they were not going to meet *Bob* and were then debriefed. The seating distance between the participant's chair and the chair where *Bob* had supposedly been sitting served as our measure of subtle, spontaneous behavior.

Results

The attitude measures were examined with one-way ANOVAs of counterattitudinal condition. The only effect to obtain was the predicted effect of counterattitudinal condition for explicit attitudes, $F(1, 27) = 12.86$, $p < .005$. Replicating the findings of Experiment 1, explicit attitudes were more positive in the control condition ($M = 0.48$, $SD = 0.80$) than in the 20 CA condition ($M = -0.51$, $SD = 0.77$), $F(1, 27) = 11.57$, $p < .005$. In contrast, implicit attitude data did not show an effect of counterattitudinal condition ($F < 1$).

The effect of counterattitudinal condition for social contact judgments was also examined with a one-way ANOVA. This analysis showed, as expected, that people reported wanting more social contact when they were in the control condition ($M = 74.53$, $SD = 15.83$) than when they were in the 20 CA condition ($M = 61.21$, $SD = 17.71$), $F(1, 27) = 4.57$, $p < .05$. Also, there was no effect of counterattitudinal condition on seating distance ($F < 1$). Thus, the counterattitudinal condition manipulation affected deliberate behavior (i.e., desire for social contact) but not the subtle behavior (i.e., seating distance).

To examine the main hypotheses, the correlation between explicit attitudes, implicit attitudes, deliberate behavior (i.e., desire for social contact), and subtle behavior (i.e., seating distance) were calculated. As expected, more positive explicit attitudes were related to greater desire for social contact ($r = .71$, $p < .001$) but were unrelated to seating distance ($r = .04$, *ns*). It is important that more positive implicit attitudes were unrelated to desire for social contact ($r = -.03$, *ns*) but were significantly related to closer seating distance ($r = -.41$, $p < .03$). Moreover, two multiple regressions analyses were conducted in which explicit and implicit attitudes served to predict desire for social contact (first analysis) and seating distance (second analysis). As predicted, explicit attitudes ($\beta = 0.70$, $p < .001$) but not implicit attitudes ($\beta = -0.01$, *ns*) predicted desire for social contact. On the other hand, implicit attitudes ($\beta = -0.41$, $p < .04$) but not explicit attitudes ($\beta = 0.02$, *ns*) predicted seating distance. Thus, explicit attitudes uniquely predicted deliberate judgments and implicit attitudes uniquely predicted subtle, spontaneous behaviors.

Discussion

Experiment 4 showed that the differential formation and change of implicit and explicit attitudes demonstrated in Experiments 1–3 have important implications for predicting behavior toward an attitude object, which in turn, reflect different systems of reason-