

# Implicit Ambivalence From Attitude Change: An Exploration of the PAST Model

Richard E. Petty  
Ohio State University

Zakary L. Tormala  
Indiana University

Pablo Briñol  
Universidad Autónoma de Madrid

W. Blair G. Jarvis  
Empirisoft Research Software

Traditional models of attitude change have assumed that when people appear to have changed their attitudes in response to new information, their old attitudes disappear and no longer have any impact. The present research suggests that when attitudes change, the old attitude can remain in memory and influence subsequent behavior. Four experiments are reported in which initial attitudes were created and then changed (or not) with new information. In each study, the authors demonstrate that when people undergo attitude change, their old and new attitudes can interact to produce evaluative responses consistent with a state of *implicit ambivalence*. In Study 1, individuals whose attitudes changed were more neutral on a measure of automatic evaluation. In Study 2, attitude change led people to show less confidence on an implicit but not an explicit measure. In Studies 3 and 4, people whose attitudes changed engaged in greater processing of attitude-relevant information than did individuals whose attitudes were not changed.

*Keywords:* attitude change, persuasion, ambivalence, implicit

The study of attitude change has one of the longest histories in social psychology (see McGuire, 1969). Literally hundreds of empirical investigations have examined ways to go about changing individuals' evaluations of people, objects, and issues (for reviews, see Eagly & Chaiken, 1993; Petty & Wegener, 1998). Yet there is little research to date that addresses the question of whether old attitudes are consequential when an attitude is changed or whether only the new (changed) attitude matters. Stated differently, if two people have the same attitude currently, does it matter if one of them used to feel differently whereas the other has always felt this way? This is the fundamental concern of the current research.

## Classic and Contemporary Approaches to Changed Attitudes

There are at least three approaches to the topic of old versus new attitudes that can be identified within the social psychological literature. The classic assumption guiding research for most of the period since the inception of the attitude construct is that when attitudes change, the old attitude disappears and is replaced with the new attitude. This assumption is clearly evident in Anderson's (1971) influential information integration theory. Consider the general equation guiding this framework:

$$\text{New Attitude} = [W_1(\text{old attitude}) + W_2(\text{new information})]/(W_1 + W_2).$$

According to this formula, a person's new attitude is a weighted average of the old attitude and the new information to which the person is exposed (and/or the new thoughts the person has about the attitude object; Anderson, 1981). Upon change, the old attitude ceases to exist because its informational value has been incorporated into the new attitude position. In brief, the underlying model begins with the widely shared assumption that representations of people, objects, and issues are tagged with evaluative nodes (e.g., good/bad) that can be retrieved upon presentation of the attitude object (e.g., Fazio, 1995; Fiske & Pavelchak, 1986). When the evaluation is changed, so, too, is the representation in memory. The traditional model of attitude change, as made explicit in the Anderson equation, is depicted in the top panel of Figure 1. Here, a person who used to like chocolate at Time 1 has replaced this favorable attitude with an unfavorable one at Time 2. If such a model were valid, then how an individual used to feel would no

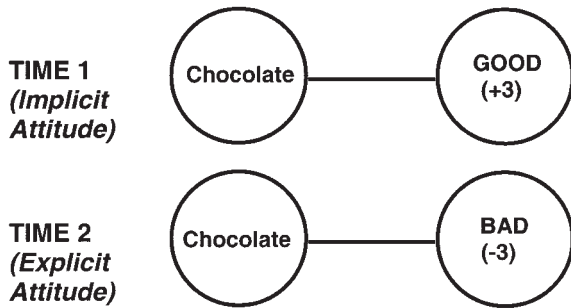
---

Richard E. Petty, Department of Psychology, Ohio State University; Zakary L. Tormala, Department of Psychological and Brain Sciences, Indiana University; Pablo Briñol, Department of Social Psychology and Methodology, Universidad Autónoma de Madrid, Madrid, Spain; W. Blair G. Jarvis, President, Empirisoft Research Software, Columbus, Ohio.

Portions of Study 1 are based on a doctoral dissertation by W. Blair G. Jarvis. This research was supported in part by Grant 752-94-0557 from the Social Sciences and Behavioral Research Council of Canada and Grant SBR 9520854 from the National Science Foundation. We thank John Cacioppo, Phillip Tetlock, Ed Crenshaw, Wendi Gardner, Brian Staub, Endel Tulving, Julie Barnes, and all the members of the 1996–2003 Groups for Attitudes and Persuasion at Ohio State University for taking the time to provide feedback on the conceptual and methodological issues raised in this article. We thank Tiffany Ito and Peter Lang for their help in acquiring the digital images used in this program of research. We also thank Nikolay Sergeevykh for his help with the data collection in Study 1 and Javier Horcajo and Ismael Gallardo for their assistance with Study 2.

Correspondence concerning this article should be addressed to Richard E. Petty, Department of Psychology, Ohio State University, 1885 Neil Avenue Mall, Columbus, OH 43210-1222. E-mail: petty.1@osu.edu

### Traditional (and *Dual Attitudes*) Models



### PAST Model (Past Attitudes are Still There)

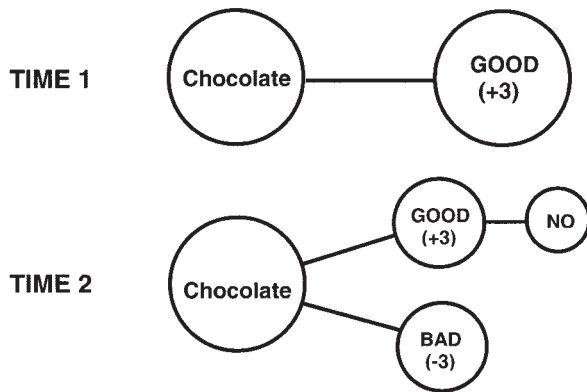


Figure 1. Top panel: Traditional and dual attitudes models of attitude change. Bottom panel: “Past attitudes are still there” (PAST) model of attitude change.

longer matter and would not have a significant influence on current evaluative responding.

An alternative to the traditional model is one that proposes that when attitudes are needed, they are largely constructed rather than retrieved from memory (e.g., Schwarz & Bohner, 2000; Wilson & Hodges, 1992). According to this model, if an evaluative response is required by the situation, people put one together on the spot on the basis of evaluative implications of any information they can retrieve and any contextual cues that are salient (e.g., the pleasant weather). Although there are various objections one might make to this model as a universal truth (e.g., see Fabrigar, MacDonald, & Wegener, 2005; Petty, Wheeler, & Tormala, 2003), the important point for our purposes is that this model, like the traditional one, holds that how an individual used to feel should have little bearing on current evaluative responding because attitudes are constructed anew on each occasion.

These views of attitude change have been challenged recently. Specifically, just as researchers have provided convincing support for the view that past behaviors can be influential in determining current behavior (e.g., Bentler & Speckart, 1979; Leone, Perugini, & Ercolani, 1999; for a review, see Ouellette

& Wood, 1998), researchers have begun to suggest that prior attitudes can also exert impact on current responses above and beyond current attitudes (e.g., Cacioppo, Marshall-Goodell, Tassinari, & Petty, 1992; Petty, Baker, & Gleicher, 1991; Petty, Gleicher, & Jarvis, 1993; Wilson, Lindsey, & Schooler, 2000). This should not happen according to the traditional model (e.g., Anderson, 1971) or the constructivist view (e.g., Schwarz & Bohner, 2000) because these perspectives stipulate that prior evaluations are incorporated into the current evaluation and have no independent existence. If prior attitudes can influence responses above and beyond current attitudes, this may suggest that when attitude change occurs, the old attitude does not disappear but rather continues to exist and to have an impact on current evaluative responding.

A recent instantiation of this idea with respect to attitude change can be found in the dual attitudes model proposed by Wilson et al. (2000). This model suggests the coexistence of two attitude systems that allow people to hold dissociated attitudes toward the same object (see also Devine, 1989; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Greenwald & Banaji, 1995). According to the dual attitudes model, when attitude change occurs, an attitude that was originally quite explicit and conscious may remain intact and be activated automatically in certain situations without the person’s awareness, thus becoming an implicit attitude (see Figure 1, top). Furthermore, consistent with much thinking about implicit and explicit attitudes (e.g., Dovidio et al., 1997), the dual attitudes model proposes that one’s new explicit attitude guides responding when an individual has sufficient time to reflect on his or her evaluations, but one’s old (implicit) attitude is more likely to have an impact in situations that require an individual to respond under time pressure or without thinking about his or her attitudes. Thus, either one’s old or one’s new attitude can be influential depending on the circumstances. As stated by Wilson et al. (2000),

Because the implicit attitude is habitual and automatic, it is the default response that is expressed when people do not have the capacity or motivation to retrieve the more recent attitude. The explicit attitude is expressed and acted on when people have the motivation and cognitive capacity to retrieve it. (p. 104)

Other researchers who have examined these issues (e.g., Greenwald & Banaji, 1995; for a review, see Dovidio, Kawakami, & Beach, 2001) have also typically assumed that when dual attitudes are formed, the resulting attitudes—called implicit and explicit—constitute independent evaluations that operate under different circumstances (see also Gaertner & Dovidio, 1986). Depending on situational factors, people will be guided by either the implicit attitude or the explicit one. In the dual attitudes conceptualization, implicit attitudes are typically assessed with measures of automatic evaluative association, such as the sequential priming measure (Fazio, Sanbonmatsu, Powell, & Kardes, 1986) or the implicit association test (IAT; Greenwald, McGhee, & Schwartz, 1998), whereas explicit attitudes are assessed with more deliberative measures, such as the semantic differential (Osgood, Suci, & Tannenbaum, 1957). In the present article, we use the terms *deliberative* and *automatic* attitudes and attitude measures to pro-

vide a closer correspondence between conceptualization and measurement.<sup>1</sup>

In fact, there is now a good deal of support for the notion that deliberative and automatic attitude measures can tap different evaluations and that these different evaluations can guide different behaviors in different situations (e.g., see Dovidio et al., 2001). Of particular relevance for the current research are data from a series of studies reported by Wilson et al. (2000) in which participants experienced attitude change through a conventional manipulation, after which their attitudes were assessed with or without diminished cognitive capacity (e.g., time pressure). Wilson et al. found that attitude reports under time pressure were more consistent with participants' prior attitudes, whereas attitude reports without time pressure were more consistent with participants' current attitudes. Also consistent with the position that dual attitudes can operate entirely independently is work on dual memory systems (see McClelland, McNaughton, & O'Reilly, 1995; Smith, 1998; Smith & DeCoster, 2000) and evidence for the physiological and motivational independence of positive and negative evaluative responding (for a review, see Cacioppo, Gardner, & Berntson, 1997).

### The PAST Model

In this article, we provide a meta-cognitive perspective on old and new attitudes, depicted in the bottom panel of Figure 1. This framework is called the "past attitudes are still there" (PAST) model (Jarvis, Tormala, & Petty, 1998; Petty & Jarvis, 1998; see also Jarvis, 1998; Petty et al., 2003). In brief, the PAST model holds that when prior attitudes are no longer considered appropriate, the individual encodes an association with the rejected evaluation that essentially tags it as false or wrong. This proposal is analogous to Gilbert and colleagues' suggestion (following the philosopher Spinoza over Descartes) that information initially held as true needs to be tagged as false to be disbelieved (Gilbert, 1991; Gilbert, Krull, & Malone, 1990; Gilbert, Tafarodi, & Malone, 1993). Only if the tag is retrieved will a person who disbelieves an assertion recognize it as false. Otherwise, the person may misremember and act on it as if it were true.

According to the PAST model as well, people who have rejected a prior attitude as false must retrieve this tag for the old evaluative habit not to operate, at least in the initial period following change. Otherwise, the prior attitude may still exert an impact on current responding. Notably, the PAST model does not require the Spinozan postulate (i.e., that all propositions are encoded as true and then subsequently disbelieved if the evidence warrants; cf. Hasson, Simmons, & Todorov, 2005) to be true in all situations. Rather, the PAST model simply asserts that when a previously held belief is tagged as false, it can later operate as if it were true if the false tag is not retrieved. According to the PAST model, prior attitudes would be most likely to impact current evaluative responding if the link between the attitude object and the prior attitude is stronger than the link between the attitude object and the current attitude (e.g., due to greater practice of the former link than the latter) or if the link to the false tag is weak or inaccessible (e.g., due to time constraints or lack of effort to retrieve it).

Research on cognitive negation clearly suggests that negation can be hard to implement and may sometimes fail (e.g., Mayo, Schul, & Burnstein, 2004). When negation fails, the original belief can exert an influence. Negation can fail for several reasons. First,

and most simply, the negation may be forgotten or may decay faster than the original assertion. This is typically given as the reason that messages (e.g., the sleeper effect; see Kumkale & Albarracín, 2004) or experimental information (e.g., the belief perseverance effect; e.g., L. Ross, Lepper, & Hubbard, 1975) labeled as false can still influence judgments. People have simply forgotten the negation but not the original information. Second, whenever people attempt to implement a negation, they also make the original assertion more accessible. In fact, Maio and Olson (1995) found that asking people to lie about their attitudes made the correct attitude more accessible. Similarly, trying to suppress a thought (e.g., "I don't want to believe that anymore") may make the original thought more accessible (e.g., see Wegner, 1989).

Finally, and perhaps most relevant, prior research suggests that even when a person is able to retrieve the negation when engaged in deliberative thinking (i.e., the negation is not lost from memory or has not decayed), the negation tag still might not be retrieved spontaneously. In one study, for example, Tybout, Calder, and Sternthal (1981) exposed people to a rumor about McDonald's hamburgers being made with worms. Even though participants reported that they believed the rumor to be false, this information had a negative impact on subsequent judgments of McDonald's compared with individuals who were not exposed to the rumor and those who were exposed but were provided with retrieval cues for their disbelief.

In short, the PAST model differs from the traditional and constructivist views of attitude change in holding that the old attitude still exists in some form and is thus available to influence responses. The PAST model shares this assumption with the dual attitudes view, but it also differs from this framework in important ways. Most notably, the PAST model suggests a more integrated relationship between old and new attitudes (see Figure 1, bottom). The PAST model holds that both old and new attitudes are linked to the same attitude object but that the old evaluation is tagged with a false label. Because both old and new evaluations are linked to the same attitude object, either or both can be activated at any point in time, depending on the strength of the linkage to the attitude object and the retrievability of the false tag.

### Implicit Ambivalence

As should be clear from our discussion so far, one fundamental prediction of the PAST model that differs at least in emphasis from prior conceptualizations of attitude change is that the joint activation of old and new attitudes is possible, and to the extent that the false tag for the old attitude is not retrieved, an ambivalent-like state should exist. This ambivalence differs from past conceptualizations of attitudinal ambivalence in a number of ways. First, traditional research on attitudinal ambivalence has focused on particular kinds of conflict—conflict between the consciously endorsed positive and negative aspects of an attitude object (e.g.,

<sup>1</sup> The confusion in terminology arises because some implicit measures (i.e., measures of which people are unaware) can be very deliberative (e.g., see Vargas, von Hippel, & Petty, 2004), and some explicit measures can be administered under time pressure, making them less likely to tap deliberative responses (e.g., see Wilson et al., 2000). For the current research, it is the deliberate-automatic distinction that is more critical for the hypotheses (for further discussion, see Fazio & Olson, 2003; Petty et al., 2003).

Kaplan, 1972) or between one's own attitude and the presumed attitudes of significant others (Priester & Petty, 2001). The PAST model focuses on conflict between one's newly endorsed attitude and the old rejected attitude.

A second issue concerns awareness of the ambivalence. Often people will be aware that their old attitude is different from their new one (though this need not be the case; see M. Ross & Conway, 1986). But are they aware of any ambivalence resulting from this discrepancy? In some cases of attitude change, people might well experience some explicit conflict or tension between their old and new views. For example, people may be aware that their attitude toward smoking has changed from positive to negative, but when they find themselves automatically reaching for a pack of cigarettes, they may realize that their behavior contradicts their anti-smoking position and suggests an underlying positivity toward cigarettes. This consciously recognized contradiction can cause explicit feelings of ambivalence. Such a situation is analogous to the conceptualization offered by Devine, Monteith, and colleagues with respect to racial prejudice. In a series of studies (e.g., Devine, Monteith, Zuwerink, & Elliot, 1991; Monteith, 1993; Monteith, Devine, & Zuwerink, 1993), they have argued that egalitarian individuals recognize that they sometimes have spontaneous negative feelings toward Blacks or engage in prejudicial behavior. When this conflict is brought into consciousness, there is enhanced cognitive activity designed to prevent the prejudice (see also Petty, Fleming, & White, 1999).

The focus of the current research is on the more common and mundane attitude change situation, which, we suspect, does not involve any explicit conflict once attitudes have changed. For example, consider a person who had one pleasant movie date with another individual and then discovers that the person is a child molester. The initial positive attitude turns to a negative one, but it seems unlikely that the person would have any ambivalence about the new negative attitude. Or imagine that a person formed a very negative impression of another person on the basis of a résumé only to find out that the résumé was put in the wrong folder and, thus, the information was not appropriately attributed to that person. The correct résumé has very positive information. Here, a negative impression turns into a positive one, but there is no logical reason to be conflicted about the new impression. Nevertheless, the PAST model holds that ambivalent-like responses are possible if the original evaluation is tagged as false and the false tag to the original impression is not easily accessible. Because people would not consciously experience conflict or doubt regarding the attitude object, however, we label such conflict, if it exists, *implicit ambivalence*. This ambivalence stems from the unrecognized conflict between the rejected old attitude and the endorsed new attitude. People may well recognize the discrepancy between their old and new attitudes (e.g., "I used to like her, but now I do not"), but it is possible that they do not recognize any conscious conflict or doubt with respect to the attitude object.

### Overview of the Current Research

The primary goal of the current research was to examine three postulated implications of the notion that people whose attitudes have changed will show signs of ambivalence compared with people who hold the same attitudes but with no change. Although there are many possible consequences of ambivalence (for a re-

view, see Thompson, Zanna, & Griffin, 1995), we focused on three common ones. First, ambivalent individuals tend to appear more neutral in their global evaluations than do unambivalent individuals because of the joint activation (or inhibition) of positive and negative information (e.g., Bargh, Chaiken, Govender, & Pratto, 1992; Kaplan, 1972). We did not expect individuals who changed their attitudes to demonstrate this neutrality on a deliberative attitude measure when there was sufficient time for reflection. We did, however, expect to see greater neutrality on a measure of automatic evaluation. Because the newly acquired false tag would not be spontaneously accessible, both positive and negative reactions would jointly influence responses, producing an overall neutrality.

Second, individuals who are ambivalent typically report greater conflict, doubt, or mixed feelings with respect to the attitude object (Priester & Petty, 1996; Thompson et al., 1995). As explained earlier, we did not expect people whose attitudes had changed to feel any conscious conflict or doubt about the attitude object. Indeed, if their attitudes sincerely changed from one evaluation to another, they would accept the new attitude as their own. However, we postulated that a more implicit measure of conflict might show signs of doubt that would not be expressed consciously.

Third, people who are ambivalent tend to engage in increased information processing with respect to the attitude object, presumably in an attempt to resolve their ambivalence (e.g., Maio, Bell, & Esses, 1996). If there is some ambivalence lurking below the surface for individuals whose attitudes have changed, they may engage in greater information processing with respect to the attitude object compared with people who have not changed.

We tested these ideas across four studies. In each study, we created initial attitudes and then changed them (or not). The goal was to look for the three signs of ambivalence noted above in comparing individuals whose attitudes had changed with those who always had the same attitude. In Study 1, we looked for evidence of neutrality on a measure of automatic evaluation but not on a measure of deliberative attitudes. In Study 2, we sought evidence of doubt on an implicit but not on an explicit measure. In Studies 3 and 4, we tested the idea that people whose attitudes have changed would engage in greater information processing with respect to the attitude object.

Across the four studies we used two different paradigms to induce attitude change. In Studies 1 and 3, we used a classical conditioning procedure (e.g., see Staats & Staats, 1958) to create an initial positive or negative attitude, and then we changed these attitudes (or not) by using an attitude similarity manipulation (see Byrne, 1971). In Studies 2 and 4, we first presented participants with positive or negative trait information about a target person (Study 2) or with information indicating that the target person had similar or dissimilar attitudes (Study 4). We then led individuals in the attitude change condition to believe that they had received the wrong information so that their attitudes would have to be reversed. Across experiments, we looked for signs of implicit ambivalence in the change versus no-change groups.

### Study 1

In Study 1, we examined the possibility that if old and new attitudes are opposite in valence, a state of ambivalence can result such that when there is little time for reflection (i.e., on a measure

of automatic evaluation), individuals will appear to be relatively neutral in their evaluative responses. As noted by Wilson et al. (2000), old attitudes, because of their prior development and use, are likely to be more accessible than are newly developed attitudes. Furthermore, in the absence of extensive practice negating one's old attitude (e.g., see Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000), retrieval of the false tag for the old attitude may be unlikely, especially in situations allowing for little thought. Thus, according to the PAST model, under conditions that allow for very thoughtful responding, people are most likely to report their new attitudes when asked. Under conditions that require quick evaluation, however, the old attitude may come to mind along with the new one (or the two attitudes could inhibit each other), producing an ambivalent-like (i.e., neutral) response. In sum, the prediction of the PAST model tested in Study 1 was that although the deliberative (explicit) attitudes of both groups (i.e., changed and unchanged) at the current time would be equivalent, this equivalence would mask differences in ambivalent responding on a more automatic measure (i.e., a sequential priming paradigm; Fazio et al., 1986).

### *Method*

#### *Participants and Overview of Procedure*

Participants were 98 undergraduates enrolled in introductory psychology courses at Ohio State University. Sessions took place in a single room with four partitioned computer workstations. When participants first arrived in the lab, they were instructed that we were conducting a variety of tasks, and they were assured that all of their responses were anonymous and would not be associated with them in any way. The first task was described as a basic survey regarding several social and political issues. Participants reported their attitudes toward 12 issues on a 7-point scale ranging from  $-3$  (*very much against*) to  $3$  (*very much in favor*). Responses to these items were used to develop the similarity manipulation that either reinforced or changed initial attitudes toward one of two target individuals (Eddie or Phil). Following this task, participants engaged in a classical conditioning procedure designed to establish initial attitudes toward the target individual. After the conditioning procedure, participants reported their explicit attitudes toward the target individual, and then they received attitude information about the target individual that either left their initial attitudes intact or changed them to the opposite valence.

Following the attitude change manipulation, participants once again reported their attitudes toward the target individual on a deliberative (explicit) measure and then a more automatic (implicit) measure. After this, participants were asked to make an explicit behavioral choice involving the target individual. At the completion of the study, all participants were probed for any suspicion regarding the experiment and were thoroughly debriefed. All participants appeared to take the procedures at face value, and none articulated the thought that the experiment concerned ambivalence. The software used to create and run the experiment was MediaLab Version 3.0 and DirectRT Version 1.0 (Jarvis, 1997).

#### *Initial (Prior) Attitude Induction: Classical Conditioning Procedure*

In the conditioning procedure that followed completion of the attitude questionnaire, participants were instructed that the primary purpose of the experiment was to examine issues having to do with visual memory. These instructions were intended to draw participants' attention away from the purpose of the conditioning task (see Zajonc, 1968). They were then instructed that a series of color images would be appearing on their

computer screens and that they were to try to remember the images.<sup>2</sup> Participants were told to turn to their computer screens where the same instructions were presented again with additional information to explain the appearance of a facial picture before each image. They received the following information:

You may or may not see the word "focus" or a small black and white picture before each image. We are examining how these affect your visual memory for the color photographs. Please pay careful attention to the screen at all times.

To create initial attitudes, we used a conditioning procedure developed through extensive pretesting. The conditioning procedure was validated to ensure that our method produced the appropriate positive and negative attitudes toward the target faces on both the deliberative and the automatic measures of attitudes used for this experiment and described in more detail below (see Jarvis, 1998, for further information about the pretesting).

The unconditioned stimulus images used to elicit evaluative reactions were chosen from the International Affective Picture System (Lang, Bradley, & Cuthbert, 1995), which contains approximately 500 digital color images ranging from extremely negative (e.g., mutilated bodies) to extremely positive (e.g., puppies). On the basis of our own pretesting and work by Ito, Cacioppo, and Lang (1998), we selected 70 of the most uniformly negative, neutral, and positive images for a total set of 210 images. On the basis of reaction time data and separate positivity and negativity ratings from our own pretesting, we avoided images that tended to elicit slow evaluative reactions (less than two standard deviations above the mean reaction time) and/or explicitly ambivalent responses (Thompson et al., 1995).

The conditioned stimuli in the conditioning trials were a set of 24 heterogeneous White male faces ranging moderately on various dimensions of appearance, such as age, attractiveness, and apparent socioeconomic status. This set was selected from an initial pool of 120 tested in an earlier study. The 24 faces chosen were rated as the most evaluatively neutral of the set (all ranging from 3.75 to 4.25 on a 7-point scale). Positivity and negativity were also both assessed by using measures that are described shortly, and only those faces that scored neutrally on both dimensions were used. From the set of 24 faces, two of them (soon to be known as Eddie and Phil) were randomly chosen to serve as the primary targets for all participants. The faces of Eddie and Phil were always paired 40 times with either positive or negative images. For half of the participants, Eddie was paired with negative images and Phil was paired with positive images. This assignment was counterbalanced for the other half of the participants. Counterbalancing produced no significant effects and is not discussed further.

To reduce awareness of the conditioned stimulus–unconditioned stimulus contingencies, the remaining 22 faces were used for filler trials. These filler faces ranged in frequency from 5 to 20 exposures but otherwise were identical in format to the target trials. The filler trials were balanced such that an equal number of filler faces were seen 5, 10, and 20 times each, and an equal number were paired with negative, neutral, and positive images. Filler faces were used to distract participants from the main purpose of the experiment, so we do not refer to them further. All trials were presented in a random order with the only constraint being that the same face would not

<sup>2</sup> Participants were informed that the task was challenging and involved some negative images and that they had the right to discontinue their participation at any point if they so desired. To accommodate any desire to withdraw, experimental credit was always given before the session began, and freedom to leave at any time was stressed. Two participants did decline to finish the session. The sample size does not include these two participants. Through the debriefing and feedback procedure, we learned that most participants thought the purpose of the research justified any temporary discomfort caused by the images.

appear within any run of three trials. On each trial, participants saw the word *FOCUS* centered on the screen for approximately 500 ms. Next, they saw a face presented for approximately 600 ms, the unconditioned stimulus image for 1,000 ms, and then a blank screen for 1,000 ms.

### *Deliberative (Explicit) Attitude Measure Following Conditioning*

Following the conditioning trials, participants were informed that we were going to give the information a chance to “settle in” before we started the visual memory task. Participants were then instructed to rate each of the faces to follow in terms of how good a man they thought he might be (meaning that he is “probably good natured, helps people, and cares about other people”) and how bad a man they thought he might be (meaning that he is “probably bad natured, does not help people, and does not care about other people”). All faces were rated together in a random order on one dimension (e.g., positive), and then again on the other (e.g., negative). The order of the two sets of ratings was random for each participant. For each set of ratings, the faces were presented sequentially on the computer, with the face presented to the right of a vertically displayed rating scale. The scale ranged from 1 (*not at all good [bad]*) to 7 (*very good [bad]*). Participants used the keyboard to enter their ratings.

### *Attitude Change (Similarity) Manipulation*

Following the explicit ratings of the faces used in the conditioning procedure (including Eddie and Phil), participants were given additional instructions that led them into the attitude change procedure. We explained that we were interested in knowing what kind of impressions people would form of the men in the pictures if they had some personal information about them. This personal information was how these men felt about various social and political issues. Participants were told that they would be reading some survey items, much like the ones they themselves had earlier completed, that had been filled out by two individuals. They were asked to try to form impressions of these people after viewing their responses. Participants were given 60 s to examine a picture of each person (Eddie and Phil) and the attitude survey completed by that person (i.e., the opinion data of Eddie and Phil). Each picture was displayed on a computer screen next to the person’s ostensible responses to 12 survey items. The survey items were selected to deal with issues important to the participants so that the conditioned attitudes could easily be changed. The issues concerned opinions regarding fraternities and sororities, belief in God, money as life’s ultimate goal, smoking in public, premarital sex, political party affiliation, affirmative action, lesbians and gay men in the military, the death penalty, freedom of choice with abortion, spending tax money to protect the environment, and sex education in school.

The similarity procedure was based on Byrne’s (1971) work showing that people have more favorable attitudes toward similar others than dissimilar others.<sup>3</sup> In accord with the Byrne paradigm, when the target was similar, his ostensible survey responses matched eight of the participant’s responses exactly. The four nonidentical responses were matched in valence but shifted one point on the scale (in the more extreme direction if possible). It was thought that this occasional shift would make the high degree of similarity seem more believable. When the target was dissimilar, his responses were the exact opposite of eight of the participant’s responses (e.g., if the participant had reported a 2, then the target’s attitude would be reported as –2). Again, to avoid a suspicious looking set of responses, the four remaining responses were generated in a similar manner but were shifted one point (in the opposite direction if possible; e.g., –2 became 3).

Of the two target people (Eddie and Phil), one was similar and one was dissimilar for each participant. Assignment of which target was similar and which was dissimilar depended on the initial attitudes induced through the conditioning procedure. Half of the participants were randomly assigned to the *congruent prior attitude* condition in which the similar target was the

one who previously had been associated with positive images and the dissimilar target was the one who had been associated with negative images. In these conditions, attitudes were not expected to change. The other half of the participants were randomly assigned to the *incongruent prior attitude* condition in which the similar target had been conditioned with negative images and the dissimilar target had been conditioned with positive images. In these conditions, attitudes were expected to change.

### *Deliberative (Explicit) Attitude Measure Following Similarity Induction*

After the similarity manipulation, participants were asked to rate Eddie and Phil on the same set of rating scales already used (i.e., ratings of how good and how bad a man they thought each target might be).

### *Automatic (Implicit) Attitude Measure*

Following the similarity manipulation and the deliberative self-report attitude assessment, participants were instructed that it was now time to begin the visual memory task. At this point, participants’ attitudes were assessed with an automatic (implicit) attitude measure based on the Fazio et al. (1986) evaluative priming paradigm. Participants were asked to rate target images or words as being either good or bad, and reaction times were recorded. Before each trial, we briefly presented one of the conditioned faces to assess its impact on response times. Participants were told to press the button marked “good” on their keyboard if the target image or word made them feel good and to press the button marked “bad” if the target image or word made them feel bad. They were told to be “fast and accurate” and that they would see a face before each image or word to be rated to “help focus attention.” The rationale for this procedure is that if a person has a positive evaluation associated with the face in memory, then seeing that face briefly should facilitate the identification of a subsequent positive stimulus as good and inhibit the identification of a subsequent negative stimulus as bad. The opposite pattern would be expected for a face associated with a negative evaluation (for additional details, see Fazio, Jackson, Dunton, & Williams, 1995).

*Priming task design.* The target stimuli used in the priming task consisted of either an image or a word that appeared in the center of the computer screen. For half of the participants, the target stimuli consisted of images, and for half, they were words. Images were drawn from the same pool of images used in the previous conditioning trials. Words were used for half of the participants to eliminate the alternative explanation that our effects were dependent on using the same images to which participants had already been exposed. Essentially, the image and word groups provide an internal replication of the basic design. Target words were drawn partly from the list of words used by Bargh et al. (1992) and partly from our own pretesting. For critical trials, only words that were clearly positive (e.g., *baby, flowers, and puppy*) or negative (e.g., *bombs, death, and spider*) were used. Neutral words were also used in filler trials to reduce suspicion. In the priming task, the faces of Eddie and Phil each served as a prime six times: three times followed by a positive stimulus (P) and three times followed by a negative stimulus (N). One face from the pair served as a prime for target stimuli in the order NPNPNP, and the other in the order PNPNNP. This design resulted in a total of 12 trials. To reduce participants’ focus on the two primary faces, the remaining 22 faces from the conditioning trials were used as filler trials. This resulted in 132 filler trials. Because trials associated with the filler faces were used to distract partic-

<sup>3</sup> Like the conditioning procedure, this manipulation was pretested extensively to ensure that attitudes (both deliberative and automatic) were more positive when people were led to believe that Eddie or Phil was similar and more negative when they were led to believe that Eddie or Phil was dissimilar.

ipants from the main purpose of the experiment, we do not refer to them further. Target stimuli for the priming trials were assigned randomly.

**Priming trials.** On each trial, the word *FOCUS* was centered and displayed for a 2.5-s interval in the center of the computer screen. Following Fazio et al. (1995), all primes were presented for 315 ms, followed by a 135-ms interval before onset of the target stimulus. Each face served as a prime for negative stimuli in half the trials and as a prime for positive stimuli in the remaining trials. To encourage both speed and accuracy, participants were shown a prompt screen reading “Please try to respond faster” if their response time was greater than 1,500 ms on any given trial, and “Please wait for the image to appear” if their response time was less than 200 ms. This screen was displayed for 2.5 s before the trials resumed.

### Choice Measure

Finally, a forced choice measure was included at the conclusion of the experiment. This assessment was to serve as a proxy for a behavioral measure of participants’ evaluations of Eddie and Phil. Specifically, participants were led to believe that Eddie and Phil were researchers in our lab and that they would be conducting follow-up interviews for this research. They were then instructed to select either Eddie or Phil as the interviewer for the postexperiment interview session. It also provided a second deliberative measure of evaluation following the automatic one.

### Summary

In sum, the full experiment was a 2 (conditioning: negative vs. positive)  $\times$  2 (similarity: similar vs. dissimilar)  $\times$  2 (counterbalancing: Eddie conditioned positive and Phil conditioned negative vs. Phil conditioned positive and Eddie conditioned negative)  $\times$  2 (priming measure: images vs. words) between-participants design. To simplify the design, for some of the analyses we turned the Conditioning  $\times$  Similarity interaction into a more conceptually meaningful congruent versus incongruent main effect. That is, when participants were conditioned to be positive toward the target (Eddie or Phil) and received the similar attitude information, and when they were conditioned to be negative and received the dissimilar attitude information, they were in the congruent condition because the conditioning manipulation should produce the same attitude as the similarity manipulation (i.e., no attitude change). The other conditions (i.e., positive conditioning with dissimilar attitude information and negative conditioning with similar attitude information) were incongruent because the similarity information was expected to change the attitude formed via conditioning. Also, because we were not interested in differences between Phil and Eddie per se, and because preliminary analyses using the full design revealed no significant results for this factor, we do not discuss this variable further. In addition, the manipulation of priming measure (i.e., images or words) did not impact response latencies, nor did it moderate any of the effects reported in the *Results* section. Thus, we do not discuss this internal replication factor further.

### Results

Figure 2 presents participants’ deliberative (explicit) evaluations of the attitude objects (Eddie and Phil) immediately following the conditioning procedure and then immediately following the similarity manipulation designed to either reinforce or change initial attitudes. The top panel presents the means for the congruent (no attitude change) conditions, and the bottom panel presents the means for the incongruent (attitude change) conditions.

#### Postconditioning Deliberative Attitudes

We first needed to ensure that the deliberative (explicit) attitudes of our congruent and incongruent samples were equally

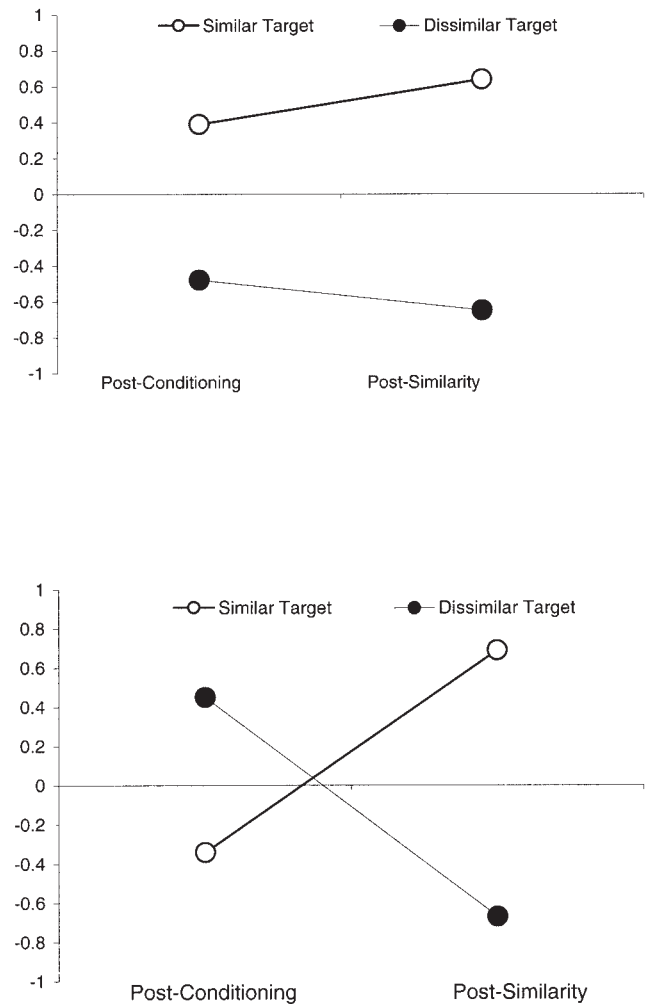


Figure 2. Top panel: Standardized deliberative attitude scores in the congruent condition. In this condition, similar targets were initially conditioned to be positive, and dissimilar targets were initially conditioned to be negative. Bottom panel: Standardized deliberative attitude scores in the incongruent condition. In this condition, similar targets were initially conditioned to be negative, and dissimilar targets were initially conditioned to be positive (Study 1).

affected by the initial conditioning manipulation (see postconditioning means in Figure 2). To examine this, we calculated participants’ initial attitudes toward Eddie and Phil by averaging the two self-report attitude measures taken immediately following conditioning (“How good . . .” and “How bad . . .”) with the negative judgment reverse scored such that higher numbers reflected more positive evaluations. Attitudes were then standardized to a mean of 0 with a standard deviation of 1. We assigned a value to each of the two resulting attitudes, which corresponded to whether the face served as the similar or dissimilar target in the subsequent similarity manipulation.

The analysis of variance (ANOVA) based on the full design revealed a main effect for conditioning such that negative conditioning produced more negative attitudes ( $M = -0.41$ ) than positive conditioning ( $M = 0.42$ ),  $F(1, 94) = 36.56$ ,  $p < .001$ . This

main effect was not qualified by any interactions (all  $F_s < 1$ ). It is important to note that this implied that the positive attitudes and negative attitudes created by the conditioning paradigm were equivalent across our two key experimental groups (i.e., congruent and incongruent).

### *Postsimilarity Deliberative Attitudes*

Next, it was important to establish that the similarity manipulation was effective in changing the attitudes of participants in the incongruent condition such that their new attitudes appeared to be equivalent to those of participants in the congruent condition (see postsimilarity means in Figure 2). To do this, we first calculated participants' new attitudes toward Eddie and Phil after the similarity manipulation. Again, this was done by averaging the two attitude scales (with the negative judgment reverse scored) and then standardizing the resulting attitudes. We again assigned a value to each of the two resulting attitudes, which corresponded to whether the face served as the similar or dissimilar target.

The ANOVA from the full design revealed a main effect for similarity such that the dissimilar target produced more negative attitudes ( $M = -0.66$ ) than the similar target ( $M = 0.66$ ),  $F(1, 94) = 146.40, p < .001$ . It is important to note that this main effect was not qualified by any interactions. Thus, the similar target was viewed more positively than the dissimilar target regardless of whether participants held a congruent or incongruent prior attitude.

### *Attitude Strength Indicators*

Although we demonstrated that current attitudes (i.e., following both the conditioning and similarity procedures) were equivalent in terms of valence and extremity regardless of whether they had changed, it seemed possible that an incongruent prior attitude could influence the strength of one's current attitude (e.g., see Petty & Krosnick, 1995). To demonstrate the equivalence of congruent and incongruent attitudes in strength, we assessed attitude strength on two indicators that seemed especially relevant for our paradigm—attitude accessibility (Fazio, 1995) and objective explicit ambivalence (Kaplan, 1972). Fazio (1993) demonstrated that attitudes of comparable extremity could produce differential priming patterns if they varied in their accessibility. Specifically, an attitude that comes to mind more quickly will produce a stronger priming effect than one that does not. Therefore, we wanted to examine whether having an incongruent prior attitude had an impact on the accessibility of current attitudes.

Participants' standardized response times to the two attitude scales were averaged to serve as an indicator of attitude accessibility. The resulting accessibility scores reflected greater accessibility (i.e., shorter response times) with lower numbers. Using the same ANOVA design as for explicit attitude extremity above, this analysis revealed no main effects or interactions ( $p_s > .18$ ), indicating that participants with incongruent prior attitudes evidenced attitude accessibility scores equivalent to those with congruent prior attitudes.

Second, objective attitudinal ambivalence represents a state of explicitly endorsing both positivity and negativity toward an object. Because two attitudes can be of equal extremity yet differ in their extent of objective ambivalence, we thought it was important to demonstrate that participants' current attitudes (i.e., postsimi-

larity induction) did not differ in this regard. We used a variety of recommended calculations to approximate the extent of objective attitudinal ambivalence (i.e., Kaplan, 1972; Priester & Petty, 1996; Thompson et al., 1995). We report the results using the Kaplan (1972) procedure, but the results were highly consistent across all measures. In the Kaplan calculation, the lower of the two (positivity and negativity) attitude extremity judgments (i.e., the extent of "conflicting reactions"; see Priester & Petty, 1996) serves as the index of ambivalence. For example, if a participant rated the target as 7 on "how good" and 4 on "how bad" he was, then this participant's ambivalence score would be 4. Using the same ANOVA design as for the attitude extremity and accessibility analyses, this analysis revealed no main effects or interactions ( $p_s > .13$ ). Thus, participants with incongruent prior attitudes ( $M = 2.36$ ) evidenced ambivalence scores equivalent to those with congruent prior attitudes ( $M = 2.47$ ).

### *Automatic (Implicit) Attitude Measure*

At this point, our paradigm had produced two groups of individuals. Both groups had the same attitude at present: favorable in the similarity case and unfavorable in the dissimilarity case. But the incongruent group had current attitudes that conflicted with prior attitudes, and the congruent group had current attitudes that did not conflict with prior attitudes. In fact, the conflict with (or change from) prior attitudes was the only way in which the congruent and incongruent groups differed. Indeed, these groups had deliberative attitudes with the same valence, extremity, accessibility, and objective level of ambivalence.

*Overall automatic attitudes.* To examine whether any evidence for ambivalent-like responding could be found, we looked at the results from the automatic attitude measure. Before submitting the raw data to analysis, we assigned missing values to response times greater than 1,500 ms and to those less than 200 ms. Together, such occurrences accounted for less than 1% of all trials (0.7%). Additionally, we excluded response times for trials on which participants responded with the incorrect key (e.g., indicating "good" to a bad word or image, or vice versa). This was consistent with the error rate reported by Fazio et al. (1995), with such trials accounting for approximately 4% of the total.

For each of our two faces, we had three "good" and three "bad" response times. Because positive and negative attitudes were each expected to influence response times to good and bad targets in opposite directions, we reverse scored participants' standardized reaction times to the bad targets such that higher scores would reflect more positive evaluations of the prime. We then averaged the six response times from each participant for each face. This new variable now reflected greater positivity toward the prime with high values and greater negativity with low values.

Figure 3 presents the attitudes toward the similar and dissimilar targets separately for participants in the congruent and incongruent prior attitudes conditions. The attitude data were submitted to a 2 (prior attitude: congruent vs. incongruent)  $\times$  2 (prime: similar vs. dissimilar) mixed ANOVA with the second factor being treated as within participants. The results of this analysis revealed a main effect for prime such that the similar prime produced more positive attitudes ( $M = 0.08$ ) than the dissimilar prime ( $M = -0.10$ ),  $F(1, 96) = 4.46, p < .04$ . More important, however, this main effect was qualified by an interaction of the prime type with prior



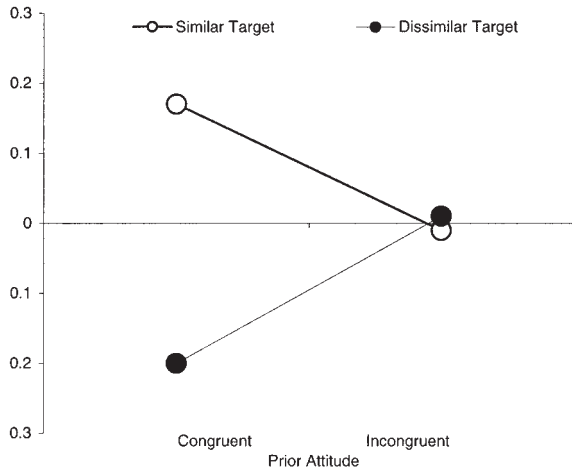


Figure 3. Standardized automatic attitudes as a function of similarity and prior attitude (Study 1).

attitude,  $F(1, 96) = 5.31, p < .03$ . Through examination of the simple main effects, it was evident that participants with congruent prior attitudes showed greater positivity toward the similar prime ( $M = 0.17$ ) than toward the dissimilar prime ( $M = -0.20$ ),  $F(1, 47) = 7.49, p < .01$ . However, participants with incongruent prior attitudes showed no greater positivity toward the similar prime ( $M = -0.01$ ) than toward the dissimilar prime ( $M = 0.01$ ),  $F(1, 49) = 0.03, p > .87$ . Thus, in the congruent (no-change) conditions, the automatic attitude task corresponded with deliberative attitudes toward the target. However, in the incongruent (attitude change) conditions, the automatic task reflected neither solely the initial attitude nor the new attitude. Rather, it reflected a more neutral (ambivalent-like) response that cannot be accounted for by models suggesting that just the old or just the new attitude would be activated.<sup>4</sup>

*Isolating positive and negative reactions.* To unpack these findings and to see whether the neutral responses of incongruent individuals could be clearly associated with ambivalence (and not simply neutrality), we resubmitted the standardized response time data to analysis without reverse scoring negative responses. This time, we conducted a 2 (prior attitude: congruent vs. incongruent)  $\times$  2 (prime: similar vs. dissimilar)  $\times$  2 (target: positive and negative) mixed ANOVA, with the second and third factors being within participants. In essence, this analysis allowed us to determine whether participants in the incongruent condition were relatively fast to respond to both positive and negative stimuli (indicating ambivalence) or relatively slow to respond to both positive and negative stimuli (which could indicate an absence of attitudes or neutrality). There was a significant three-way interaction,  $F(1, 96) = 4.46, p < .04$ , which involved a significant two-way interaction in the congruent condition,  $F(1, 49) = 7.49, p < .01$ , but not in the incongruent condition,  $F(1, 49) = 0.03, p > .87$  (see Figure 4).

As depicted in the top panel of Figure 4, individuals in the congruent conditions were relatively fast to respond to positive stimuli when primed with similar targets and to negative stimuli when primed with dissimilar targets, suggesting strong associations to positivity and negativity in the appropriate conditions. On

the other hand, individuals in the congruent conditions were relatively slow to respond to negative stimuli when primed with similar targets and to positive stimuli when primed with dissimilar targets. This suggests the absence of these links. Of course, these patterns are responsible for the overall automatic positive and negative attitudes obtained in the congruent conditions (see Figure 3).

As depicted in the bottom panel of Figure 4, participants in the incongruent conditions were relatively fast to respond to both positive and negative stimuli regardless of similarity. In fact, participants in the incongruent conditions—which did not differ from each other,  $F(1, 49) = 1.73, p > .19$ —were just as fast as the two fast congruent response groups,  $t(96) = 0.13, p > .90$ , and significantly faster than the two slow congruent response groups,  $t(96) = -2.47, p < .02$ . Overall, then, the data suggested that participants who experienced attitude change responded relatively quickly to both positive and negative stimuli, consistent with the notion of implicit ambivalence. This relatively quick responding to both positive and negative stimuli produced the apparent neutrality evident on the overall automatic attitude (see Figure 3).

### Choice Measure

Following the automatic attitudes task, participants completed the behavioral choice measure described earlier. Because the behavioral measure followed the priming procedure, the measure could serve as evidence that participants' current attitudes had in fact persisted through the duration of the priming task. That is, if congruent and incongruent attitudes differed on the choice task, this could suggest that even though congruent and incongruent attitudes were equivalently strong prior to the priming procedure, the priming procedure itself produced differences in strength between the two kinds of attitudes. If behavioral choices were comparable to the initial deliberative measure, this would suggest that this was not the case. Thus, following the automatic attitude assessment, participants were told that Eddie and Phil were actually lab assistants and that they could choose which of the two they would like to work with in a follow-up interview. Participants were equally and overwhelmingly likely to choose the similar target over the dissimilar target regardless of whether they held a congruent (92% chose similar) or incongruent (96% chose similar) prior attitude (chi-square was nonsignificant). Had participants felt neutral (or explicitly ambivalent) toward the target (as suggested by the automatic measure), one would have expected each target person to be chosen about half the time.

### Discussion

Study 1 provided some initial support for a key prediction from the PAST model. We found that people whose attitudes recently changed appeared to hold explicit attitudes similar to people whose attitudes had not changed (in terms of valence, extremity, and strength), yet they evidenced significant discrepancies in their evaluative responses when assessment occurred with an automatic

<sup>4</sup> We also examined the distribution of responses to see whether the neutral pattern overall was the result of some participants expressing the old attitude whereas others were expressing the new attitude. However, the distribution of responses was normal rather than bimodal.

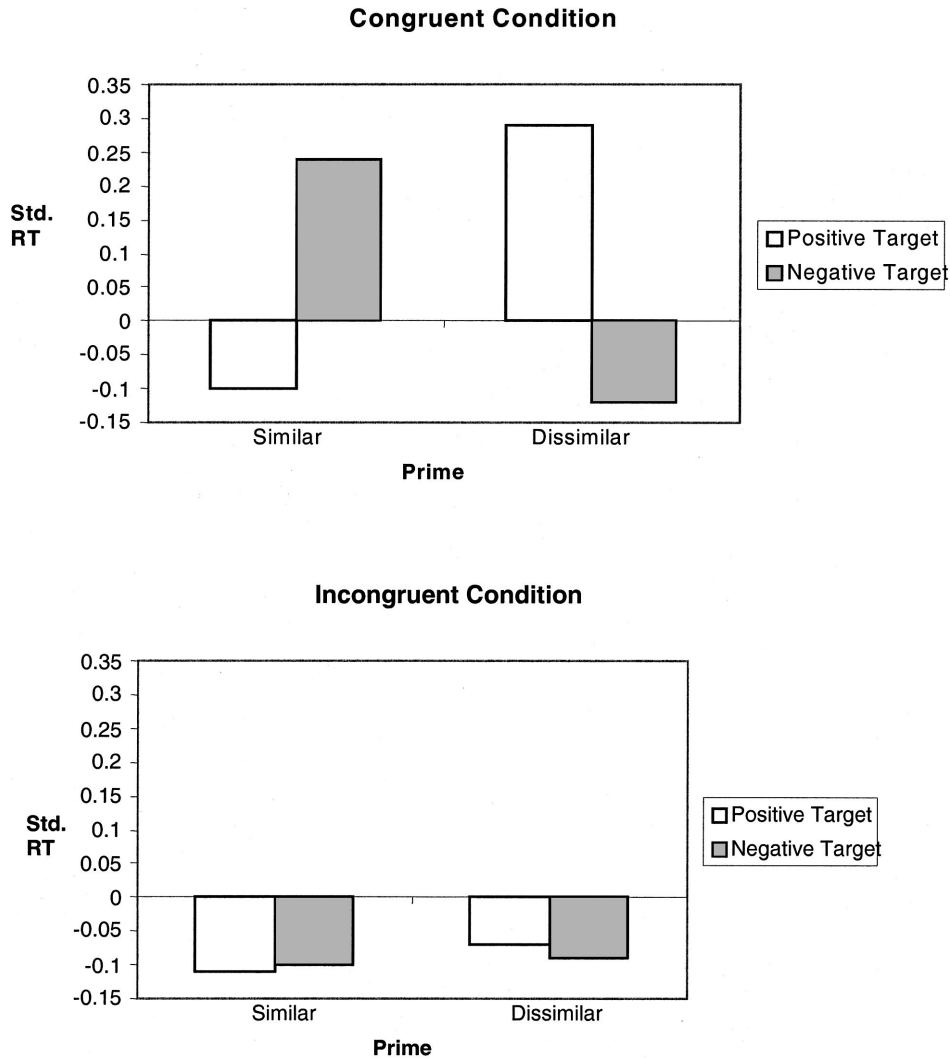


Figure 4. Standardized response times (Std. RT) in Study 1 as a function of congruence, prime, and target. Lower values indicate faster response times, or stronger associations.

attitude measure. Specifically, participants whose attitudes had recently changed appeared to have neutral attitudes according to the evaluative priming procedure. Moreover, this neutrality appeared to reflect ambivalence, as indicated by the simultaneous high accessibility of both positive and negative responses. Because of the methodological structure of our study, we were able to plausibly explain this variance as being a function of people's prior attitudes—a possibility not permitted by traditional attitude change models. Also important, our findings suggest that the activation and impact of old (automatic-implicit) and new (deliberative-explicit) attitudes on judgment are not necessarily an either-or phenomenon, as emphasized in much prior research (see reviews by Dovidio et al., 2001; Greenwald & Banaji, 1995; Wilson et al., 2000). On the contrary, our findings evinced a more dynamic interplay of these attitudes, consistent with the idea that they can jointly influence evaluative responses when time for deliberative responding is not permitted.

It is important to note that the PAST model does not dispute the notion that old and new attitudes can sometimes operate sepa-

rately. When will an ambivalent (neutral) response be observed on an automatic measure when attitudes have changed, and when will the old attitude be more likely to be observed (e.g., see Wilson et al., 2000)? According to the PAST model, it depends on several things (see Figure 1, bottom panel). First, one must consider the strength of the various associative paths (i.e., the initial evaluative link, the new evaluative link, and the false tag). The stronger the initial link is compared with the new link and the weaker the false tag is, the more likely it is that the old attitude will be observed on an automatic measure or in a spontaneous situation. In the current procedure, we wanted to create a strong initial link by using conditioning as well as a strong new link by using a powerful similarity task. Furthermore, the false tag was likely to be weak because participants never had to explicitly reject their old attitudes. Some participants may not even have created a false tag. Thus, the conditions were likely ripe for joint activation of old and new attitudes. Our point is not that joint activation is the only or even the most likely outcome on automatic measures when attitudes have changed. Rather, we suggest that it is one possible and

heretofore unconsidered possibility. More important for our purposes is the idea that people whose attitudes have changed can experience a state of implicit ambivalence. In our remaining studies, we sought to provide further evidence of this notion.

## Study 2

In the preceding study, we assessed the extent of explicit ambivalence by measuring individuals' positive and negative reactions to an attitude object. Since Kaplan's (1972) pioneering research, this has been the standard way to assess what has come to be called objective or operative ambivalence (e.g., Bassili, 1996). In Study 1, people whose attitudes had changed showed no evidence of ambivalence by this criterion. However, evidence for ambivalence was obtained on an implicit measure of attitudes that assessed the degree of association between the focal object and positive and negative words or pictures (Fazio et al., 1995). When attitudes had not changed, the target object was associated more strongly with either positive or negative stimuli. When attitudes had changed, however, the target object was associated equally strongly with both positive and negative stimuli. We argued that this pattern was consistent with a state of implicit ambivalence.

To provide more definitive evidence for our conceptualization, in Study 2 we sought an alternative way to assess implicit ambivalence. Rather than inferring ambivalence from an automatic attitude measure in this study, we developed an implicit measure of confidence or doubt regarding the attitude object. This measure, explained in more detail shortly, was essentially an IAT (Greenwald et al., 1998) in which we assessed whether the attitude object was less associated with confidence when attitudes changed than when they did not.

This implicit measure was compared with an explicit one. Unlike Study 1, in which we used a measure of objective ambivalence, in Study 2 we used a common measure of *subjective* ambivalence. Although Study 1 showed that people were not aware of any conflicting positive and negative reactions to the attitude object about which they experienced change, they might still have been aware of some lingering conflict or doubt about the attitude object even if they did not know its source. Indeed, the two types of measures of ambivalence (objective and subjective) are typically correlated at less than .5 (Priester & Petty, 1996; Thompson et al., 1995), so it is possible that although objective ambivalence (endorsement of positive vs. negative attributes) would not be affected by an attitude change induction, people might still feel more subjectively ambivalent about their changed attitudes. If the ambivalence is implicit, however, we would not expect people to report any greater feeling of doubt on a direct or explicit measure after experiencing attitude change compared with no change.

In addition to using some new measures of ambivalence in Study 2, we also developed a new procedure to create attitude change to establish greater generality for our conceptualization. In particular, the findings of Study 1 may have been due to the fact that classically conditioned attitudes tend to be affective in nature, whereas similarity-based attitudes arguably are more cognitive in nature. Because of this, it is possible that the evidence of implicit ambivalence we have uncovered so far is confined to special cases of affective-cognitive discrepancy (see Cacioppo et al., 1997; Thompson et al., 1995). Although the results would still be interesting and important if they were confined to cases of affective-

cognitive incongruence, they would be more limited in applicability than if they applied more generally to cases of attitude change, as the PAST model implies.

Also germane is the issue of whether the predictions of the PAST model are limited to situations in which old and new attitudes are stored in different memory systems. A number of recent reviews strongly implicate at least two separate memory systems that could be relevant to understanding the key finding from Study 1 (e.g., see McClelland et al., 1995; Sloman, 1996; Smith & DeCoster, 2000; Smolensky, 1988). Across various conceptualizations, one memory system is generally considered to be associative in nature in that it represents the cumulative experience between concepts, and the other is generally considered to be interpretive in nature and able to represent learning from as little as a single instance. With respect to attitudes, the associative memory system may become involved when there are many trials linking an object to an evaluation. Because evaluations based on associative memory are a consequence of many experiences, they are activated automatically but are slow to change (see Smith & DeCoster, 2000). In contrast, interpretive memory can learn rapidly from a single symbolic experience, but it is generally expected to require more effortful and conscious processing to be activated.

This conceptualization has some intriguing implications for the present research. First, if one assumes that prior attitudes will often be based on more trials or experiences than any given new attitude, prior attitudes could plausibly remain in associative memory, whereas relatively new attitudes could be based more on interpretive memory, allowing the two attitudes to exist simultaneously, though in different memory systems. The empirical procedure of Study 1 paralleled this idea. Initial attitudes were based on 40 associative conditioning trials, whereas the relatively new attitudes were based on a single event in which participants learned and interpreted specific information about the target.

Consequently, one might wonder whether the joint operation of two attitudes is dependent on their existence in separate memory systems. In Study 2, we sought to address this issue. To do so, we used an attitude change paradigm in which old and new attitudes were based on the same type of information, thus having the same basis, and were presumably developed and stored within the same memory system. We expected that implicit ambivalence would still result from attitude change even when both old and new attitudes shared a cognitive basis and were presumably stored in the interpretive memory system.

To examine this possibility, we manipulated initial attitudes toward two target individuals by providing participants with information about them. This information was designed to create initially positive attitudes toward one person and negative attitudes toward the other. We then told participants (or not) that the information they had received was accidentally transposed, so they should mentally reverse it. We tested the PAST model prediction that even though participants would presumably feel no explicit ambivalence due to this "mistake," the fact that they had to change attitudes (labeling the old attitude as false) would be sufficient to produce indications of doubt on an implicit measure of ambivalence—the IAT. Unlike Study 1, however, both the original and the changed attitudes in this study were based on the same (cognitive) information and presumably resided in the same (interpretive) memory system.

## Method

### Participants and Procedure

Seventy undergraduates enrolled in introductory psychology courses at the Universidad Autónoma de Madrid participated in exchange for course credit. Sessions were conducted in a single room containing eight partitioned computer stations. Participants were led to believe that they were taking part in a study on person perception and that they would be provided with information about two different people that we would like them to memorize. Participants were then presented with a sheet of paper that provided trait and behavioral information (described below) about the two individuals. One individual was described in very positive terms, and one was described in very negative terms. Then, some individuals received the attitude change induction, and some did not. Next, participants reported their attitudes toward the two individuals followed by measures of deliberative (explicit) and automatic (implicit) subjective ambivalence.

*Initial attitude induction.* There were two target individuals in this study. Both individuals were given a first, middle, and last name: José Luis Martín and Juan Antonio Pérez. To create initial attitudes toward these individuals, we provided participants with trait and behavioral information. For all participants, initial information was positive regarding José Luis Martín and negative toward Juan Antonio Pérez. Specifically, participants received a sheet of paper that described José Luis Martín as sociable, intelligent, hardworking, responsible, caring, empathic, loyal, committed to his friends, and an active volunteer in nonprofit organizations. In contrast, Juan Antonio Pérez was described as intolerant, stubborn, lazy, irresponsible, violent, mean, racist, and unfaithful. Similar procedures have been used to create attitudes toward people in previous research (e.g., Hastie & Park, 1986).

*Attitude change manipulation.* Immediately following the initial information about the target individuals, participants received a piece of paper containing the attitude change manipulation or not, depending on random assignment. Specifically, participants in the change condition were led to believe that the information they just received about José and Juan had been accidentally transposed and needed to be mentally reversed to be accurate. More specifically, participants received a piece of paper containing the following text:

Our apologies . . . We recently discovered that due to errors in copying the information, the files describing Juan Antonio Pérez and José Luis Martín were transposed. That is, the information you just read about these individuals was reversed, so you received incorrect information for each person. As a result, what you just learned about them needs to be reversed. Everything you learned about Juan is, in fact, the information about José. Conversely, everything you learned about José actually corresponds to Juan. Please take a moment to think back to the information to make sure you've now got it right. We are very sorry for this mistake.

The purpose of this information was to create an explicit, cognitive attitude change manipulation that would create new attitudes toward Juan and José that were opposite in valence but based on the same information as attitudes in the control (no-change) condition. Participants in the control condition did not receive any error message.

### Dependent Measures

*Attitudes.* After receiving the trait information and being told of the error (or not), participants completed a measure of attitudes toward José Luis and Juan Antonio. The measure consisted of one item common in impression formation research: To what extent do you like José Luis/Juan Antonio? Participants responded to this item for each target person on a scale ranging from 1 (*not at all*) to 9 (*very much*).

*Deliberative subjective ambivalence.* Immediately following the attitude measure, participants completed a measure of subjective ambivalence. Two items (adapted from Priester & Petty, 1996) constituted this measure: To what extent do you feel any confusion when you think about José Luis and Juan Antonio? To what extent do you feel secure in your opinions of José Luis and Juan Antonio? Participants responded to these items on scales ranging from 1 (*not at all*) to 9 (*very much*).

*Automatic subjective ambivalence.* Following the measure of deliberative subjective ambivalence, we assessed automatic subjective ambivalence by using an adaptation of the IAT (Greenwald et al., 1998). In this IAT, participants classified target concepts (represented by *José/Juan* or *other*) and attributes (represented by *confidence* or *doubt* categories of words) by using two designated keys. The *José/Juan* category was represented by the words *José, Luis, Martín, Juan, Antonio,* and *Perez*. The *other* category was represented by the words *Pedro, Pablo, García, Enrique, Francisco,* and *Rodríguez*. Attributes related to confidence included the words *confident, certain, secure, sure,* and *firm*. In contrast, attributes related to doubt included the words *doubtful, hesitant, skeptical, conflicted,* and *ambiguous*.

The test was very similar to the original IAT used by Greenwald et al. (1998). There were seven blocks of trials in total. Blocks 1, 2, and 5 were practice blocks in which participants made single categorizations (*José/Juan* vs. *other*, or *confidence* vs. *doubt*). In the remaining blocks, participants discriminated between *confidence* versus *doubt* words and *José/Juan* versus *other* words on separate trials within the same block. In Block 4, participants used one response key to indicate if a word belonged to the *doubt* or *other* categories and the other key to indicate if the word belonged to the *confidence* or *José/Juan* categories. In Block 7, participants used one response key to indicate if a word belonged to the *doubt* or *José/Juan* categories and the other key to indicate if the word belonged to the *confidence* or *other* categories. Blocks 3 and 6 were combined blocks that served as practice for Blocks 4 and 7. Only data from Blocks 4 and 7 were used to compute IAT scores. The main dependent variable (IAT score) was computed by subtracting participants' average response latencies during Block 4 from their average response latencies during Block 7. Positive differences on this index indicated faster automatic associations between *José/Juan* and *confidence* than between *other* and *confidence*.

Following typical procedures, stimuli in the IAT task appeared within a centered white window. Reminder labels were positioned above the stimuli on the left and right side of the screen. These reminders read "*José/Juan*" and "*Other*" for single target-classification blocks and "*confidence*" and "*doubt*" for single attribute-classification blocks. Mixed target + attribute blocks were also accompanied by appropriate labels (e.g., "*confidence* or *other*" and "*doubt* or *José/Juan*"). Incorrect classifications were followed by error feedback (i.e., the word *ERROR* flashed on the screen). At the end of each practice block, summary feedback was provided that informed participants of their average response latency and percentage of errors for that block. Within each block, stimuli were randomly selected without replacement, and no more than two consecutively presented stimuli belonged to the same category. To correct for anticipatory responses and momentary inattention, latencies shorter than 300 ms and longer than 3,000 ms were recorded as 300 and 3,000 ms, respectively. Response latencies were then log transformed to normalize the distribution prior to computing the measure of implicit ambivalence. Further details about the IAT procedure can be found in Greenwald et al. (1998).

## Results and Discussion

### Attitudes

We began by creating an attitudinal variable for the target who was associated with positive traits (*José Luis* for the control group and *Juan Antonio* for the attitude change group), and also for the target who was associated with negative traits (*Juan Antonio* for

the control group and José Luis for the attitude change group), and then submitted these data to a 2 (target: positive vs. negative)  $\times$  2 (prior attitude: congruent vs. incongruent) mixed ANOVA. In this analysis, target was a within-participants factor and congruency was a between-participants factor. We expected the positive target to be liked more than the negative target, but we did not expect participants who had changed their attitudes (i.e., received the error manipulation) to evaluate the positive and negative targets any differently than participants who had not changed their attitudes.

Consistent with expectations, only a significant main effect for the target person emerged such that participants reported more positive evaluations for the positive ( $M = 7.95$ ,  $SD = 0.18$ ) than for the negative ( $M = 1.46$ ,  $SD = 0.13$ ) target,  $F(1, 67) = 542.13$ ,  $p < .0001$ . It is important to note that neither the main effect for congruency,  $F(1, 67) = 0.003$ ,  $p = .96$ , nor the interaction was significant,  $F(1, 67) = 0.29$ ,  $p = .58$ . Thus, participants liked the individual associated with the positive traits more than the individual associated with negative traits to the same extent regardless of whether the associations were the initial ones or had to be reversed from the initial ones.

### *Deliberative Subjective Ambivalence*

After reverse scoring one item, responses to the two ambivalence questions were significantly positively correlated ( $r = .65$ ,  $p < .001$ ), so we averaged them to form a single composite index of deliberative subjective ambivalence. Higher scores indicated less ambivalence. Again, we did not expect participants who had earlier changed their attitudes to report any greater doubt or ambivalence compared with participants who had not changed. Consistent with this view, there was no difference in subjective ambivalence across the incongruent ( $M = 3.31$ ,  $SD = 1.46$ ) and congruent ( $M = 3.67$ ,  $SD = 1.83$ ) conditions,  $F(1, 69) = 0.82$ ,  $p = .37$ .

### *Automatic Subjective Ambivalence*

As expected, the automatic ambivalence data revealed a different pattern. As predicted by the PAST model, participants showed significantly less certainty with respect to the target names in the incongruent condition ( $M = 248.75$ ,  $SD = 243.53$ ) than in the congruent condition ( $M = 378.64$ ,  $SD = 290.56$ ),  $F(1, 69) = 4.65$ ,  $p < .04$ . Again, lower values on the IAT index reflected weaker associations between the target individuals and confidence words. Thus, although they reported the same attitudes and the same degree of confidence/doubt in their opinions at the explicit level, participants showed greater evidence of ambivalence on the implicit measure after changing rather than not changing their attitudes.

## Study 3

We had now developed two different paradigms for creating changed and unchanged attitudes that are similar in valence, extremity, and explicit ambivalence but that appear to differ in implicit ambivalence. In the first paradigm, attitudes were classically conditioned to be positive or negative and then changed (or not) with a similarity procedure. In the second paradigm, attitudes

were initially created with some (cognitive) information about the target and then changed (or not) by the experimenter's error. Although the paradigms differ in many ways, the findings were conceptually similar. In both studies, individuals whose attitudes had changed reported the same final attitudes toward the target and the same degree of explicit objective (Study 1) or explicit subjective (Study 2) ambivalence on traditional deliberative measures. However, in both studies, there was evidence of ambivalence on measures of automatic association.

In Studies 3 and 4, we sought to examine another potential consequence of implicit ambivalence from attitude change, and we aimed to demonstrate this consequence within each of the general paradigms used thus far. Prior research has revealed that when people are ambivalent at an explicit level (i.e., they are ambivalent and know it), they tend to engage in increased information processing, presumably to reduce their feelings of ambivalence (e.g., Bell & Esses, 2002; Jonas, Diehl, & Bromer, 1997; Maio et al., 1996). In one study, for instance, Maio et al. (1996) presented individuals who were or were not ambivalent in their attitudes toward Asians with a message containing strong or weak arguments in favor of immigration from Hong Kong. The primary result was that the quality of the arguments in the message had a larger effect on immigration attitudes of individuals who were high rather than low in ambivalence. The enhanced argument quality effect indicated greater message processing (Petty & Cacioppo, 1986).

In Study 3, we explored the possibility that discrepancies between old and new attitudes could produce a similar outcome. As in Study 1, we produced initial attitudes toward a target person (Eddie) by using a classical conditioning procedure and then changed attitudes at the explicit level by using a similarity manipulation. Following this manipulation and corresponding attitude change, we presented participants with persuasive information about Eddie as a potential job candidate that was either strong or weak. We predicted that when participants' prior (conditioned) and current (similarity-based) attitudes were incongruent, they would show a greater argument quality effect than when these attitudes were congruent. In other words, when attitudes toward Eddie had changed, participants would act as if they were ambivalent about Eddie by processing information about him to a greater extent. If they did so, then their evaluations of Eddie would be more influenced by the quality of the information presented about him. When attitudes had not changed, evaluations of Eddie were not expected to be influenced as much by argument quality.

## *Method*

### *Participants and Procedure*

Participants were 77 undergraduates enrolled in introductory psychology courses at Ohio State University. As in Study 1, all sessions were conducted in a room with four partitioned work areas, and up to four people participated during any given session. Overall, the procedure was very similar to that used in Study 1. When participants arrived, they reported their attitudes toward the same issues as before (e.g., smoking in public, lesbians and gay men in the military) on 7-point scales ranging from 1 (*very much against*) to 7 (*very much in favor*). The attitudes questionnaire was administered via computer by using MediaLab (Jarvis, 2000). Responses were written to a data file that was accessible for the explicit attitude change (i.e., similarity) manipulation. Participants next completed

the conditioning phase of the study followed by the similarity phase. After attitudes had been formed and changed (or not) with respect to the target person, participants were given additional information about the target person, and their extent of information processing was assessed.

*Initial attitude induction.* After the attitude questionnaire (used for the similarity manipulation), we instructed participants that the primary purpose of the session was to examine issues having to do with visual memory. To create the initial attitudes, we used the same conditioning procedure as in the first experiment, but in this study, we focused on just one target face (Eddie from Study 1) because both faces showed the same results in Study 1. On the basis of random assignment, the face of Eddie was paired 40 times with negative images or 40 times with positive images (as was the case in Study 1).

*Similarity (attitude change) manipulation.* Participants were instructed at the end of the conditioning trials that they needed to give the information a chance to settle in before the visual memory test. Participants were then told that we were interested in the impressions they would form of these men if they had some personal information about them (i.e., their responses to the attitude questionnaire). Participants were told that they would be receiving information about just one of the men to conserve time in the experimental session. Participants then viewed a picture of Eddie, which was presented with his ostensible responses to the attitude survey. The similarity of these responses to participants' responses was manipulated as in the first experiment.

Following Study 1, half of the participants were assigned to the congruent (i.e., no attitude change) condition in which Eddie was made to appear dissimilar when he had previously been associated with negative images and similar when he had been associated with positive images. The other half of the participants were assigned to the incongruent (i.e., attitude change) condition in which Eddie was made to appear dissimilar when he had previously been associated with positive images and similar when he had been associated with negative images.

*Argument quality manipulation.* After the attitude change manipulation, participants were led to believe that the Ohio State University Department of Psychology was searching for a new faculty member. They were told that some of the men they had seen in this experiment were actually applicants for the position and that the search committee was seeking feedback from both undergraduate and graduate students. To obtain feedback on one of the applicants, Eddie, we told participants we would provide them with some information pertaining to his qualifications. Participants were told that Ohio State has an excellent psychology department and that it was important that the new hire be able to maintain this high standard in both teaching and research. Following this introduction, participants were presented with a summary of the ostensible résumé (i.e., curriculum vitae [CV]) of Eddie. The core information it contained was randomly assigned to be strong or weak. The CV information is presented in Table 1.

### Attitude Measure

After reading the synopsis of Eddie's CV, participants were told that it was important to record their evaluations of Eddie as a potential job candidate for Ohio State University. Participants were told that there were no right or wrong answers to this question and that their responses would remain totally anonymous. Two items were used to measure attitudes toward Eddie as a job candidate. Participants were asked how good they thought that Eddie might be for the job and to what extent they would be willing to hire him. Both items were responded to on scales ranging from 1 (*not at all*) to 7 (*very much*). These items were highly correlated ( $r = .88$ ,  $p < .001$ ) and thus were averaged to form a single attitude score.

### Results and Discussion

As in Study 1, we created a congruency factor based on whether the initial conditioning and subsequent similarity manipulation

Table 1  
*Core Information From the Strong and Weak Argument (Vita) Conditions*

| Condition | Argument  |
|-----------|---|
| Strong    | M.A., Psychology, Florida State University (1996)<br>Ph.D., Psychology, Stanford University (2001)<br>10 professional affiliations to international associations and societies of psychology<br>7 national honors and awards in psychology<br>2 edited books on psychology topics<br>6 book chapters in psychology<br>3 articles in the top journal in the field<br>Scored 9.8 points out of 10 on the Quality Research Score Index<br>Teaching in psychology is exceptional in quality: 4.75 points (scale from 1 to 5 points) |
| Weak      | M.A., Education, Australian College of Education, Sydney (2000)<br>B.A., Education, Australian College of Education, Sydney (1996)<br>Has not defended his Ph.D. dissertation yet<br>2 book chapters in psychology<br>2 articles in a medium-quality journal in the field of school testing<br>Scored 5.1 points out of 10 on the Quality Research Score Index<br>Teaching in psychology is average in quality: 3.2 points (scale from 1 to 5 points)   |

*Note.* M.A. = master of arts; Ph.D. = doctor of philosophy; B.A. = bachelor of arts.

produced conditions of attitude change or not. We then submitted the attitude data to a 2 (congruency: congruent vs. incongruent)  $\times$  2 (argument quality: strong vs. weak)  $\times$  2 (initial conditioning: positive vs. negative) ANOVA. A main effect for argument quality,  $F(1, 69) = 14.43$ ,  $p < .001$ , showed that participants who received the strong CV held more favorable attitudes toward Eddie as a job candidate ( $M = 4.99$ ,  $SD = 1.33$ ) than did participants who received the weak CV ( $M = 3.98$ ,  $SD = 1.16$ ). As predicted, though, this main effect was qualified by an interaction between attitude congruence and argument quality,  $F(1, 69) = 14.47$ ,  $p < .001$ . Consistent with the PAST model predictions, argument quality made a larger difference when participants' old and new attitudes toward Eddie were incongruent,  $F(1, 69) = 30.20$ ,  $p < .001$ , than when they were congruent ( $F < 1$ ; see Figure 5, top panel).

Unexpectedly, there was also a two-way interaction between the direction of the conditioning and argument quality,  $F(1, 69) = 9.70$ ,  $p < .01$ . This interaction indicated that there was a stronger argument quality effect when Eddie had initially been conditioned to be negative,  $F(1, 69) = 24.30$ ,  $p < .001$ , than when he had been conditioned to be positive ( $F < 1$ ) and may reflect individuals' general propensity to process negative or untrustworthy individuals more than trustworthy ones (e.g., Priester & Petty, 1995). Notably, the initial conditioning did not qualify the two-way interaction between congruency and argument quality, and there were no other significant effects in this analysis.

In Study 3, we obtained the predicted pattern of results according to the PAST model. In short, participants' evaluative responses to a job candidate were more affected by the strength of his

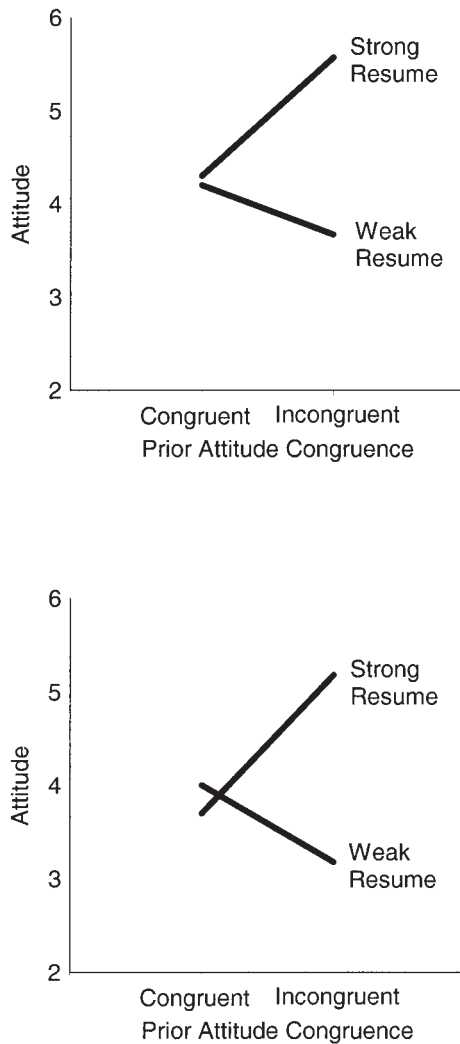


Figure 5. Top panel: Interaction on attitudes between congruency and argument quality (Study 3). Bottom panel: Interaction on attitudes between congruency and argument quality (Study 4).

credentials (i.e., argument quality) when their old and new attitudes were incongruent rather than congruent. This pattern suggests that people engaged in greater information processing when they had recently undergone attitude change than when they had not. Because the design of the prior attitude and attitude change manipulations closely followed the design of Study 1, we can reasonably infer that participants' attitudes, prior to the argument quality manipulation, were no different in valence, extremity, accessibility, or explicit objective ambivalence. Instead, the findings are most compatible with the notion that participants whose old and new attitudes were incongruent were ambivalent at an implicit level, and they thought more deeply about Eddie's credentials, presumably to resolve the implicit ambivalence. When old and new attitudes were congruent, there appeared to be little motivation to process the information, as suggested by the absence of an argument quality effect under these conditions. Instead, participants may have given only a cursory look at the CV and

provided a moderate evaluation, perhaps relying on a numerosity heuristic (i.e., counting the available evidence rather than evaluating its merits; see Petty & Cacioppo, 1984). It is not uncommon for argument quality not to matter when people are devoting little cognitive effort to the evaluation (Eagly & Chaiken, 1993; Petty & Cacioppo, 1986).

#### Study 4

To this point, the results are consistent with the idea that attitude change can produce a state of implicit ambivalence, resulting in apparent neutrality on an automatic measure (Study 1), a reduction in implicit confidence about the attitude object (Study 2), and increases in information processing (Study 3). The information-processing effect was demonstrated in Study 3 by using a paradigm in which initially conditioned attitudes were changed with more cognitive information. Although Study 2 suggested that this paradigm was not required for obtaining evidence of implicit ambivalence, in our final study we wanted to demonstrate the generality of the behavioral consequences of implicit ambivalence by showing that the effect would generalize to the paradigm introduced in Study 2. In this paradigm, attitudes initially created with cognitive information were changed because of an experimental error.

#### Method

##### Participants and Procedure

Participants were 45 undergraduates enrolled in introductory psychology courses at Ohio State University. Sessions were conducted in a single room containing eight partitioned computer stations, and up to eight people participated in any given session. When participants arrived, they were welcomed by an experimenter and asked to complete a questionnaire assessing their attitudes toward a variety of issues. The attitude survey was identical to the surveys administered at the outset of Studies 1 and 3 and served as the basis for a similarity induction that was used to create participants' initial attitudes. After initial attitudes were created with the similarity induction, as in Study 2, participants were informed (or not) about the experimenter's error in reversing the information about the target person. Following a measure of explicit subjective ambivalence, participants were given additional information about the target person, and their extent of information processing was assessed.

*Initial attitude induction.* As an introduction to the experiment, participants were led to believe that the Ohio State University Department of Psychology was searching for a new faculty member. They were told that the department was seeking feedback from both undergraduate and graduate students about two particular candidates under consideration. Furthermore, participants were instructed that they would be provided with both personal and academic information about the candidates to help them make their decisions. Following this introduction, participants were exposed to ostensible personal information about two individuals, Eddie and Phil. Specifically, participants received the attitude survey responses of Eddie and Phil, presented in the same manner as in Studies 1 and 3 (the order of presentation was random for each participant). The computer randomly selected either Eddie or Phil to have attitudes that were similar to the participant. The other person's attitudes were dissimilar. Thus, if Eddie was similar, then Phil was dissimilar, and vice versa.

*Attitude change manipulation.* Immediately following the initial survey (similarity) information pertaining to Eddie and Phil, participants were exposed to the attitude change manipulation (or not, according to random assignment). This manipulation was essentially the same as that used in

Study 2. In brief, some participants were led to believe that the survey information they just received about the job candidates, Eddie and Phil, had been accidentally transposed and needed to be mentally reversed to be accurate. Thus, they were told that everything they had learned about Eddie actually applied to Phil and vice versa. Participants in the control, or no-change, condition did not receive this error information. As in Study 2, the purpose of this procedure was to create an explicit, cognitive attitude change manipulation that would create new attitudes toward Eddie and Phil that were opposite in valence but based on the same information as attitudes in the control (no attitude change) condition.

*Subjective ambivalence measure.* After receiving the survey responses of Eddie and Phil and being told to reverse the information (or not), participants completed a measure of subjective ambivalence. Specifically, participants were asked, "To what extent do you feel conflict when you think about Eddie (Phil)?" Participants reported their subjective conflict toward Eddie and Phil on separate scales, ranging from 1 (*not at all*) to 9 (*very much*). This measure was adapted from Priester and Petty (1996).<sup>5</sup>

*Argument quality manipulation.* Once participants completed the subjective ambivalence measure, they were informed of Ohio State University's high standard of excellence for hiring, as in Study 3. Then we instructed participants that they would now be exposed to additional information about the two candidates. In actuality, to simplify the experimental design, participants received information only about Eddie in this phase of the experiment. As in Study 3, participants were presented with a summary of the job candidate's ostensible CV. The information in the CV was randomly assigned to be either strong or weak—that is, it was designed to make Eddie seem highly qualified or not so qualified, respectively. This manipulation was identical to that used in Study 3 and is presented in Table 1.

*Attitudes toward Eddie as a job candidate.* After viewing Eddie's CV, participants were told that it was important to assess their evaluation of him as a potential job candidate. Participants were told that there were no right or wrong answers and that their responses would remain totally anonymous. Attitudes toward Eddie as a job candidate were then recorded by using a single item. Specifically, participants were asked, "To what extent do you think Eddie would be good for the Psychology Department position?" Responses were given on a scale ranging from 1 (*not at all good*) to 7 (*very good*).

## Results and Discussion

### Deliberative Subjective Ambivalence

Although Study 2 suggested that our reversal manipulation did not produce any subjective ambivalence on an explicit measure, we examined this possibility once again, this time with a measure focusing on explicit conflict with respect to the attitude object, Eddie. As in the prior studies, we created a congruency factor based on whether the initial attitude was modified with the reversal manipulation (incongruent) or not (congruent). A 2 (congruency: congruent vs. incongruent)  $\times$  2 (argument quality: strong vs. weak) between-participants ANOVA revealed that there were no main effects for either argument quality ( $F < 1$ ) or congruency condition,  $F(1, 41) = 1.85, p > .18$ , indicating that felt conflict was essentially the same across the incongruent condition ( $M = 5.91, SD = 1.60$ ) and the congruent condition ( $M = 5.17, SD = 1.99$ ). Also important, there was no interaction between congruency and argument quality,  $F(1, 41) = 2.35, p > .13$ .

### Attitudes

Attitudes, reported following examination of Eddie's CV, were also submitted to a 2 (congruency)  $\times$  2 (argument quality)

ANOVA. As in Study 3, we expected an interaction between the congruency and argument quality factors. More specifically, our hypothesis was that there would be a significantly larger argument quality effect for participants who had earlier changed their attitudes (i.e., received the reversal manipulation) than for participants whose attitudes had not changed.

Our findings were compatible with these predictions. To begin with, there was a marginally significant main effect for argument quality,  $F(1, 41) = 3.68, p < .06$ , indicating that attitudes toward Eddie as a job candidate were more favorable in the strong vita condition ( $M = 4.45, SD = 1.77$ ) than in the weak vita condition ( $M = 3.61, SD = 1.41$ ). Of greatest importance, however, was that this marginal main effect was qualified by a significant two-way interaction with congruency,  $F(1, 41) = 6.36, p < .02$  (see bottom panel of Figure 5). Argument quality had a larger impact on attitudes toward Eddie as a job candidate in the incongruent (attitude change) condition,  $F(1, 41) = 9.65, p < .01$ , than in the congruent (no-change) condition ( $F < 1$ ).

### Summary

The results of Study 4 provided an important extension of the findings from the first three experiments. In this study, we conceptually replicated the finding from Study 3, in which apparent attitude change led to an increase in information-processing activity, consistent with the notion that participants experienced some implicit ambivalence. We also replicated the finding from Study 2 by showing that following attitude change (vs. no change), there were no differences in subjective ambivalence. As in Study 2, although participants did not indicate that they felt conflicted, they acted as if they experienced some conflict with respect to the attitude object (i.e., they engaged in greater scrutiny of attitude-relevant information). Study 4 also extended Study 3 by demonstrating that the behavioral (i.e., information-processing) consequences of attitude change are confined neither to special cases of affective-cognitive ambivalence nor to special cases in which old and new attitudes are stored in separate memory systems. On the contrary, the PAST model appears to have a wider range of potential application to attitude change scenarios more generally.

<sup>5</sup> In Study 4 (unlike Study 2), we did not assess attitudes immediately after the initial attitude induction or after the subsequent attitude change manipulation. Participants were led to believe at the beginning of the session that we were interested in their feedback about a job candidate based on both personal and academic information. Although it would have been useful to demonstrate that prior to the argument quality manipulation, attitudes were equal in valence and extremity in the change and no change conditions, we wanted to create conditions in this experiment that would facilitate participants' attempt to synthesize the personal and academic information into a single global attitude. We felt that if we measured attitudes after the personal information and participants already knew academic information was on the way, we might lead participants to explicitly store two separate impressions, which would not provide a strong test of the PAST model prediction that both evaluations continue to impact evaluative responding. In any case, the data from Study 2 show that participants are capable of reversing the information and reporting the same explicit attitudes as participants who do not have to reverse the information.



## General Discussion

Across four studies, we demonstrated that people who used to hold one attitude but then shifted to another attitude looked similar on traditional explicit measures of attitudes and attitude strength to individuals who held the same attitude all along. This is not surprising, as decades of research on persuasion have shown that one can successfully modify people's evaluations of various objects and issues. Furthermore, changed attitudes did not appear to be held with any greater doubt or ambivalence than unchanged attitudes when we used traditional explicit measures of both objective (Study 1) and subjective (Studies 2 and 4) ambivalence.

Why did people not recognize any ambivalence when their attitudes changed in our studies? As noted earlier, we do suspect that people will sometimes feel ambivalent when their attitudes change. For example, people can move from an unambivalent pro-chocolate cake attitude, in which only the pros are recognized, to an ambivalent anti-chocolate cake attitude, in which some cons are simply added to and outweigh the pros. In the current research, however, we attempted to establish situations in which explicit ambivalence would be unlikely. In the conditioning studies (1 and 3), participants started out with a vague (but real) positive or negative feeling toward the target individuals but then learned information about the targets' attitudes on highly important topics (e.g., belief in God). It seems perfectly sensible for participants to put aside their initial vague feelings and focus on the highly diagnostic information about the targets' attitudes. In Studies 2 and 4, the attitude change occurred simply as a result of an experimental error. If the traits associated with the target person simply were wrong, there is no logical reason to feel any explicit conflict about changing to the correct view. In general, if people have truly changed their attitudes from one to another, there may often be no reason to feel ambivalent. Indeed, since the inception of scientific attitude change research, we could not find any suggestion that changing one's attitudes is necessarily associated with ambivalence.<sup>6</sup>

Yet, although no ambivalence was observed on direct assessments of this construct, the picture was dramatically different on more implicit and indirect assessments. In all four experiments, participants showed signs of ambivalence on these measures when their attitudes were changed. In Study 1, individuals whose attitudes had changed appeared more neutral on an automatic attitude measure, reflecting equivalently fast responding to positive and negative stimuli. In Study 2, individuals whose attitudes had changed showed evidence of reduced confidence about the attitude object on an IAT compared with people whose attitudes had not changed. In Studies 3 and 4, individuals whose attitudes had changed engaged in greater processing of information relevant to the attitude object than did people whose attitudes had not changed. All of these responses are ambivalent-like reactions that were not expected from prior conceptualizations of attitude change. Next, we turn to some limitations and implications of our findings.

### *Methodological Considerations*

#### *Were the Changes at Time 2 Genuine?*

Our analysis rests on the assumption that in the incongruent conditions of our studies, we produced attitudes that were genu-

inely changed at Time 2 by our manipulations. However, one might argue that the similarity and reversal manipulations did not create a genuine attitude but rather a constructed and temporary evaluative response that was tapped by the explicit measure. If this were indeed the case, it could be argued that attitude change did not actually occur. If the new response was a temporary construction and not a new attitude per se, it could simply have decayed before the response time facilitation task (Study 1) or the evaluation of Eddie as a job candidate (Studies 3 and 4). We offer several responses to this suggestion. First, in Study 1, the attitudes induced through the similarity manipulation affected not only the attitudes reported right after the similarity manipulation but also the behavioral choice measure, which followed the priming task. It seems unlikely that the initial response was a temporary construction that decayed before the response time task and then was reconstructed for the choice task. Second, even if this were assumed, one would still have to account for the fact that the attitudes induced through the similarity paradigm were evident in the response time task. Because the priming task is presumed to prevent the influence of controlled or constructed evaluative thought, the possibility of the similarity-based attitudes being temporary constructions and still affecting these responses seems limited.

In Study 2, we found that people whose attitudes had changed because of an experimenter's error showed a reduction in confidence on an implicit measure compared with people who had not changed. If the former individuals had not really changed their attitudes, they should have responded to the implicit confidence measure in a manner similar to those who did not receive the reversal manipulation. Yet their responses were quite different. Finally, with respect to Studies 3 and 4, it is not clear why the second attitude would interact with the initial attitude to determine the extent of information processing about the attitude object if the second attitude were simply a temporary construction.

#### *Were the Attitudes in Congruent and Incongruent Groups Really the Same at Time 2?*

Our analysis also rests on the assumption that participants held equivalent deliberative (explicit) attitudes following the attitude change procedure regardless of whether their prior attitudes were congruent or incongruent with their current evaluations. In Study 1, we demonstrated this by showing that the attitudes of individuals in the change group did not differ from the no-change participants' attitudes in terms of valence, extremity, accessibility, objective ambivalence, or choice behavior. In Study 2, we showed that changed attitudes did not differ from unchanged attitudes in

<sup>6</sup> One reviewer of this article suggested that participants must be lying on the subjective ambivalence measures. However, it is not clear why they would lie on this measure but not on the attitude measures. Furthermore, prior research has demonstrated that people are very willing to report possessing both positive and negative reactions (objective ambivalence) and feeling ambivalent (subjective ambivalence) on a host of diverse attitude objects, such as their parents, their country, and their school (e.g., Bargh et al., 1992; Larsen, McGraw, & Cacioppo, 2001; Priester & Petty, 1996, 2001). It is not clear why participants would be reticent to express ambivalence regarding the target individuals here. Finally, the fact that we stressed the complete anonymity of responses in all studies also reduced any motivation to lie.

terms of valence, extremity, or subjective confidence. In Study 4, we provided further evidence for the equivalence of these attitudes by showing that they did not differ in self-reported feelings of conflict. These were critical demonstrations to rule out the possibility that the strength of participants' current attitudes (as assessed with traditional explicit measures) caused the ambivalent-like responding. Although it actually would be quite an interesting and novel finding if people were aware of feeling less certain of changed attitudes compared with unchanged attitudes of equivalent valence, this did not seem to be the case.

It is important to note that although we showed that our groups did not differ on any of the traditional attitude strength measures we assessed, it is useful to emphasize that our claim of equivalence is based on null findings. Consequently, there are a number of factors that could have influenced this conclusion. For example, it could be that we had insufficient power to detect effects on these measures, though our manipulations did produce statistically significant effects on other measures (e.g., deliberative and automatic attitude assessments). Also, it is certainly possible that attitudes could still have differed in one of many unmeasured ways as there a large number of indicators of attitude strength (see Petty & Krosnick, 1995). However, we demonstrated equivalence by several well-established criteria that seemed most directly relevant to our framework and most likely to have been affected by our manipulations of prior and current attitudes. Nevertheless, exploration of the conditions under which people are or are not aware of feeling conflicted about changing their attitudes would be fruitful for further inquiry.

### *Directions for Future Research*

#### *Further Consideration of Dual Memory Systems*

In the present research, we attempted to examine whether the PAST model would still hold when old and new attitudes existed within the same memory system. In Studies 2 and 4, we found that even when people's old and new attitudes were presumably both stored in interpretive memory, evidence for implicit ambivalence was obtained. In future work, it would be interesting to examine whether the same effects obtain when old and new attitudes are both stored in associative memory. For example, one could use a counterconditioning paradigm in which initial attitudes would be formed through a process of conditioning over many trials (as in the current research), but then these attitudes would be reversed by conditioning over many new trials. This would parallel the concept of a new attitude becoming highly practiced over time and consequently also represented in associative rather than interpretive memory. The question would then be whether a prior attitude could still impact evaluative responding when the new attitude reaches some criterion level of association. Research on classical conditioning suggests that this is indeed possible (e.g., Baeyens, Eelen, & Crombez, 1995), but new research examining the implications of this for implicit ambivalence would be quite useful.

#### *Conditions Facilitating the Impact of Prior Attitudes*

According to the PAST model (see Figure 1, bottom), the influence of prior attitudes would be facilitated by the following: (a) a failure to explicitly reject an existing attitude when a new

opposing one is formed (i.e., failing to tag it false), (b) a failure to retrieve the false tag if and when the prior attitude subsequently comes to mind, or (c) an inability to inhibit the application of the prior attitude even if it is recognized as false. In future research, it would be useful to tease these possibilities apart and to determine the extent to which each of them affects the impact of prior attitudes. It could be particularly useful to examine the role of explicit rejection of the prior attitude at the time of attitude change. For example, prior to our response time task in Study 1, imagine that we had reminded participants of their prior attitudes and asked them whether they still felt that way. By doing this, participants would have been given the opportunity to explicitly reject their prior attitudes—in essence, tagging them as false. It is interesting to note that the accessibility of the prior attitude would also be increased (see Maio & Olson, 1995) along with the false association tagged to that attitude. Although we know that accessibility facilitates the potential impact of an attitude, the rehearsal of its rejection should have an opposing effect by facilitating the automatic recognition of it as false when it does come to mind (Kawakami et al., 2000). The accessibility of the tags marking prior attitudes as false might be as important for determining some behavior as the accessibility of the new attitude—especially when the prior attitude remains more accessible than the current attitude. It would also be informative to know when people spontaneously apply false tags. One possibility is that false tags are more likely to be applied when an attitude changes its valence (as in the current research) than when attitudes stay positive or negative and just change their degree. In any case, future exploration of false tags (and the possibility of true tags) applied to one's evaluations is an avenue of research suggested by the PAST model.

#### *Attitude–Behavior Consistency*

In the past three decades, a number of models have been developed to account for the frequently observed discrepancy between attitudes and behavior. As a result, when attitudes do not guide behavior, investigators tend to cite research that describes how an attitude can sometimes become impotent (e.g., Davidson & Jaccard, 1979; Fazio, 1995; Fishbein & Ajzen, 1975). The PAST model provides a complementary account of attitude–behavior discrepancy that does not diminish the power of the attitude construct. Under conditions where prior attitudes are influential but only current attitudes have been measured, it seems unwise to attribute a low attitude–behavior correlation to an inability of attitudes to predict behavior. The problem may be that an influential prior attitude was simply not assessed.

Although it is well documented that past behavior, particularly in terms of habit (e.g., Triandis, 1971), can be an important determinant of current behavior (for a review, see Ouellette & Wood, 1998), there is no research looking at the power of past attitudes to guide behavior. For example, Sutton and Hallet (1989) found that following attitude change toward seat belt use, participants' old seat belt wearing habits remained an important influence above and beyond participants' new attitudes. If the PAST model is correct, it is possible that participants' prior attitudes, rather than prior behavior per se, could be responsible for this effect as prior behavior would likely have been correlated with prior attitudes. Indeed, the idea that prior attitudes can operate like evaluative habits (e.g., Petty et al., 1993; see also Payne, Jacoby,

& Lambert, 2005) is very consistent with the PAST model. However, because prior attitudes were not assessed in past studies on behavioral habits, we cannot tell whether they were an influential factor. In future research, it would be useful to examine whether past behavior can still predict current behavior after controlling for prior attitudes.

### Conclusion

In the present research, our intention was not to suggest that old attitudes never disappear or that prior attitudes will forever remain impactful. Rather, our intention was to demonstrate the danger of presuming that apparent attitude change as assessed by typical procedures necessarily reflects a literal changing, updating, or destruction of the prior attitude. In some cases, current evaluative responding can be explained with the knowledge of how people felt in the past. In this sense, our model is compatible with the dual attitudes model approach to attitude change proposed by Wilson et al. (2000; see also Dovidio et al., 2001; Greenwald & Banaji, 1995). We departed, however, in suggesting that when prior and current attitudes conflict, they can create a state of implicit ambivalence and sometimes have a joint influence on current responding. For example, they can simultaneously affect automatic evaluative responses as operationalized by an evaluative priming task and implicit confidence as assessed with an IAT. Moreover, they can interact to determine the amount of information processing in which an individual engages. Given these consequences, we encourage future researchers interested in persuasion or attitude-behavior consistency to consider the role of prior attitudes whenever possible.

### References

- Anderson, N. H. (1971). Integration theory and attitude change. *Psychological Review*, *78*, 171–206.
- Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.
- Baeyens, F., Eelen, P., & Crombez, G. (1995). Pavlovian associations are forever: On classical conditioning and extinction. *Journal of Psychophysiology*, *9*, 127–141.
- Bargh, L. A., Chaiken, S., Gvender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology*, *62*, 893–912.
- Bassili, J. N. (1996). Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology*, *71*, 637–653.
- Bell, D. W., & Esses, V. M. (2002). Ambivalence and response amplification: A motivational perspective. *Personality and Social Psychology Bulletin*, *28*, 1143–1152.
- Bentler, P. M., & Speckart, G. (1979). Models of attitude-behavior relations. *Psychological Review*, *86*, 452–464.
- Byrne, D. (1971). *The attraction paradigm*. New York: Academic Press.
- Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, *1*, 3–25.
- Cacioppo, J. T., Marshall-Goodell, B. S., Tassinary, L. G., & Petty, R. E. (1992). Rudimentary determinants of attitudes: Classical conditioning is more effective when prior knowledge about the attitude stimulus is low than high. *Journal of Experimental Social Psychology*, *28*, 207–233.
- Davidson, A. R., & Jaccard, J. (1979). Variables that moderate the attitude-behavior relation: Results of a longitudinal survey. *Journal of Personality and Social Psychology*, *37*, 1364–1376.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.
- Devine, P. G., Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology*, *60*, 817–830.
- Dovidio, J. F., Kawakami, K., & Beach, K. R. (2001). Implicit and explicit attitudes: Examination of the relationship between measures of intergroup bias. In R. Brown & S. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes* (pp. 175–197). Oxford, England: Blackwell Publishers.
- Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, *33*, 510–540.
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace.
- Fabrigar, L. R., MacDonald, T. K., & Wegener, D. T. (2005). The origin and structure of attitudes. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes and attitude change* (pp. 79–124). Mahwah, NJ: Erlbaum.
- Fazio, R. H. (1993). Variability in the likelihood of automatic attitude activation: Data reanalysis and commentary on Bargh, Chaiken, Gvender, and Pratto (1992). *Journal of Personality and Social Psychology*, *64*, 753–758.
- Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Mahwah, NJ: Erlbaum.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027.
- Fazio, R. H., & Olson, M. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology*, *54*, 297–327.
- 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.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison Wesley.
- Fiske, S. T., & Pavelchak, M. A. (1986). Category-based versus piecemeal-based affective responses: Developments in schema-triggered affect. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 167–203). New York: Guilford Press.
- Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61–89). Orlando, FL: Academic Press.
- Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, *46*, 107–119.
- Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology*, *59*, 601–613.
- Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology*, *65*, 221–233.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.
- Hasson, U., Simmons, J. P., & Todorov, A. (2005). Believe it or not: On

- the possibility of suspending belief. *Psychological Science*, 16, 566–571.
- Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review*, 93, 258–268.
- Ito, T. A., Cacioppo, J. T., & Lang, P. J. (1998). Eliciting affect using the International Affective Picture System: Trajectories through evaluative space. *Personality and Social Psychology Bulletin*, 24, 855–879.
- Jarvis, W. B. G. (1997). MediaLab (Version 3.0 Beta) and DirectRT (Version 1.0 Beta) [Computer software]. Columbus, OH: Empirisoft.
- Jarvis, W. B. G. (1998). *Do attitudes really change?* Unpublished doctoral dissertation, Ohio State University.
- Jarvis, W. B. G. (2000). MediaLab [Computer software]. Columbus, OH: Empirisoft.
- Jarvis, W. B. G., Tormala, Z. L., & Petty, R. E. (1998, May). *Do attitudes really change?* Paper presented at the annual meeting of the American Psychological Society, Washington, DC.
- Jonas, K., Diehl, M., & Bromer, P. (1997). Effects of attitudinal ambivalence on information processing and attitude-intention consistency. *Journal of Experimental Social Psychology*, 33, 190–210.
- Kaplan, K. J. (1972). On the ambivalence–indifference problem in attitude theory and measurement: A suggested modification of the semantic differential technique. *Psychological Bulletin*, 77, 361–372.
- Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, 78, 871–888.
- Kumkale, G. T., & Albarracín, D. (2004). The sleeper effect in persuasion: A meta-analytic review. *Psychological Bulletin*, 130, 143–172.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1995). *International Affective Picture System (IAPS): Technical manual and affective ratings*. Gainesville: Center for Research in Psychophysiology, University of Florida.
- Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology*, 81, 684–696.
- Leone, L., Perugini, M., & Ercolani, A. P. (1999). A comparison of three models of attitude–behavior relationships in the studying behavior domain. *European Journal of Social Psychology*, 29, 161–189.
- Maio, G. R., Bell, D. W., & Esses, V. M. (1996). Ambivalence in persuasion: The processing of messages about immigrant groups. *Journal of Experimental Social Psychology*, 32, 513–536.
- Maio, G. R., & Olson, J. M. (1995). The effect of attitude dissimulation on attitude accessibility. *Social Cognition*, 13, 127–144.
- Mayo, R., Schul, Y., & Burnstein, E. (2004). “I am not guilty” versus “I am innocent”: Successful negation may depend on the schema used for its encoding. *Journal of Experimental Social Psychology*, 40, 433–449.
- McClelland, J. L., McNaughton, B. L., & O’Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102, 419–457.
- McGuire, W. J. (1969). The nature of attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (2nd ed., Vol. 3, pp. 136–314). Reading, MA: Addison Wesley.
- Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice reduction effects. *Journal of Personality and Social Psychology*, 65, 469–485.
- Monteith, M. J., Devine, P. G., & Zuwerink, J. R. (1993). Self-directed versus other-directed affect as a consequence of prejudice-related discrepancies. *Journal of Personality and Social Psychology*, 64, 198–210.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.
- Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, 124, 54–74.
- Payne, B. K., Jacoby, L. L., & Lambert, A. J. (2005). Attitudes as accessibility bias: Dissociating automatic and controlled processes. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 393–420). Oxford, England: Oxford University Press.
- Petty, R. E., Baker, S., & Gleicher, F. (1991). Attitudes and drug abuse prevention: Implications of the elaboration likelihood model of persuasion. In L. Donohew, H. E. Sypher, & W. J. Bukoski (Eds.), *Persuasive communication and drug abuse prevention* (pp. 71–90). Hillsdale, NJ: Erlbaum.
- Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on response to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology*, 46, 69–81.
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19, 123–205.
- Petty, R. E., Fleming, M. A., & White, P. (1999). Stigmatized sources and persuasion: Prejudice as a determinant of argument scrutiny. *Journal of Personality and Social Psychology*, 76, 19–34.
- Petty, R. E., Gleicher, F. H., & Jarvis, W. B. G. (1993). Persuasion theory and AIDS prevention. In J. B. Pryor & G. Reeder (Eds.), *The social psychology of HIV infection* (pp. 155–182). Hillsdale, NJ: Erlbaum.
- Petty, R. E., & Jarvis, W. B. G. (1998, October). *What happens to the old attitude when attitudes change?* Paper presented at the annual meeting of the Society for Experimental Social Psychology, Lexington, KY.
- Petty, R. E., & Krosnick, J. A. (1995). *Attitude strength: Antecedents and consequences*. Mahwah, NJ: Erlbaum.
- Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 1, pp. 323–390). New York: McGraw-Hill.
- Petty, R. E., Wheeler, S. C., & Tormala, Z. L. (2003). Persuasion and attitude change. In T. Millon & M. J. Lerner (Eds.), *Handbook of psychology; Vol. 5. Personality and social psychology* (pp. 353–382). Hoboken, NJ: Wiley.
- Priester, J. M., & Petty, R. E. (1995). Source attributions and persuasion: Perceived honesty as a determinant of message scrutiny. *Personality and Social Psychology Bulletin*, 21, 637–654.
- Priester, J. M., & Petty, R. E. (1996). The gradual threshold model of ambivalence: Relating the positive and negative bases of attitudes to subjective ambivalence. *Journal of Personality and Social Psychology*, 71, 431–449.
- Priester, J. M., & Petty, R. E. (2001). Extending the bases of subjective attitudinal ambivalence: Interpersonal and intrapersonal antecedents of evaluative tension. *Journal of Personality and Social Psychology*, 80, 19–34.
- Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, 32, 880–892.
- Ross, M., & Conway, M. (1986). Remembering one’s own past: The construction of personal histories. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 122–144). New York: Guilford Press.
- Schwarz, N., & Bohner, G. (2000). The construction of attitudes. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intrapersonal processes* (pp. 436–457). Oxford, England: Blackwell Publishers.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3–22.
- Smith, E. R. (1998). Mental representation and memory. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 391–445). New York: McGraw-Hill.
- Smith, E. R., & DeCoster, J. (2000). Dual process models in social and

- cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, 4, 108–131.
- Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, 11, 1–74.
- Staats, A. W., & Staats, C. K. (1958). Attitudes established by classical conditioning. *Journal of Abnormal and Social Psychology*, 57, 37–40.
- Sutton, S. R., & Hallet, R. (1989). Understanding seat-belt intentions and behavior: A decision-making approach. *Journal of Applied Social Psychology*, 19, 1310–1325.
- Thompson, M. M., Zanna, M. P., & Griffin, D. W. (1995). Let's not be indifferent about (attitudinal) ambivalence. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 361–386). Mahwah, NJ: Erlbaum.
- Triandis, H. C. (1971). *Attitude and attitude change*. New York: Wiley.
- Tybout, A. M., Calder, B. J., & Sternthal, B. J. (1981). Using information processing theory to design marketing strategies. *Journal of Marketing Research*, 28, 73–79.
- Vargas, P. T., von Hippel, W., & Petty, R. E. (2004). Using partially structured attitude measures to enhance the attitude–behavior relationship. *Personality and Social Psychology Bulletin*, 30, 197–211.
- Wegner, D. M. (1989). *White bears and other unwanted thoughts*. New York: Viking Press.
- Wilson, T. D., & Hodges, S. D. (1992). Attitudes as temporary constructions. In L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 37–65). Hillsdale, NJ: Erlbaum.
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107, 101–126.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9(2, Pt. 2), 1–27.

Received December 31, 2004

Revision received July 8, 2005

Accepted August 1, 2005 ■