A Process Model of Affect Misattribution

Personality and Social
Psychology Bulletin
36(10) 1397–1408

DOI: 10.1177/0146167210383440
http://pspb.sagepub.com
SAGE

B. Keith Payne[1], Deborah L. Hall[2], C. Daryl Cameron[1], and Anthony J. Bishara[3]

Abstract

People often misattribute the causes of their thoughts and feelings. The authors propose a multinomial process model of affect misattributions, which separates three component processes. The first is an affective response to the true cause of affect. The second is an affective response to the apparent cause. The third process is when the apparent source is confused for the real source. The model is validated using the affect misattribution procedure (AMP), which uses misattributions as a means to implicitly measure attitudes. The model illuminates not only the AMP but also other phenomena in which researchers wish to model the processes underlying misattributions using subjective judgments.

Keywords

misattribution, priming, affect misattribution procedure, multinomial model, implicit attitudes

Received January 8, 2010; revision accepted May 11, 2010

The origin of our thoughts and feelings is all too often a mystery. It is important to solve because the meaning of those thoughts and feelings depends on their source. For example, when someone feels unhappy, it is important to know whether the feeling of unhappiness is caused by one's job, one's marriage, or a rainy day. The changes made would be very different. Yet experiences do not come complete with return addresses, and so people must make inferences about their causes. When those inferences are mistaken, interesting illusions follow.

Misattributions are sometimes responsible for false memories, as when eyewitnesses confuse the suggestion of an interrogator for memories of an actual event (Loftus, 1975). The déjà vu experience may result from a different sort of misattribution, in which a scene is processed fluently because of an initial quick glance, and that fluency is misattributed to having experienced the event before (Brown & Marsh, 2009; Jacoby & Whitehouse, 1989). Wegner (2002) has proposed that the experience of conscious will is a similar misattribution, in which conscious thoughts preceding actions are mistaken for the true causes of action, which are typically unconscious.

Beyond these cognitive illusions, misattributions have been used to explain affective biases, such as the tendency to express greater life satisfaction on sunny days than on rainy days (Schwarz & Clore, 1983). In related research on the misattribution of arousal, subjects were more aroused by erotic films if they had recently exercised (Cantor, Zillman, & Bryant, 1975), and hikers were more attracted to an experimenter after crossing a precarious footbridge than before crossing it (Dutton & Aron, 1974). This kind of misattribution may underlie the tendency to make more cautious decisions after experiencing incidental fear (Beer, Knight, & D'Esposito, 2006; Lerner & Keltner, 2001), more risky decisions following experiences of anger (Lerner & Keltner, 2001), and more severe moral judgments after experiencing irrelevant disgust (Schnall, Haidt, Clore, & Jordan, 2008). Understanding the mechanisms behind misattributions is an important step toward explaining how people can become confused about what they feel, why they feel that way, and what it means for how they should act.

In this article we develop a multinomial model for separating the processes underlying misattributions, with a focus on misattributions of affect. We begin our analysis by observing that any misattribution involves three elements. First, there is the true cause of the thought or feeling. Second, there is the apparent cause. And third, there is the mistaking of one for the other. For example, in Schwarz and Clore's (1983) study subjects were called either on a sunny day or a rainy day (the actual cause of the subjects' feelings), and they were

[1]University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[2]Arizona State University, Glendale, AZ, USA
[3]College of Charleston, Charleston, SC, USA

Corresponding Author:
B. Keith Payne, University of North Carolina at Chapel Hill, Department of Psychology, Campus Box 3270, Chapel Hill, NC 27599
Email: payne@unc.edu

asked about how their daily lives were going (the apparent cause). When subjects' attention was drawn to the real cause by first asking how the weather was, the weather did not influence judgments. But when there was no mention of the weather, it affected judgments of life satisfaction (a misattribution of the apparent cause for the real cause).

Formal models have been developed for the study of attributions in memory (Batchelder & Reifer, 1990; Jacoby, Bishara, Hessels, & Toth, 2005). However, these models presuppose correct and incorrect responses, and they are designed to model accuracy data. No similar models have been developed for subjective judgments such as life satisfaction, risky decisions, moral judgments, attitudes, and other such judgments of interest for social and personality psychology. We validate the model using the affect misattribution procedure (AMP), which uses affect misattributions as a means to measure automatic influences of attitudes (Payne, Cheng, Govorun, & Stewart, 2005). This provides an opportunity not only to model affect misattributions in subjective judgments but also to apply the model to advance implicit attitude measurement.

The Affect Misattribution Procedure

The AMP relies on the fact that people have difficulty disentangling their affective responses to two events occurring in close proximity in time and space. When this happens, people confuse the sources of their affective responses. Murphy and Zajonc (1993) presented affectively charged pictures too briefly to be identified, followed by Chinese pictographs that subjects rated for pleasantness. They found that ratings of the pictographs were influenced by the valence of the prime photos. Payne and colleagues (2005) modified this paradigm by presenting the prime photos visibly and by using a binary judgment of pleasantness to maximize the ambiguity of the judgment. In addition, they added a direct warning to subjects to avoid any influence of the primes on their judgments. This instruction was intended to place intentionally controlled response strategies in opposition to automatic influences of the primes. To the extent that subjects are able to strategically avoid the influence of the prime, they should show no priming effects. But to the extent that primes affect responding despite intentions, the task provides evidence for automatic influences, counter to control attempts.

Evidence for construct validity comes from studies showing that the AMP is significantly related to explicit attitude measures, but only in the absence of motivations to control responses. Such motivations tend to distort explicit measures but leave AMP scores relatively unaffected (Gawronski, Peters, Brochu, & Strack, 2008; Imhoff & Banse, 2009; Payne, Burkley, & Stokes, 2008; Payne, Govorun, & Arbuckle, 2008). Predictive validity has been established in studies showing that AMP scores predict judgment and behavioral outcomes after controlling for explicit measures in studies of alcohol drinking (Payne, Govorun, et al., 2008), smoking (Payne, McClernon, & Dobbins, 2007), and the influence of racial prejudice on voting in the 2008 presidential election (Payne et al., 2009). The AMP provides an ideal paradigm for studying affect misattributions because it elicits robust misattribution effects in a within-subjects design that provides high statistical power for testing the fit of the proposed model. In addition, applying the model to the AMP provides a useful means for separating automatic affective responses toward the primes from other component responses. Disentangling these processes helps refine implicit measures of social cognition.

Disentangling Processes Underlying Implicit Tests

A number of formal models, such as the process dissociation procedure, have been used to estimate the contribution of automatic and controlled processes to implicit task performance (Jacoby, 1991; Payne, 2001). Related models include the quadruple process model (quad model; Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005) and the diffusion model of Klauer, Voss, Schmitz, and Teige-Mocigemba (2007; for discussions of the relationships between these models, see Bishara & Payne, 2009; Payne & Bishara, 2009). Formal models of this type have helped clarify a number of theoretical questions. For example, Govorun and Payne (2006) found that participants showed more race bias on an implicit task when their self-control resources were depleted. This finding presents a puzzle for perspectives that view implicit tests as reflecting purely automatic responses. Why, after all, should depleting self-control resources influence a purely automatic effect? A process dissociation analysis revealed that the effect of depletion was mediated by reductions in the controlled component of task performance. Self-control depletion did not increase automatic prejudice, but instead it interfered with the ability to prevent that prejudice from influencing behavior (also see Amodio, 2009; Amodio et al., 2004; Lambert et al., 2003).

Quantitative models have also clarified the sources of group differences in implicit test performance. It has frequently been observed that Whites show more anti-Black implicit bias than Blacks do and that older adults show more bias than younger adults do. It might be tempting to assume the same explanation for both of these effects: that Whites and older people are simply more prejudiced against Blacks. But the results of studies using process models have shown that the two effects are driven by distinct mechanisms. Stewart, von Hippel, and Radvansky (2009) used process dissociation to compare age and race differences in implicit bias. Older adults showed poorer control than young adults, but they did not differ in automatic race bias. In contrast, the difference in implicit test scores for White and Black respondents was driven entirely by differences in automatic attitudes (also see

Gonsalkorale, Sherman, & Klauer, 2009). These studies suggest that age differences in implicit prejudice reflect age-related declines in cognitive control, whereas differences between racial groups reflect different automatic responses.

As with models of memory attributions, the models developed for implicit attitude measures are limited to certain kinds of outcomes. Process dissociation and quad models have been developed for compatibility tasks in which the outcome is accuracy. Klauer and colleagues' (2007) diffusion model can be used with accuracy and response time data. However, none of the currently available process models are appropriate for use with subjective judgments. The model proposed here is validated using the AMP, but it is intended to apply to the wide range of misattribution phenomena in which outcomes are subjective judgments.

A Model for Separating Component Processes in Affect Misattribution

We formulated a multinomial model, which represents unobserved cognitive processes in a branching tree diagram. Each process is represented as a probability, which can be interpreted as the likelihood that the process contributes to behavioral responses. The tree diagram depicts the set of conditional probabilities that together are posited by the model to produce the observed patterns of responses. Because the behavioral responses are empirical data gathered in an experiment, the model allows researchers to solve the set of joint probability equations. The model's parameters are then compared to the data using a goodness-of-fit test to assess how well the model fits the data.

To understand the model, it is important to note that subjects are instructed to evaluate the pictographs without influence from the prime photos. The model claims that whenever subjects are successful at distinguishing their response to the prime from their response to the pictograph, the response is driven by their evaluation of the pictograph. But when subjects confuse their reaction to the prime with their reaction to the pictograph, the response is driven by the reaction to the prime.

The model depicted in Figure 1 assumes that for any judgment, a misattribution occurs with probability M, or it does not occur with probability 1 – M. The misattribution parameter in this model shares some similarity with the control parameter (C) in process dissociation models, in that misattribution reflects a failure to carry out the intended task by distinguishing between relevant cues (the pictograph) and irrelevant cues (the prime). However, the parameter is different in the sense that no speeded compatibility task is involved. Failures presumably occur because of an inability to distinguish the source of one's affect rather than because of a failure to resolve interference in a speeded test. If a misattribution occurs, then the response is driven by affect toward the prime. The affective response is favorable with probability A

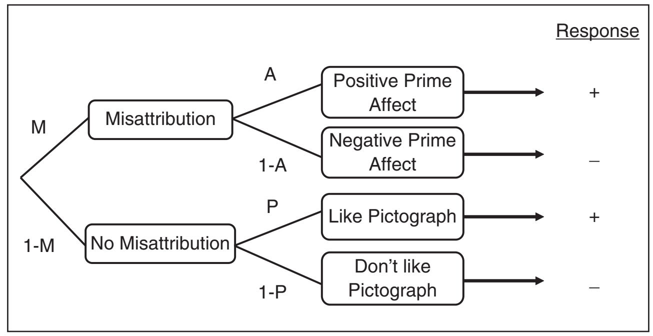

Figure 1. Multinomial process tree model of the affect misattribution procedure

Note: In the response column, + signifies a "pleasant" response and – signifies an "unpleasant" response.

(positive affect toward the prime) or unfavorable with probability 1-A. If no misattribution occurs, the response is driven by evaluations of the pictograph. The evaluation is favorable with probability P or unfavorable with probability 1-P. The model proposes that the A, M, and P parameters are sufficient to describe performance on the AMP. These parameters correspond to our general analysis of misattributions, in that A represents affective responses to the prime (the true cause of affect), P represents evaluations of the pictograph (the apparent cause of affect), and M represents confusing one for the other.

At least two criteria are important in evaluating models. First, the model must provide a good statistical fit to the data. And second, the model must produce theoretically meaningful process estimates. That is, the process estimates should be influenced by variables in ways that are theoretically predicted. In the model tests reported below, we manipulated variables expected to selectively influence individual parameters as a means of validating the model. To evaluate the model, we tested whether the model could successfully recover evaluations of the primes when the pictographs were presented slowly, thus reducing the likelihood of making a misattribution. In the first experiment, we used clearly pleasant and unpleasant primes to test whether the model could accurately estimate known affective reactions to the primes. In the second experiment, we tested whether estimates of affective reactions to the primes could predict individual differences in behavioral intentions as well or better than relying on task performance alone.

Experiment 1

In this experiment we examined the validity of the model by manipulating the pleasantness of the primes using photos that were normatively liked (e.g., puppies) and disliked (e.g., snakes). We also manipulated the likelihood of making a misattribution by varying the time for which the pictograph was presented. Prior research demonstrated that affective priming was reduced when the pictograph was presented for long durations (e.g., Payne et al., 2005). When the pictograph is presented slowly, respondents can presumably focus more on the features of the pictograph itself and better distinguish between their affective reactions to the pictographs as opposed to the primes. Finally, we manipulated the likelihood that subjects liked the pictograph themselves by selecting pictographs that tended to be more or less liked based on pretesting. If the model is valid, then the pleasantness of the primes should selectively influence the A parameter. The pictograph presentation time should selectively influence the M parameter. And the pleasantness of pictographs should selectively influence the P parameter. Critically, once the model is used to account for differences in misattributions (M), the presentation time of the pictographs should not influence affective reactions toward the primes (A). That is, although long presentation times are expected to obscure the effects of the primes in task performance, the model should be able to recover affective responses to the primes independent of the presentation rate.

Method

Participants. For partial course credit, 68 undergraduate students (24 men and 44 women) participated in the study.

Procedure. On arrival at the lab, participants were seated at a computer and introduced to the computer-based AMP. On each trial of the AMP, a pleasant or unpleasant photo (the prime) appeared on the center of the computer screen for 75 ms, followed by a Chinese pictograph, whose duration was manipulated as described below. Following the pictograph, a black-and-white pattern mask was presented that appeared as "static." The mask remained on the screen until participants indicated a pleasant or unpleasant response, at which point the next trial began. Participants were instructed to try their absolute best to ignore the photo and to sort each pictograph into the more pleasant half or the less pleasant half by pressing two keys marked "pleasant" and "unpleasant." The task consisted of 48 trials, plus one practice trial to demonstrate the trial sequence. Following the AMP, participants completed some demographic measures, indicated whether they could read the pictographs (none could), and were debriefed about the purpose of the study.

Design and materials. The design of the study was a 2 (prime: pleasant vs. unpleasant) × 2 (pictograph: pleasant vs. unpleasant) × 2 (pictograph duration: 100 ms vs. 1,000 ms) mixed design, with the duration of the pictograph manipulated between subjects and the other factors manipulated within subjects.

Chinese pictographs. To measure the pleasantness of the pictographs we used data from a pilot study.[1] The purpose of the pictographs is to serve as affectively neutral stimuli, to which feelings toward the primes may be misattributed. In manipulating the pleasantness of the pictographs, it was important not to use items that were so clearly pleasant or unpleasant that they would no longer be considered ambiguous. We therefore selected items whose pleasantness ratings were just slightly above or below average ratings. The overall average proportion of pleasant ratings was 57%; the more pleasant set was rated 64% pleasant and the less pleasant set was 51% pleasant.

Affective primes. In all, 24 positive images and 24 negative images were selected from the International Affective Picture System to serve as affective primes (Lang, Bradley, & Cuthbert, 1997). Each image in the database had been rated on a 1 to 9 pleasantness scale, with the pleasant items averaging 7.88 and the unpleasant items averaging 3.33.

Results

The main hypothesis was that the duration of the Chinese pictographs would moderate the impact of primes versus pictographs on responses. Primes were expected to have a larger impact when the duration was short compared to when it was long. But the pleasantness of the pictograph itself was expected to have a larger impact when the duration was long. For purposes of implicit attitude measurement, these effects are nuisance variables that interfere with the ability to assess automatic affective responses to the primes. However, it was expected that parameter estimates derived from the multinomial model should uncover affective responses to the primes, uncontaminated by features of the pictographs and independent of their duration. By separating the influence of the primes from the influence of the pictographs, the model was expected to reveal lower rates of misattribution (not less intense affective responses to the primes) in the long presentation condition. We first report the behavioral results, followed by the model testing.

Affective priming in behavioral responses. Priming was scored by computing the percentage of pleasant responses to the pictographs in each condition. Results were analyzed using a 2 (prime: pleasant vs. unpleasant) × 2 (pictograph: pleasant vs. unpleasant) × 2 (pictograph duration: fast vs. slow) analysis of variance (ANOVA). As expected, the main effect of prime pleasantness was significant, indicating that overall, the pictographs were rated more pleasant when they were preceded by pleasant primes than unpleasant primes, $F(1, 66) = 30.37$, $p < .001$. In addition, a significant main effect of pictograph pleasantness confirmed that pleasant pictographs were rated as more pleasant than unpleasant pictographs, $F(1, 66) = 51.21$, $p < .001$. These main effects were qualified by two predicted interactions. As displayed in Figure 2, prime pleasantness had a stronger effect when the pictographs were presented quickly, $F(1, 34) = 20.57$, $p < .001$, $\eta_p^2 = .38$, than when they were presented slowly, $F(1, 32) = 10.53$, $p < .01$, $\eta_p^2 = .25$. The expected Prime Pleasantness × Pictograph Duration interaction was marginally significant, $F(1, 66) = 3.60$, $p = .06$.

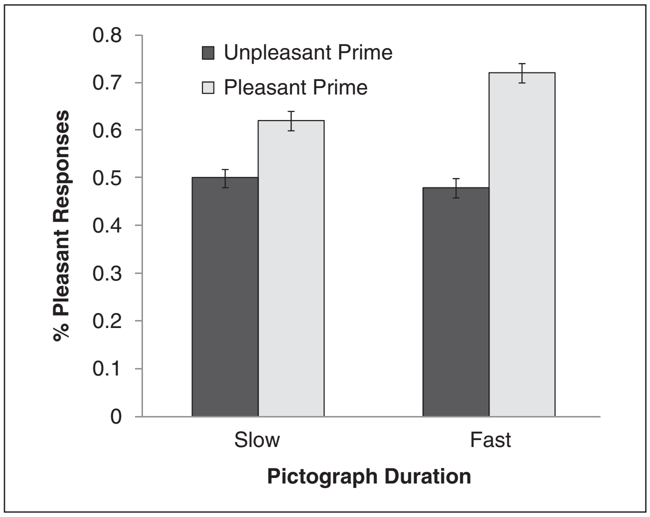

Figure 2. Proportion of pleasant and unpleasant responses as a function of prime pleasantness and pictograph duration in Experiment 1
Note: Error bars represent 1 *SEM*.

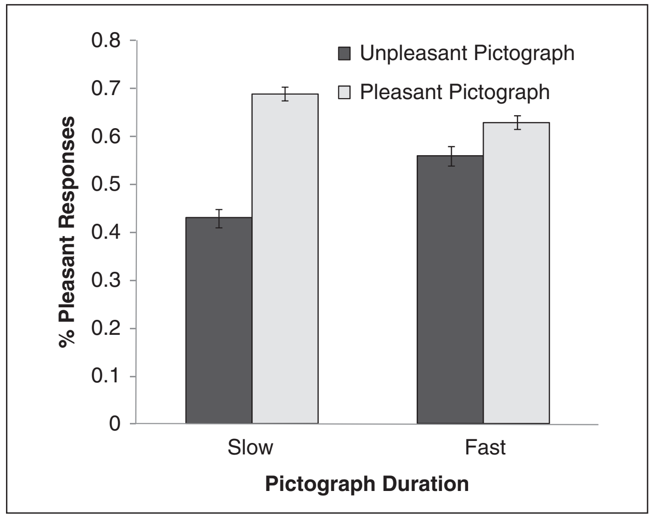

Figure 3. Proportion of pleasant and unpleasant responses as a function of pictograph pleasantness and pictograph duration in Experiment 1
Note: Error bars represent 1 *SEM*.

The duration manipulation had the opposite effect for pictograph pleasantness. As displayed in Figure 3, pictograph pleasantness had a larger impact when the pictographs were presented slowly, $F(1, 32) = 55.16$, $p < .001$, $\eta_p^2 = .63$, than when they were presented quickly, $F(1, 34) = 4.88$, $p < .05$, $\eta_p^2 = .13$. The Pictograph Pleasantness × Pictograph Duration interaction was significant, $F(1, 66) = 18.66$, $p < .001$. To summarize, these patterns suggest that as the duration became longer, the influence of the primes diminished and the influence of the pictograph features increased.

In addition to these predicted effects, there was also an unanticipated Prime Pleasantness × Pictograph Pleasantness × Pictograph Duration interaction, $F(1, 66) = 4.98$, $p < .05$. We resist drawing conclusions from this interaction because it did not replicate in Experiment 2. But for completeness, the full pattern of means is displayed in Table 1. This interaction indicates that when the pictographs were presented quickly, prime pleasantness and pictograph pleasantness had additive effects, with no interaction, $F(1, 34) = 2.06$, $p = .16$. But when the pictographs were presented slowly, there was a marginally significant interaction between prime pleasantness and pictograph pleasantness, $F(1, 32) = 2.88$, $p = .10$. In the slow duration condition, prime valence had a somewhat larger effect on evaluations of pleasant pictographs, $F(1, 32) = 11.43$, $p < .01$, $\eta_p^2 = .26$, compared to unpleasant pictographs, $F(1, 32) = 3.30$, $p = .08$, $\eta_p^2 = .09$. This interaction reflects a difference in the magnitude of priming effects, rather than a qualitative difference, because affective priming was present (at marginally significant levels) even in the slow duration/unpleasant pictograph condition.

Tests of the multinomial model. We fitted the multinomial model to the data from Experiment 1 and tested the model fit using a G^2 statistic. The equations that make up the model are included in the appendix. We began by fitting a saturated model that had a unique A, M, and P parameter for every cell in the 2 × 2 × 2 design. (The saturated model cannot be tested because it has more parameters than data cells and hence no degrees of freedom.) We then constrained the parameters based on theoretical assumptions and examined whether the model continued to fit under these constraints. Specifically, we allowed M to vary across short versus long symbol duration but not across pleasant versus unpleasant primes or pleasant versus unpleasant pictographs. Second, we allowed A to vary across pleasant and unpleasant primes but constrained it to be equal across pleasant and unpleasant pictographs and long versus short duration. Finally, we allowed P to vary across pleasant versus unpleasant symbols but constrained it to be equal across prime and duration conditions. The model constrained in this way tests whether each parameter can be cleanly dissociated from the other parameters. We refer to this particular set of parameter constraints as the "model of interest" and focus on it as a point of comparison.

The model of interest fit the data well, $G^2(2) = 4.29$, $p > .05$, critical value = 5.99. As displayed in Table 2, the parameter estimates showed the expected patterns. To determine which manipulations significantly influenced parameter estimates, we examined a series of nested models with additional parameter constraints. Compared to the model of interest, constraining the M parameter to be the same across durations led to a significant increase in G^2, $\Delta G^2(1) = 49.86$, $p < .001$. In other words, the misattribution rate, M, was significantly higher when the pictographs were presented quickly than when presented slowly. Furthermore, positive affect toward the prime, A, was significantly higher when the

Table 1. Probability of a Pleasant Response, Experiment 1

	Unpleasant prime		Pleasant prime	
	Unpleasant pictograph	Pleasant pictograph	Unpleasant pictograph	Pleasant pictograph
Slow duration	.39	.61	.46	.77
Fast duration	.43	.52	.70	.74

Table 2. Parameter Estimates for the Model of Interest, Experiment 1

	Slow duration			
	Unpleasant prime		Pleasant prime	
Parameter	Unpleasant pictograph	Pleasant pictograph	Unpleasant pictograph	Pleasant pictograph
A	.47	.47	.74	.74
M	.43	.43	.43	.43
P	.29	.75	.29	.75
	Fast duration			
	Unpleasant prime		Pleasant prime	
	Unpleasant pictograph	Pleasant pictograph	Unpleasant pictograph	Pleasant pictograph
A	.47	.47	.74	.74
M	.87	.87	.87	.87
P	.29	.75	.29	.75

Note: A = positive affect toward prime; M = misattribution rate; P = evaluation of pictograph.

prime was pleasant than when it was unpleasant, $\Delta G^2(1) = 124.69$, $p < .001$. Finally, liking for the pictographs, P, was significantly higher for the pleasant pictographs than for the unpleasant pictographs, $\Delta G^2(1) = 120.10$, $p < .001$.

To further examine the constraints in the model of interest, it was also compared to models where the constraints were loosened. For example, when A for the unpleasant prime was allowed to vary separately for slow and fast durations, the model fit did not significantly improve, $\Delta G^2(1) < .01$, $p > .99$. In other words, A for the unpleasant prime was not significantly influenced by duration. Likewise, relaxing any of the other individual parameter constraints in the model of interest failed to significantly improve model fit, all $\Delta G^2(1) < 3.24$, all ps > .05. Even if two parameter constraints were relaxed simultaneously, the fit of the model of interest was so good that it was impossible to significantly improve the fit, all $\Delta G^2(2) < 4.29$, all ps > .05.

Discussion

Results of Experiment 1 suggest that the proposed multinomial model successfully captured the processes underlying performance in the AMP. The model fit the data well, and the pattern of estimates derived from the model showed theoretically predicted patterns. Considering behavioral responses, faster presentation allowed greater influence of the primes, but slower presentation allowed greater influence of the pictographs. The model successfully explained these differences as resulting from different rates of misattribution in the fast versus slow presentation conditions. Compared to the model, behavioral performance underestimated the size of the affective priming effect, particularly in the slow presentation condition. The difference between behavioral responses and model estimates in the long presentation condition can be accounted for by the higher rate of misattributions when the pictograph presentation was fast.

These results suggest that when the pictograph was presented slowly, the primes were not evaluated any differently; instead, their impact on evaluations of the pictograph was masked because judgments were then based heavily on the qualities of the pictograph itself. By applying the model, we were able to correct for this masking effect and recover the underlying affective reactions to the primes. Having demonstrated that the model can adequately explain affective priming data, we next tested whether the model could uncover individual differences in attitudes that were predictive of behavior.

Experiment 2

In this experiment, we sought to replicate the main finding of Experiment 1, that the model could separate affective reactions to primes from the influences of the pictographs and misattribution rates. However, rather than examining unanimously pleasant and unpleasant items, we examined individual differences in attitudes. We took advantage of the 2008 U.S. presidential election to test whether automatic responses toward the candidates would predict voting intentions. Previous research has shown that implicit and explicit responses toward political candidates tend to be strongly related (Nosek, 2005; Payne et al., 2005). As a result, we could examine the relationship between implicit responses and self-reported voting intentions without accounting for social desirability or other factors that interfere with self-reports of behavioral intentions. Data were collected in the weeks immediately before the election.

Method

Participants. For partial course credit, 55 undergraduate students (11 males and 44 females) participated in the study.

Procedure and design. Participants were recruited to take part in a study on political attitudes. After demographic information was collected, subjects were asked for whom they intended to vote. The options included Barack Obama, John McCain, and "another candidate." All subjects expressed an intention to vote for either Obama or McCain.

Next, a version of the AMP was administered in which photographs of John McCain and Barack Obama served as the primes. On each of 60 trials, a photograph of either McCain or Obama appeared on the screen for 75 ms, followed by a blank screen for 125 ms, a Chinese pictograph, and then a black-and-white pattern mask. Participants were instructed that the prime photos could influence their judgments, and they were warned to try their absolute best not to let the photos affect their pleasantness judgments of the pictographs. The presentation duration of the Chinese pictographs was manipulated between subjects (100 ms or 1,000 ms). The candidate photos and the pleasantness of the Chinese pictographs were manipulated within subjects, yielding a 2 (prime: McCain vs. Obama) × 2 (pictograph pleasantness: pleasant vs. unpleasant) × 2 (pictograph duration: short vs. long) mixed design.

Materials. The same pictograph selections as in Experiment 1 were used to manipulate the pleasantness of the target items. We first selected photos of the candidates based on a pilot test. Selected photos were judged to be typical of how the candidates normally look, and they were matched to minimize differences in attractiveness.[2] The final materials included eight matched photos of each candidate.

Results

No participants indicated that they knew the semantic meaning of the Chinese pictographs. One participant's response times to the AMP were more than three standard deviations

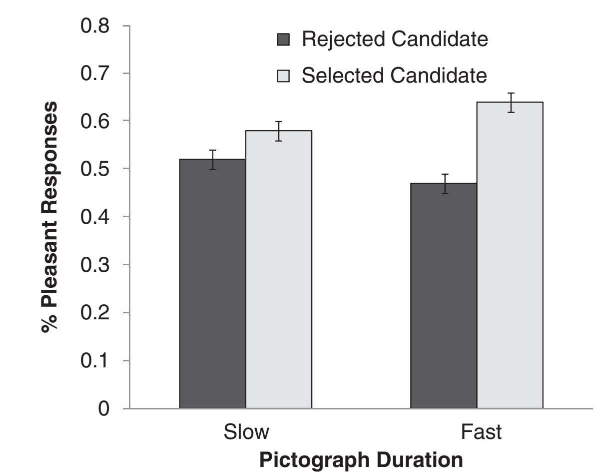

Figure 4. Proportion of pleasant and unpleasant responses as a function of prime candidate and pictograph duration in Experiment 2

Note: Error bars represent 1 *SEM*.

longer than the mean response time, and this subject's data were excluded from analyses.

Behavioral responses. Voting intentions were very close to actual election results, with 53% of subjects intending to vote for Obama and 47% intending to vote for McCain (Obama won the popular vote by 53% to 46%). We recoded the candidate primes based on voting intentions to reflect each subject's selected candidate and rejected candidate. We then analyzed the proportion of pleasant responses on the AMP using a 2 (prime: selected vs. rejected candidate) × 2 (pictograph pleasantness: pleasant vs. unpleasant) × 2 (pictograph duration: fast vs. slow) ANOVA. We expected responses to the candidate primes to be consistent with voting intentions and that this difference would be larger when pictographs were presented quickly. In addition, we expected that pictograph pleasantness would have a larger impact when the pictographs were presented slowly.

As predicted, a significant main effect of the candidate prime indicated that the selected candidate primed more positive responses than the rejected candidate, $F(1, 53) = 16.33$, $p < .001$, $\eta_p^2 = .24$. This effect demonstrates that AMP responses were consistent with voting intentions.[3] Also as expected, the predictive validity of the AMP was qualified by the duration of the pictograph: Prime × Pictograph duration, $F(1, 53) = 4.26$, $p < .05$, $\eta_p^2 = .07$ (see Figure 4). Simple effects tests showed that the main effect of selected versus rejected candidate primes was significant in the fast condition, $F(1, 26) = 12.84$, $p < .001$, $\eta_p^2 = .33$, but only marginally significant in the slow condition, $F(1, 27) = 3.37$, $p = .08$, $\eta_p^2 = .11$.

In addition, subjects responded more favorably to pleasant pictographs than unpleasant pictographs, $F(1, 53) = 15.16$, $p < .001$, $\eta_p^2 = .22$. This effect was moderated by

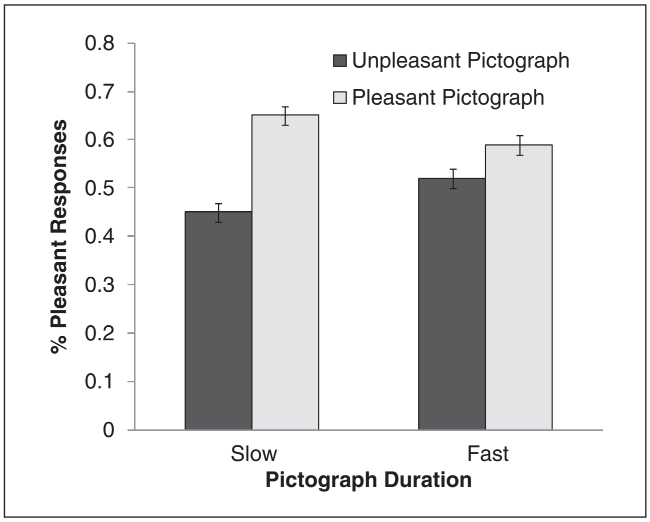

Figure 5. Proportion of pleasant and unpleasant responses as a function of pictograph pleasantness and pictograph duration in Experiment 2
Note: Error bars represent 1 *SEM*.

pictograph duration, $F(1, 53) = 4.26$, $p < .05$, $\eta_p^2 = .07$ (see Figure 5). Pictograph pleasantness had a significant effect when pictographs were presented slowly, $F(1, 27) = 13.02$, $p < .001$, $\eta_p^2 = .33$, but only a marginally significant effect when presented quickly, $F(1, 26) = 2.75$, $p = .10$. These effects demonstrate that priming responses were consistent with behavioral intentions but that the predictive validity of the AMP was greater when the pictographs were presented quickly. When they were presented slowly, the pleasantness of the pictographs had greater impact, replicating Experiment 1.

Tests of the multinomial model. We fitted the model using the same procedure as in Experiment 1, by first estimating a fully saturated model and then constraining the parameters based on theoretical predictions. As in Experiment 1, we constrained the model such that the M parameter was allowed to vary only between long and short pictograph presentations, the P parameter was allowed to vary only between pleasant and unpleasant pictographs, and the A parameter was allowed to vary only for selected and unselected candidate primes. This model of interest fit the data very well, $G^2(2) = 0.61$, $p > .05$, critical value = 5.99, and it generated meaningful parameter estimates, as displayed in Table 3. The A parameter was significantly higher for the selected candidate than the rejected candidate, $\Delta G^2(1) = 50.57$, $p < .001$. The M parameter was significantly higher for fast than slow durations, $\Delta G^2(1) = 26.80$, $p < .001$. Finally, the P parameter was significantly higher for pleasant pictographs than unpleasant ones, $\Delta G^2(1) = 76.01$, $p < .001$. In addition, the fit of the model of interest was so good that it was impossible to significantly improve on it by relaxing the parameter constraints, all $\Delta G^2 < 0.61$, all *p*s > .05.

Discussion

Behavioral responses in the AMP were reliably associated with voting intentions, but this relationship was attenuated when the pictographs were presented slowly. If we had looked only at responses under the slow presentation conditions, we might have concluded that affective responses to the primes were not good predictors of attitudes or behavioral intent. However, an alternative explanation is that affective responses to the primes were masked by intentional evaluations of the pictographs themselves (consistent with task instructions). By this account, subjects had attitude-consistent affective reactions, but these reactions did not carry over to evaluations of the pictographs when the pictographs were presented slowly. By applying the multinomial model, we were able to successfully remove the influence of the pictographs themselves to produce a more accurate estimate of attitudes. Model estimates of affective responses to the primes differentiated voters about as well as behavioral priming effects in the fast presentation condition, and considerably better than behavioral responses in the slow presentation condition. The evidence for this conclusion comes from the fact that for behavioral responses the preference-consistent priming effect was qualified by duration, but A was not different in the fast versus slow presentation conditions. These findings suggest that the model may be especially useful when studying individuals or groups who differ in vulnerability to misattributions, as we discuss in more detail below.

General Discussion

In two experiments, we validated a new model of the processes underlying misattributions in the AMP. In Experiment 1, we used clearly pleasant and unpleasant primes to establish the meanings of each model parameter. The A parameter, meant to reflect affective responses to the primes, was influenced selectively by the valence of the primes. The M parameter, meant to reflect the rate of misattributions of affect from the primes to the pictographs, accounted for the effects of presentation duration. Finally, the P parameter, meant to reflect evaluations of the pictographs, was influenced selectively by the pleasantness of the pictographs as defined by a pretest. In Experiment 2, we applied the model to individual differences in political attitudes. The model successfully distinguished subjects based on their voting intentions. Here, too, the A parameter was unaffected by the duration of the pictographs, showing that underlying evaluations of the primes could be recovered even when they were hidden in behavioral responses.

Implications for the Mechanisms of Misattribution

Our model provides a new framework for studying a broad set of misattribution effects in which there is no correct

Table 3. Parameter Estimates for the Model of Interest, Experiment 2

	Slow duration			
	Prime: Rejected candidate		Prime: Selected candidate	
Parameter	Unpleasant pictograph	Pleasant pictograph	Unpleasant pictograph	Pleasant pictograph
A	.45	.45	.67	.67
M	.25	.25	.25	.25
P	.41	.68	.41	.68
	Fast duration			
	Prime: Rejected candidate		Prime: Selected candidate	
	Unpleasant pictograph	Pleasant pictograph	Unpleasant pictograph	Pleasant pictograph
A	.45	.45	.67	.67
M	.77	.77	.77	.77
P	.41	.68	.41	.68

Note: A = positive affect toward prime; M = misattribution rate; P = evaluation of pictograph.

answer. Several phenomena of interest to social psychologists may reflect a similar processing architecture. For example, classic findings on the misattribution of arousal (Cantor et al., 1975; Dutton & Aron, 1974), misattributions of emotion and mood (Schachter & Singer, 1962; Schwarz & Clore, 1983), and mood as input (Martin, Ward, Achee, & Wyer, 1993) seem likely to follow similar logic.

The model provides a testable theoretical account of the relationships among underlying processes that produce misattribution effects. That is, the model specifies a particular set of contingencies in which the apparent cause influences judgments only when a misattribution occurs; otherwise, the judgment is driven by the real source. Other processes and different relationships between processes are possible and could be evaluated by testing competing models against each other.

The model also aids in testing hypotheses about specific processes. For example, many misattribution paradigms are based on the assumption that directing attention to the true source of one's feelings reduces the likelihood that a misattribution will be made, whereas directing attention to an apparent cause increases the likelihood of misattribution. This assumption was reflected in the design of Schwarz and Clore's (1983) study, in which participants were either asked about the weather *before* they reported their happiness (which prevented misattribution from taking place) or attention was drawn to the weather as a potential determinant of one's happiness afterward. Yet considering these outcomes as opposite poles of a continuum ignores the possibility that misattribution, and the influences of the real and apparent causes, might operate independently of each other. Does directing attention to the real cause have its effects by reducing misattributions, by increasing the influence of the real cause, or by decreasing the influence of the apparent cause (or a combination of these)? Applying the model to classic paradigms such as misattributions of emotion (Schachter & Singer, 1962; Schwarz & Clore, 1983) and misattributions of arousal (Cantor et al., 1975; Dutton & Aron, 1974) would help clarify the role (or roles) of important variables such as the focus of attention and the ambiguity of the situation.

Implications for Implicit Measurement

In addition to advancing basic understanding of the processes driving misattributions, the model provides a useful tool for improving the quality of implicit measurement. In our studies, the model improved the sensitivity and predictive power of implicit attitudes measured with the AMP. Estimates of A (compared to behavioral AMP scores) showed increased sensitivity to the valence of primes in Experiment 1 and increased ability to differentiate Obama voters from McCain voters in Experiment 2. In both studies, we manipulated the presentation rate of the pictographs as a means of influencing the likelihood of misattributions. In effect, we created an obstacle to implicit attitude measurement to test whether the model could overcome that obstacle. In most contexts, of course, researchers would not deliberately include such a manipulation, but natural obstacles to measurement are common in ordinary use.

Consider the findings described in the introduction that older adults and younger adults differed in controlled components of implicit task performance (Gonsalkorale, Sherman, et al., 2009; Stewart et al., 2009). Given that older adults (Hashtroudi, Johnson, & Chrosniak, 1989; Henkel, Johnson, & De Leonardis, 1998; Jacoby et al., 2005) and children (Lindsay, Johnson, & Kwon, 1991) often have difficulty with source attribution, it is likely that age may also influence misattributions in the AMP. The extremity of priming effects among young adults may be underestimated compared to older adults and children, but applying the process model proposed here offers a potential solution. By accounting for differences in misattribution across age groups, the model may be able to more accurately measure attitudes. Other individual differences might also moderate the degree to which AMP responses reflect purely automatic reactions. For example, individuals with good attention control and executive functioning might be better able to avoid the influence of the primes as they are instructed to do. The model developed here shows promise for removing these influences to better isolate affective reactions to the primes.

An important next step is extending the model to estimate processes at the level of individual subjects. Our studies were designed to evaluate the model at the group level, which did not allow enough degrees of freedom to test the model for individual participants. Although this is beyond the scope of the present research, individual estimates would provide more flexibility by allowing researchers to examine individual difference correlations in addition to group comparisons. However, now that the model has been validated, it should be possible to generate individual estimates by manipulating both pictograph pleasantness and pictograph duration as within-subject variables. The main consideration is to ensure that the number of data cells for each subject is greater than the number of free parameters. For example, a 2 (prime) × 2 (pictograph pleasantness) × 2 (pictograph duration) within-subjects design would generate eight data cells for each subject. If two A parameters, two M parameters, and two P parameters were estimated for each subject the model would be identifiable, with two degrees of freedom. Validating such an individual-level model is an important direction for future research.

Conclusion

We validated a new process model of affect misattributions that successfully separated affective reactions to two sources (prime photos and target pictographs) and the likelihood of misattributing one for the other. The model can, in principle, be applied to any case of misattribution. Because people do not necessarily know the causes of their thoughts and feelings, they often confuse the causes and thus are influenced by logically irrelevant factors. When a person feels more attracted to a date because he or she has just watched an action movie, or when a person feel unsatisfied with life because it is rainy outside, that person is in the grip of a misattribution. Our model sheds light on exactly how subjects feel about the real cause of their reactions, how they feel about the apparent cause, and the extent to which they have confused one with the other.

Appendix

Modeling Methods and Equations

To ensure convergence on best fitting parameters and the smallest possible G^2, multinomial models were implemented with a quasi-Newton optimization method and multiple sets of random starting parameters. This was performed using Microsoft Excel's Solver add-on. Alpha was set to .05. With this alpha, power to detect small effect sizes (w = .1; Cohen, 1977) always exceeded .999 (Faul, Erdfelder, Lang, & Buchner, 2007).

The following general equations were used for the model:

$$p(\text{Pleasant response}) = \text{M}\cdot\text{A} + (1\text{–M})\cdot\text{P} \qquad \text{(A1)}$$

$$p(\text{Unpleasant response}) = \text{M}\cdot(1\text{–A}) + (1\text{–M})\cdot(1\text{–P}) \qquad \text{(A2)}$$

Declaration of Conflicting Interests

The authors declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Financial Disclosure/Funding

The authors received the following financial support for the research and/or authorship of this article: The first author acknowledges support from National Science Foundation Grant 0615478.

Notes

1. Pretest data were from a pilot study for the American National Election Studies (ANES) that included the affect misattribution procedure (AMP) as a measure of racial attitudes. The pilot study used a nationally representative sample of 538 American adults. The version of the AMP used in the ANES included photos of White and Black men as primes that were randomly paired with the pictographs. For the present study, we collapsed across the primes because we were interested in the average evaluation of each pictograph. The pilot study included 71 unique pictographs. The rate of "pleasant" responses for the pictographs (averaged across participants) ranged from 46% to 75% (M = 57%, SD = 7%). We selected 30 pictographs from the upper half of this distribution, resulting in a set with an average of 64% pleasant responses; and we selected 30 from the lower half of the distribution, resulting in a set with 51% pleasant responses. In this way, we selected relatively more pleasant and less pleasant pictographs that were nonetheless close to neutral ratings.
2. The photos were pretested as part of an ANES pilot study. We first selected 17 photos of each candidate from Internet sites. All photos showed the candidate in business attire, and the photos included a wide range to poses and background contexts. In the pilot study, 160 subjects rated each photo on typicality and attractiveness. Two items measured typicality: "How unusual is this picture of [candidate]?" (answered using five response options from *not at all unusual* to *extremely unusual*) and "How different does [candidate] look in this picture from the way he usually looks?" (answered using five response options from *not at all different* to *extremely different*). Two additional items measured attractiveness: "How attractive does [candidate] look in this picture?" (answered using five response options from *not attractive at all* to *extremely attractive*) and, "In this picture do you think [candidate] looks . . ." (answered using seven response options from *extremely good* to *extremely bad*). We first identified photographs that were rated highest in

typicality for each candidate (i.e., those with average ratings above 3 when the items were scored such that higher values represent greater typicality). This resulted in 12 typical photos of each candidate. Among typical photos, Obama's were rated as significantly more attractive than McCain's, $p < .05$. To reduce differences in attractiveness, we excluded the four most attractive Obama photos and the four least attractive McCain photos. This resulted in 8 photos of each candidate that were equivalently high on typicality (Obama $M = 3.31$, McCain $M = 3.30$) and relatively close on attractiveness (Obama $M = 4.21$, McCain $M = 3.98$; difference = 0.23, $p = .09$).

3. A supplementary analysis treated voting intention as a between-subjects variable rather than recoding the primes as selected versus rejected candidates. This analysis revealed a significant Voting Intention × Prime interaction, $p < .01$, reflecting preference-consistent priming. There was no main effect of prime, indicating that there was no net preference for either candidate. This crossover interaction indicates that implicit preferences were symmetrical: Obama voters displayed automatic preferences for Obama by about the same degree as McCain voters displayed automatic preferences for McCain. This symmetry justifies collapsing the preferences together into selected and rejected candidates for the purpose of model testing.

References

Amodio, D. M. (2009). Intergroup anxiety effects on the control of racial stereotypes: A psychoneuroendocrine analysis. *Journal of Experimental Social Psychology*, *45*, 60-67.

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science*, *15*, 88-93.

Batchelder, W., & Riefer, D. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97*, 548-564.

Beer, J. S., Knight, R. S., & D'Esposito, M. (2006). Controlling the integration of emotion and cognition. *Psychological Science*, *17*, 448-453.

Bishara, A. J., & Payne, B. K. (2009). Multinomial process tree models of control and automaticity in weapon misidentification. *Journal of Experimental Social Psychology*, *45*, 524-534.

Brown, A. S., & Marsh, E. J. (2009). Creating illusions of past encounter through brief exposure. *Psychological Science*, *20*, 534-538.

Cantor, J., Zillman, D., & Bryant, J. (1975). Enhancement of experienced sexual arousal in response to erotic stimuli through misattribution of unrelated residual excitation. *Journal of Personality and Social Psychology*, *32*, 69-75.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York, NY: Academic Press.

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. (2005). Separating multiple processes in implicit social cognition: The quad-model of implicit task performance. *Journal of Personality and Social Psychology*, *89*, 469-487.

Dutton, D. G., & Aron, A. P. (1974). Some evidence for heightened sexual attraction under conditions of high anxiety. *Journal of Personality and Social Psychology*, *30*, 510-517.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Gawronski, B., Peters, K., Brochu, P., & Strack, F. (2008). Understanding the relations between different forms of racial prejudice: A cognitive consistency perspective. *Personality and Social Psychology Bulletin*, *34*, 648-665.

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology*, *45*, 410-414.

Govorun, O., & Payne, B. K. (2006). Ego depletion and prejudice: Separating automatic and controlled components. *Social Cognition*, *24*, 111-136.

Hashtroudi, S., Johnson, M. K., & Chrosniak, L. D. (1989). Aging and source monitoring. *Psychology and Aging*, *4*, 106-112.

Henkel, L., Johnson, M., & De Leonardis, D. (1998). Aging and source monitoring: Cognitive processes and neuropsychological correlates. *Journal of Experimental Psychology: General*, *127*, 251-268.

Imhoff, R., & Banse, R. (2009). Ongoing victim suffering increases prejudice: The case of secondary anti-Semitism. *Psychological Science*, *20*, 1443-1447.

Jacoby, L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513-541.

Jacoby, L., Bishara, A., Hessels, S., & Toth, J. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older Adults. *Journal of Experimental Psychology: General*, *134*, 131-148.

Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, *118*, 126-135.

Klauer, K., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the Implicit Association Test: A diffusion-model analysis. *Journal of Personality and Social Psychology*, *93*, 353-368.

Lambert, A. J., Payne, B. K., Shaffer, L. M., Jacoby, L. L., Chasteen, A., & Khan, S. (2003). Stereotypes as dominant responses: On the "social facilitation" of prejudice in anticipated public contexts. *Journal of Personality and Social Psychology*, *84*, 277-295.

Lang, P. J., Bradley, M. M., & Cuthbert, B. (1995). International Affective Picture System. Gainesville: Center for Research in Psychophysiology. University of Florida.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*, 146-159.

Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology*, *52*, 297-318.

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, *7*, 560-572.

Martin, L. L., Ward, D. W., Achee, J. W., & Wyer, R. S. (1993). Mood as input: People have to interpret the motivational implications of their moods. *Journal of Personality and Social Psychology*, *64*, 317-326.

Murphy, S., & Zajonc, R. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, *64*, 723-739.

Nosek, B. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, *134*, 565-584.

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*, 181-192.

Payne, B. K., & Bishara, A. J. (2009). An integrative review of process dissociation and related models in social cognition. *European Review of Social Psychology*, *20*, 272-314.

Payne, B. K., Burkley, M., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology*, *94*, 16-31.

Payne, B., Cheng, C., Govorun, O., & Stewart, B. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*, 277-293.

Payne, B., McClernon, F., & Dobbins, I. (2007). Automatic affective responses to smoking cues. *Experimental and Clinical Psychopharmacology, 15*, 400-409.

Payne, B. K., Govorun, O., & Arbuckle, N. L. (2008). Automatic attitudes and alcohol: Does implicit liking predict drinking? *Cognition and Emotion*, *22*, 238-271.

Payne, B. K., Krosnick, J. A., Pasek, J., Lelkes, Y., Akhtar, O., & Tompson, T. (2009). Implicit and explicit prejudice in the 2008 American presidential election. *Journal of Experimental Social Psychology*, *46*, 367-374.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69*, 379-399.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality Social Psychology Bulletin*, *34*, 1096-1109.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513-523.

Stewart, B., von Hippel, W., & Radvansky, G. (2009). Age, race, and implicit prejudice: Using process dissociation to separate the underlying components. *Psychological Science*, *20*, 164-168.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.