Journal of Personality and Social Psychology
2005, Vol. 89, No. 3, 277–293


0022-3514/05/$12.00 DOI: 10.1037/0022-3514.89.3.277

# An Inkblot for Attitudes: Affect Misattribution as Implicit Measurement

B. Keith Payne, Clara Michelle Cheng, Olesya Govorun, and Brandon D. Stewart
The Ohio State University

Misattributions people make about their own affective reactions can be used to measure attitudes implicitly. Combining the logic of projective tests with advances in priming research, the affect misattribution procedure (AMP) was sensitive to normatively favorable and unfavorable evaluations (Experiments 1–4), and the misattribution effect was strong at both fast and slow presentation rates (Experiments 3 and 4). Providing further evidence of validity, the AMP was strongly related to individual differences in self-reported political attitudes and voting intentions (Experiment 5). In the socially sensitive domain of racial attitudes, the AMP showed in-group bias for Black and White participants. AMP performance correlated with explicit racial attitudes, a relationship that was moderated by motivations to control prejudice (Experiment 6). Across studies, the task was unaffected by direct warnings to avoid bias. Advantages of the AMP include large effect sizes, high reliability, ease of use, and resistance to correction attempts.

*Keywords:* attitude, implicit, automatic, projective, measurement, affective priming

People are notorious for their misattributions, especially when it comes to the workings of their own minds. Hikers have been known to mistake their fear over a precarious bridge for the allure of a stranger (Dutton & Aron, 1974). People misinterpret the fleeting pleasure of a sunny day as enduring life satisfaction (Schwarz & Clore, 1983). Eyewitnesses mistake the suggestions of misleading interrogators for their own memories (Wells & Loftus, 2003). Even people's feelings of conscious agency can be misattributed from observations of their own behavior (Gazzaniga, 1985; Wegner, 2003). The prevalence of misattributions across so many psychological domains suggests that it is an everyday aspect of mental life.

The idea behind classic projective tests such as the Rorschach inkblot test and the Thematic Apperception Test (TAT) can be seen as a kind of misattribution. Given an ambiguous event to interpret, people are believed to imbue the event with personal sources of meaning. Although the content comes (unconsciously) from the person, he or she perceives it as a quality of the event. This article reports a series of studies harnessing the human propensity for misattribution to achieve an implicit measure of attitudes in the spirit of projective tests.

Projective tests have a long and controversial history. A recent review (Lilienfield, Wood, & Garb, 2000) concluded, "the substantial majority of Rorschach and TAT indexes are not empirically supported. The validity evidence with human figure drawings is even more limited. With a few exceptions, projective indexes have not consistently demonstrated incremental validity" (p. 27). Nonetheless, that same review reported that more than 80% of clinical psychologists continue to report using the Rorschach at least occasionally. Projective tests are simply intuitively compelling to many people. Can an effective projective measure of attitudes be created? We propose to do so by integrating the ambiguous interpretations of classic projective tests with the precision and control of priming experiments that have recently flourished in social cognition research.

The logic of the procedure is as follows: People are asked to make evaluative judgments in an ambiguous judgment situation. For each judgment, they are exposed to an attitude object (a prime; say, President George W. Bush) that gives rise to a positive or negative evaluative reaction. They are also presented with a judgment target that is ambiguous in how it should be evaluated (e.g., an abstract symbol). They are instructed to avoid expressing any influence of the prime, and only to evaluate the symbol. However, to the extent that individuals misattribute their reactions from the attitude object to the target, the President will nonetheless systematically bias evaluations of the symbol. This kind of affective transfer has been documented in previous studies (e.g., Murphy & Zajonc, 1993). Working backward, to the extent that President Bush engenders a positive versus negative influence on judgments of the symbol, we may infer a favorable reaction to the President. The task is literally projective, because in misattributing reactions caused by the President to the symbol, participants are projecting their own psychological state onto an ambiguous external source. Although previous research has established that affective reactions can influence judgments of other ambiguous events, we are not aware of any research that has explored the implications for attitude measurement. This article explores these implications by introducing an *affect misattribution procedure* (AMP) for implicitly measuring attitudes.

## Some Key Concepts

The use of affective misattributions for the purpose of implicit attitude measurement depends on a few key ideas that should be

B. Keith Payne, Clara Michelle Cheng, Olesya Govorun, and Brandon D. Stewart, Department of Psychology, The Ohio State University.

This research was supported by National Institute of Mental Health Grant 1 R03 MH070673-01A1 to B. Keith Payne. We thank Richard Petty, Larry Jacoby, and Elizabeth Marsh for thoughtful comments on drafts of this article. We also thank Jessica Lang Nowinski for providing the photo of Ella Rose Nowinski used in Figure 1.

Correspondence concerning this article should be addressed to B. Keith Payne, who is now at the Department of Psychology, University of North Carolina at Chapel Hill, CB# 3270, Davie Hall, Chapel Hill, NC 27599-3270. E-mail: payne@unc.edu

clarified. The first is *misattribution*. We conceptualize misattribution as mistaking an effect of one source for the effect of another. Projection is a special case of misattribution, in which the true source is the self and the effect is attributed to an external entity.

The second key concept is *affect*. By affect we mean a rudimentary pleasant or unpleasant reaction (Frijda, 1999; Russell, 2003). Although affect is usually subjectively experienced, it is the product of underlying processes that may be either conscious or unconscious. Simple affective reactions (what Russell, 2003, calls "unattributed affect") are distinguished from emotions in that they are not (yet) appraised as having a specific source or a particular meaning in a given context. As a result, simple affective reactions are available to be attributed or misattributed to various sources.

Affect is one key component of attitudes. We take a broad view of attitudes, as an evaluation of some entity along positive and/or negative dimensions. We resist describing attitudes along a single continuum because evidence suggests that attitudes can be ambivalent and, hence, include simultaneous positive and negative components (e.g., Cacioppo & Berntson, 1994). Of course, attitudes are not limited to affective reactions. For instance, researchers have distinguished between affect, beliefs, and behaviors as separate components of attitudes (Zanna & Rempel, 1988). The present study focuses on the misattribution of affect, as opposed to other psychological states or reactions. However, we recognize that misattributions might well occur for other processes or psychological states, and if so, those misattributions may provide an opportunity for indirect measurement in those domains.

The affect misattribution procedure may be described as an implicit measure in at least two senses. First, it is implicit in the sense that it is an indirect measure (see Campbell, 1950). Participants are not directly asked to report their attitudes, but instead attitudes are inferred from behavior. Second, the AMP measures influences of attitudes on behavior that persist *in opposition to* participants' intentions. Task requirements demand that participants not express any evaluations of the attitude objects in their performance. Those evaluations that bias performance nonetheless are taken to reflect automatic (i.e., unintentional) influences of attitudes. The relationship of the AMP to issues of automaticity and control and to consciousness and unconsciousness are picked up again in the General Discussion, after the procedure and data have been fully presented.

## From Affect Misattribution to Implicit Attitude Measurement

There is an important relationship between participants' ability (or inability) to identify the sources of their reactions and the implicit character of the attitude measure. Only if the participant is unable to separate his or her reaction to the President from his or her reaction to the symbol will any misattribution occur (Jacoby, Kelley, & Dywan, 1989; Winkielman, Zajonc, & Schwarz, 1997). In other words, misattribution is likely to take place only when participants are unable to monitor and control the influence of their attitudes toward the President on their judgments. Therefore, any misattribution observed is likely to be an implicit or automatic reflection of attitudes, in the sense that participants would be unable to monitor and control the expression of their attitudes via the misattribution.

The potential for measuring attitudes outside of participants' monitoring and control is a major objective of implicit attitude research. By limiting monitoring and control, implicit measures can be extremely useful for studying socially sensitive attitudes such as prejudice (Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998; Wittenbrink, Judd, & Park, 1997). Implicit measures are also useful for predicting stigmatized behaviors such as drug, alcohol, and tobacco use (Palfai & Ostafin, 2003; Sherman, Rose, Koch, Presson, & Chassin, 2003; Wiers, Van Woerden, Smulders, & De Jong, 2002). These implicit measures benefited from the widespread availability of computers, which allow a large number of repeated observations, careful experimental control, and precise timing of stimuli and responses. Precisely because they are so useful in these regards, implicit measurement techniques have enjoyed an explosion of research attention in recent years. In view of the fact that several new implicit techniques have been developed, before describing the affect misattribution procedure the following section considers some characteristics important in evaluating the contributions of new implicit measures. We believe that the AMP offers advantages compared with currently available implicit measures in validity, reliability, sensitivity, and ease of use.

## Important Properties of Implicit Measures

### *Validity*

Every implicit measure in the published literature has shown evidence of validity in one form or another. Although reviewing this evidence is beyond the scope of the present article, readers are referred to Fazio and Olson (2003) and Greenwald and Nosek (2001) for recent reviews. In the present research, we sought to establish the validity of the AMP in three ways. First, we show that it is sensitive to evaluations of items that are near-universally regarded as favorable or unfavorable. Second, we show that the AMP predicts behavioral intentions. Third, we show that the AMP is associated with explicitly measured attitudes when participants are unmotivated to conceal their attitudes. This is done by first selecting a topic that is not socially sensitive (Experiment 5) and then by measuring individual differences in motivations using a topic that is socially sensitive (Experiment 6).

### *Reliability*

According to classical measurement theory, reliability reflects the ratio of "true score" variance to "error" variance. As such, the reliability of a measure sets the upper limit on possible correlations with other variables. Currently, evidence of reliability among implicit measures is mixed. Reliability estimates for the Implicit Association Test (IAT) have ranged from quite high (Hoffman, Gawronski, Gschwendner, Le, & Schmitt, in press) to quite low (Bosson, Swann, & Pennebaker, 2000; Cunningham, Preacher, & Banaji, 2001). Priming procedures often produce relatively low reliabilities (e.g., Kawakami & Dovidio, 2001). Although reliability estimates are varied, values for implicit methods are frequently lower than what is conventionally accepted for explicit measures. Conventional standards hold that explicit measures should have reliability coefficients of .80 or higher. We see no reason why implicit measures should not be evaluated by the same standards.

The present research establishes the reliability of the AMP by computing internal consistency. In each experiment, the AMP displayed good reliability.

### *Sensitivity and Effect Size*

The IAT is the most widely used implicit measure, and a key reason is the instrument's large effect sizes. Large effect sizes make it easy to detect effects, and easy to replicate them. Of course, it is important that the effect size relates to the strength of the attitude. In the presence of a true attitude, a larger effect size can indicate a more sensitive measure. The present studies demonstrate consistently large effect sizes using the AMP.

### *Ease of Use*

A final practical consideration is ease of use. Various psychophysiological techniques have been used as indirect measures of attitudes (e.g., Ito & Cacioppo, 2000; Phelps et al., 2000). However, in many cases these measures are expensive and demand considerable labor, expertise, and technology, making them impractical for many uses. Other techniques such as reaction time priming and the implicit association test are more easily applied, using only the technology of personal computers. However, these measures still require several stages with different instructional sets, sometimes involving cover stories to obscure the purpose of the task. The AMP is also administered using personal computers. However, as we show ahead, the AMP requires only a single phase with one set of instructions, requires no deception, and can be completed in less than 5 min. Considerations of validity, reliability, sensitivity, and ease of use are important when evaluating any measure, implicit or otherwise. In the following, we describe the development of the AMP and how it fares against these criteria.

## Overview

To elicit misattributions, we developed a procedure based on that used by Murphy and Zajonc (1993), albeit with several alterations. Murphy and Zajonc presented positively or negatively valenced prime pictures followed by Chinese pictographs. Participants were asked to rate their liking for the pictographs on 10-point scales. This study found that when the primes were presented too briefly to be consciously identified (4 ms), the valence of the prime influenced ratings of the pictographs (presented for 2,000 ms). However, this prime-consistent effect was eliminated when the primes were presented visibly (1,000 ms). A primary conclusion of this study was that affective reactions to the primes could be misattributed to the pictographs, but only when participants could not identify the real source of their affect.

Unlike Murphy and Zajonc's (1993) procedure, in our procedure we presented the primes visibly rather than subliminally. We made this choice because we were interested in studying the impact of the primes despite correction attempts rather than in trying to avoid correction attempts. To maximize the ambiguity of the judgment being made, we had participants make a dichotomous judgment of each pictograph as "more or less pleasant than the average pictograph" rather than judge it on a continuous rating scale. Most critically, we explicitly warned some participants not to let the prime images influence their evaluations of the pictographs. A great deal of research has shown that when people are aware of a potentially biasing influence, they often adjust their judgment to eliminate or even reverse the bias (e.g., Strack, Schwarz, Bless, Kubler, & Wanke, 1993; for reviews, see Martin, Seta, & Crelia, 1990; Wegener & Petty, 1997; Wilson & Brekke, 1994). We reasoned that if responses reflected true misattributions, they would be difficult to monitor and control. Any effects that persist despite such a warning would provide strong evidence for the unintentional expression of attitudes. Indeed, the misattributions proved remarkably immune to warnings.

The AMP, then, consists of an affect-laden prime followed by an ambiguous target. Participants classify the target as relatively pleasant or unpleasant (see Figure 1). In some conditions, participants are warned not to be influenced by the primes. Influences of the primes on target evaluations are used to assess participants' attitudes toward the prime objects. Six studies are presented to validate the affect misattribution procedure. The first two experiments demonstrate evaluative misattributions from consensually favorable or unfavorable objects. These studies also establish the AMP's resistance to monitoring and control attempts using a warning manipulation. Experiments 3 and 4 examined the timing of primes, targets, and the interval between the two in an effort to better understand how the measure operates. A fifth experiment assessed individual differences in attitudes toward presidential candidates. Because participants are willing and able to publicly express their political attitudes, we used this topic to validate the AMP by comparing participants' priming responses to their self-reported attitudes and voting intentions. In a final, sixth experiment, we applied the AMP to the socially sensitive topic of racial attitudes. We compared participants' priming responses with their self-reported racial attitudes, taking into account individual differences in motivations to be unprejudiced.

## Experiment 1

In our initial study, we expected participants to misattribute their affective reactions, but only when the potential for bias was not salient. We expected that once their attention was called to the potential bias, they would correct their judgments. Therefore, we created two experimental groups, a warning condition and a no-warning condition. Both groups completed the priming task, in which affect-laden pictures were paired with ambiguous pictographs. Unanimously pleasant and unpleasant prime pictures were selected from a normed set of stimuli (Lang, Bradley, & Cuthbert, 1995). Examples of pleasant pictures included a smiling baby and puppies. Unpleasant pictures included such items as a spider and a handgun. The Appendix lists the specific images used.

### *Method*

#### *Participants*

Participants were 33 introductory psychology students who participated in the study for partial course credit (participant sex was not recorded). Data from 1 Chinese-speaking participant were removed from the analysis, leaving a final sample of 32.

#### *Design*

The experiment was a 3 (pleasant vs. unpleasant vs. neutral primes) × 2 (warning vs. no-warning) design with prime manipulated within partic-

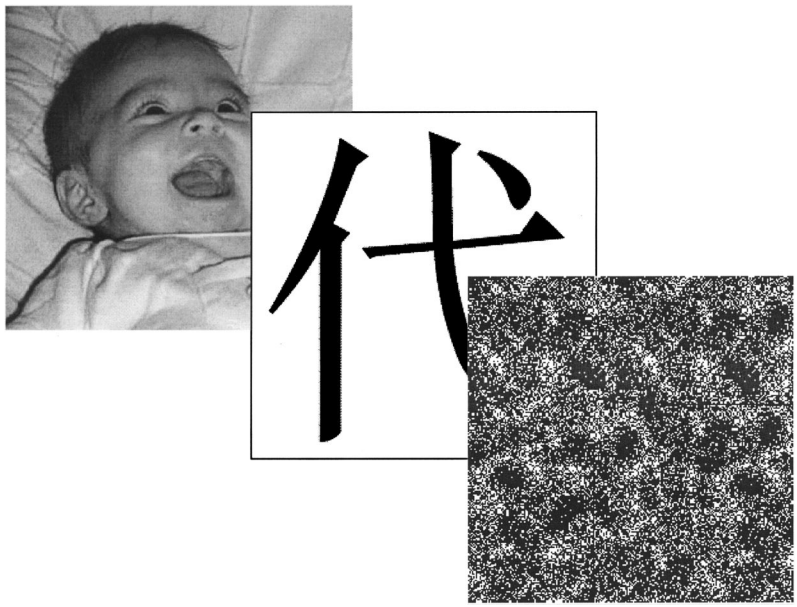

*Figure 1.* Representative stimuli used in the affect misattribution procedure.

ipant and warning between participants. The dependent variable of primary interest was the proportion of pictographs participants judged as pleasant in each prime condition.

### *Procedure*

Participants were seated in front of a computer and were informed that the study examined "how people make simple but quick judgments." Participants were told that they would see pairs of pictures flashed one after the other, the first one being a real-life image and the second being a Chinese character. They were told that the real-life image simply served as a warning signal for the Chinese character and that they should do nothing with the real-life image. Instead, their job was to judge the visual pleasantness of each Chinese pictograph. Participants were instructed to press a key labeled *unpleasant* if they judged the Chinese pictograph to be less visually pleasing than average and a key labeled *pleasant* if they judged it to be more visually pleasing than average. In addition, participants were instructed to respond quickly.

Participants were randomly assigned to the warning or no-warning condition. Participants in the warning condition were additionally told that, "Sometimes, the photographs presented prior to the Chinese characters can bias your responses on those characters. Thus, please try to make sure that your responses are not influenced by the photographs."

From Lang et al.'s (1995) normed ratings, 12 positive images (rating $M = 7.88$, $SD = 1.40$) and 12 negative images (rating $M = 3.33$, $SD = 1.63$) were selected for primes. The positive and negative images were matched on arousal ratings ($M$s = 5.04 and 4.97, $SD$s = 2.39 and 2.25, respectively) to ensure that any differences across prime conditions were driven by valence rather than arousal. An image of a gray square was created as a neutral prime.

During each trial of the priming task, the prime image appeared in the center of the screen for 75 ms, followed by a blank screen for 125 ms, and then a Chinese pictograph for 100 ms. Following the pictograph, a pattern mask consisting of black and white "noise" appeared until the participant responded (see Figure 1). The next trial began as soon as participants made a response. Participants completed a total of 36 randomly ordered trials, with 12 each of positive, negative, and neutral primes. Thirty-six different Chinese characters were used as targets. Each pictograph was paired with a prime picture in a new random order generated by the computer program for each participant. The task lasted approximately 4 min. Following the priming task participants were debriefed.

### *Results*

Results were analyzed using a 3 (prime) × 2 (warning) analysis of variance (ANOVA). As displayed in Figure 2, there was a clear main effect of prime valence, $F(2, 60) = 10.18$, $p < .01$. Participants were most likely to judge the pictograph as pleasant following a pleasant prime and were least likely to do so following an unpleasant prime. Responses following neutral primes were intermediate between these two. Post hoc comparisons showed that all three prime conditions were significantly different from each other (all $F$s > 6.5, $p$s < .05). This effect was not qualified by the warning condition, $F(2, 60) = 0.40$, $p = .68$. Moreover, the priming effect was significant for both the no-warning condition,

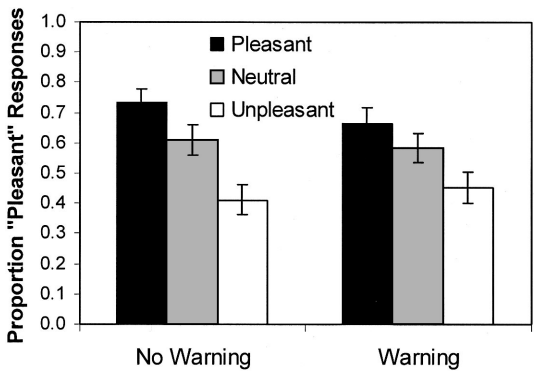


*Figure 2.* Proportion of "pleasant" responses as a function of prime pleasantness and warning condition, Experiment 1. Error bars represent one standard error.

$F(2, 30) = 5.62$, $p < .01$, and the warning condition individually, $F(2, 30) = 4.67$, $p < .02$.

We calculated Cohen's effect size to examine the size of the priming effect. We computed this statistic by subtracting the proportion of pleasant responses on pleasant-prime trials from the proportion of pleasant responses on unpleasant-prime trials and dividing that difference by the pooled standard deviation. By convention, an effect size of 0.80 or greater is considered large. The effect size in the present experiment was 1.13.

To examine the reliability of the priming measure, we used Cronbach's alpha. A set of 12 difference scores was created and treated as individual "items" (for a conceptually similar approach using reaction time variables, see Bosson et al., 2000). This procedure is akin to taking two randomly selected halves of trials and conducting a split-half analysis. One could go further by splitting the procedure into 3rds, 4ths, and so on. We split the test into 12ths to obtain the most fine-grained analysis afforded by the procedure. First, each trial was scored as +1 for a pleasant judgment or 0 for an unpleasant judgment. The score on each randomly selected unpleasant-prime trial was subtracted from a randomly selected pleasant-prime trial. Each trial was used in only one pair. This created a set of 12 difference scores that could each range between +1, 0, and −1. For a given pair of trials, if a participant responded in a prime-consistent way by judging "pleasant" on the pleasant-prime trial and "unpleasant" on the unpleasant-prime trial, that participant would receive 1 for that difference score. If the participant responded either "pleasant" or "unpleasant" on both trials, he or she would score 0. Finally, if the participant responded in a prime-inconsistent way, he or she would receive −1. This analysis revealed a high level of reliability ($\alpha = .85$).

### *Discussion*

Results showed that the affective valence of the prime pictures influenced participants' evaluations of the pictographs. The priming effect proved to be large and internally consistent from trial to trial. Surprisingly, this effect persisted with a blatant warning against being biased by the primes.

Theories of judgmental correction indicate that in order to correct judgments, people must (a) have an appropriate lay theory of how they might be biased, (b) be motivated to correct for its effects, and (c) have the opportunity or capacity to adjust their response (Martin et al., 1990; Wegener & Petty, 1997; Wilson & Brekke, 1994). One interpretation of the warning's ineffectiveness is that participants could not monitor the impact of their prime-based reactions on their target judgments. That is, people could not tell when they were being influenced by the prime and when they were not.

However, correction theories suggest three other factors that might also explain the ineffectiveness of the warning. One possibility is that the warning was not salient or blatant enough to make participants aware of the bias. Participants may have overlooked the warning, or they may have had incorrect lay theories about how they might be biased. A second possibility is that despite being aware of the potential for bias, participants were not motivated enough to change their behavior. A third possibility is that participants did not have the opportunity to correct their judgments because the task instructions encouraged fast responding. To rule out these alternatives (1 = *lack of awareness* and 3 = *lack of opportunity*), we replicated Experiment 1 without instructing participants to respond quickly and with a more blatant warning. Not only was the warning more salient, but it also spelled out exactly how the primes might bias judgments, in case any participants did not have the correct lay theory. The question of participants' motivation and effort to correct is addressed in Experiments 5 and 6.

## Experiment 2

### *Method*

Forty-three introductory psychology students participated in the experiment in exchange for partial course credit. Experiment 2 was identical to Experiment 1, with the exception that participants in the warning condition were given a more explicit correction instruction:

> It is important to note that having just seen a positive image can sometimes make you judge the drawing more positively than you otherwise would. Likewise, having just seen a negative image can make you judge the drawing more negatively. Because we are interested in studying how people can avoid being biased, **please try your absolute best not to let the real-life images bias your judgment of the drawings!** Give us an honest assessment of the drawings, regardless of the images that precede them.

The warning appeared in large 20-point font, and the emphasized section was in bold type. Participants were not asked to respond quickly. The salient and blatant warning was expected to ensure that all participants were aware of the potential for bias and were motivated to avoid bias and that they had the correct theory about how they might be biased. Because responses were self-paced, we expected participants to have ample capacity to adjust their judgments if they were so inclined. All other aspects of the experiment were the same as in the first experiment.

### *Results*

As illustrated in Figure 3, results were virtually identical to those of Experiment 1. Participants were most likely to judge pictographs as pleasant after a pleasant prime, followed by the neutral prime, and least likely after an unpleasant prime. The main effect of prime was significant, $F(2, 82) = 44.61$, $p < .01$. Post hoc comparisons showed that all three prime conditions were significantly different ($F$s $> 14.0$, $p$s $< .001$). This effect was not

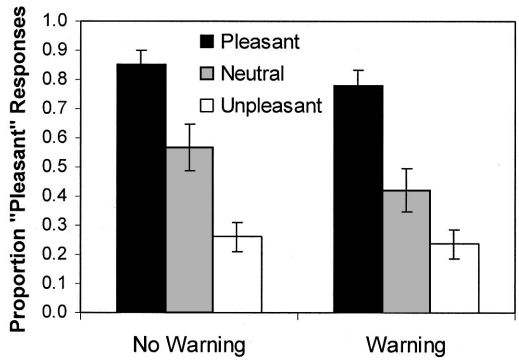


*Figure 3.* Proportion of "pleasant" responses as a function of prime pleasantness and warning condition, Experiment 2. Error bars represent one standard error.

qualified by the warning condition, as evidenced by the lack of a Prime × Warning interaction, $F(2, 82) = 0.58$, $p = .56$. The priming effect was significant in both the no-warning condition, $F(2, 42) = 25.05$, $p < .01$, and the warning condition, $F(2, 40) = 20.30$, $p < .01$. The effect size in this experiment was larger than in Experiment 1 (Cohen's $d = 2.44$). Finally, internal consistency across trials was approximately equal to that in Experiment 1 (Cronbach's $\alpha = .81$).

### Discussion

The results of this experiment closely replicated the results of Experiment 1 despite a salient and blatant warning. We believe this to be strong evidence that performance on this task is an implicit reflection of participants' evaluations of the primes. We suspect that the warning's ineffectiveness was because participants could not accurately monitor when they were incorporating their affective reactions to the primes into their judgments of the targets.

One potential reason for this poor monitoring and correction may be that the prime and target events were presented very rapidly. The timing of stimulus events can have powerful effects on how much control participants exert over their responses in priming tasks. In particular, there is reason to expect three aspects of the stimulus timing to have important effects. First is the length of time the prime is presented. Participants might be better able to correctly attribute their reactions to the prime when it is presented for longer intervals. By correctly identifying the source of affect, they might avoid misattributing the affect to the targets (Murphy & Zajonc, 1993). Second is the interval between prime and target (often labeled the stimulus–onset–asynchrony; SOA). On the basis of the results of dozens of priming studies, one might expect the priming effect to be reduced at longer SOAs (e.g., Neely, 1977). This is because with a longer interval participants have more time to implement a careful, deliberate strategy after the prime is presented but before their response to the target. Moreover, the reaction to the prime might also dissipate quickly, reducing the priming effect by passive decay with time. A third aspect of timing is the length of time that the target is presented. Like the duration of the prime, longer target duration could increase the ability to sort out one's reactions to the prime from reactions to the target. Also, a longer presentation might allow more time to implement a corrective strategy.

In Experiment 3 we tested the effect of prime duration and SOA separately. In the fourth experiment we tested the effect of target duration. These studies were aimed at shedding some light on how and why the AMP "works." Are the effects we have observed thus far all because of the specific (relatively fast) timing parameters we chose? Or are the effects relatively indifferent to timing? Both of the following studies showed that the AMP priming effect was remarkably robust even at relatively long time intervals within the ranges we tested.

## Experiment 3

In this experiment we manipulated both prime duration and SOA by including three SOAs (100, 500, and 1,500 ms). Within each SOA, the prime was presented for either 75 ms or the full time interval. All participants received a blatant warning to avoid the influence of the primes. On the basis of previous research, we expected the priming effect to be reduced when the primes were presented longer and when the interval between prime and target was longer. If these expectations were correct, they might explain, in part, why participants seemed to have such difficulty in avoiding affect misattributions in our procedure.

### Method

#### Participants

Forty-seven undergraduates participated for course credit. Initial analyses revealed that 2 participants failed to follow instructions, responding to every target with the same response (one all pleasant, the other all unpleasant). These 2 participants were removed from the reported analyses, leaving 45 participants (26 women and 19 men).

#### Design and Procedure

This experiment used the same affective pictures as the previous two studies. All participants were given the strong warning used in Experiment 2. The design was 3 (prime valence: pleasant, neutral, unpleasant) × 3 (SOA: 100, 500, 1,500 ms) × 2 (interval: filled vs. unfilled), with prime valence and SOA manipulated within participant and filled versus unfilled intervals manipulated between participants. The SOA varied from trial to trial in a random order. For the filled-interval group, the prime always remained on the screen for the full SOA. For the unfilled-interval group, the prime appeared for 75 ms. The remaining time between the offset of the prime and the onset of the target was filled with blank screen. Following the target presentation of 100 ms, a pattern mask appeared until participants made a response. The task lasted approximately 4 min. All other aspects of the priming task were the same as those reported in Experiment 2.

### Results

Results were analyzed using a 3 (prime valence) × 3 (SOA) × 2 (interval) ANOVA. Figure 4 displays the results. Two effects emerged. First, prime valence had a significant effect, replicating previous studies. Participants were more likely to evaluate the pictographs as pleasant after a pleasant prime, followed by a neutral prime, and they were least likely to evaluate them as pleasant after an unpleasant prime, $F(2, 86) = 30.18$, $p < .01$, Cohen's $d = 1.47$. This prime effect was qualified by the predicted Prime Valence × SOA interaction, $F(4, 172) = 2.99$, $p < .05$. Although the priming effect was highly significant at each level of SOA (all $p$s $< .001$), there was a significant reduction as the SOA became longer.

Contrary to our prediction, the prime duration did not have any significant effects (all $p$s $> .20$). Although the priming effect in the filled-interval condition was numerically smaller than that in the unfilled condition, they were not statistically different. Even when the prime remained on the screen for 1,500 ms., the misattribution effect was not substantially different than when the prime was presented for only 75 ms. Further, post hoc tests showed that the difference between neutral primes and pleasant primes was significant in the unfilled-interval condition, $F(1, 19) = 9.62$, $p < .01$, but not the filled-interval condition, $F(1, 24) = 0.85$, $p = .37$, although the two-way Prime Valence × Interval interaction was not significant, $F(2, 86) = 1.59$, $p = .21$. No other interactions were significant. This difference between neutral and pleasant primes was not predicted, and we resist drawing conclusions based on this effect, given that the overall interaction was not reliable.

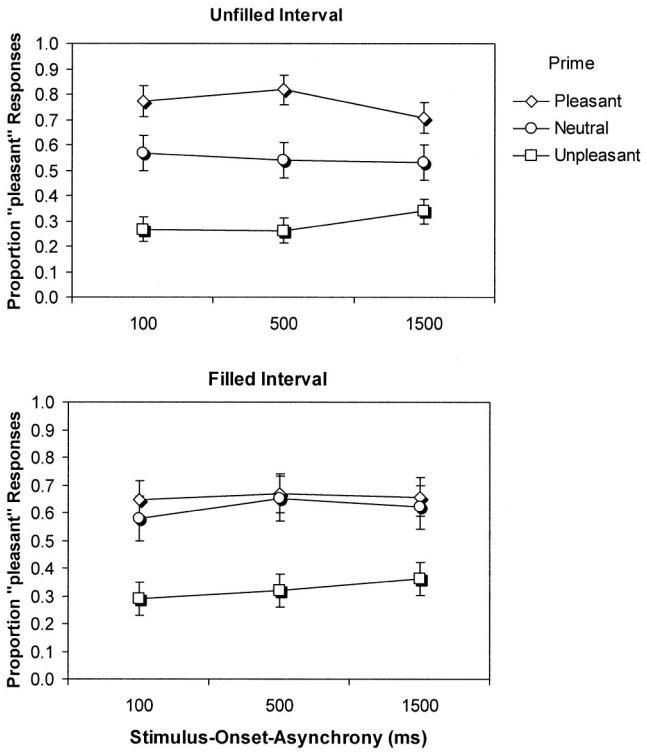


*Figure 4.* Proportion of "pleasant" responses by prime pleasantness, stimulus-onset asynchrony (SOA) and whether the SOA interval was filled with the prime image, Experiment 3. Error bars represent one standard error.

Reliability was calculated as described in previous studies, yielding an overall alpha of .95. When reliability was computed for each SOA condition separately, coefficients were .86, .88, and .87 for the short, medium, and long SOAs, respectively.

### *Discussion*

Results showed that the time interval between the onset of the prime and target diminished the priming effect by a small amount. Although the reduction was statistically reliable, the priming effect was remarkably strong even when the SOA was long and that interval was completely filled with the prime image. By the standards of sequential priming research, these time figures range from values considered to severely limit controlled processing to values considered to allow ample controlled processing. It is interesting that although affective priming paradigms that rely on response latencies tend to show affective priming only at very short SOAs (Hermans, De Houwer, & Eelen, 2001), the misattribution procedure was not limited to these short time intervals. On one hand, the fact that the priming effect was significantly reduced at long time intervals suggests that participants may gain some ability to correct for the influence of the primes at slow time scales, or this might reflect a passive decay with time. On the other hand, the fact that the priming effect persisted from very short time intervals to very long time intervals suggests that a large component of the priming effect is independent of timing, at least within this range.

Given that prime duration and SOA had such small effects, the next study examined whether the length of the target presentation was critical for the misattribution. A very fast presentation of the target might give participants little basis for judging the pictographs. Without time to inspect the pictograph features, they may have to rely on other cues, including their own affective states. Alternatively, allowing ample time to examine the features of the pictograph may allow people to avoid the effects of the prime. To test this possibility, we ran a fourth experiment varying the duration that the target was presented.

## Experiment 4

### *Method*

#### *Participants*

Forty-six introductory psychology students participated in the experiment for partial course credit. Six participants who spoke Chinese were removed from the analyses. Two additional participants were dropped for failing to follow instructions by giving the same response to every item. This left a final sample of 38.

#### *Design*

The study had a 2 (prime: pleasant vs. unpleasant) × 3 (target duration: 100 ms vs. 750 ms vs. 2,000 ms) factorial design. We eliminated the neutral prime condition because it was not central for any of our questions of interest. Both prime valence and target duration factors were manipulated within participants, with each block containing a single target duration. The blocks were counterbalanced for order, with participants randomly assigned to one of the six orders. We expanded the number of primes by additionally selecting 12 positive (pleasantness rating $M = 7.69$, $SD = 0.38$) and 12 negative images (rating $M = 2.72$, $SD = 0.60$) from the normed set used in the previous studies (Lang et al., 1995). As before, the newly selected positive images ($M = 1.48$, $SD = 0.26$) did not differ from the negative images ($M = 1.62$, $SD = 0.23$) on the ratings of arousal. We also expanded the number of target pictographs to 96 so that a unique pictograph could be used on each trial.

#### *Procedure*

As in the previous experiments, participants were told that they would see and evaluate Chinese characters preceded by pleasant and unpleasant images. Participants completed three blocks of trials, in which target duration was manipulated in a counterbalanced order. Each block contained 32 randomly ordered trials, half of which included positive primes and the other half negative primes. All participants were given notice that the time the character appeared on the screen would change with each set of trials. At the beginning of each block, participants were informed and/or reminded that the images may bias their evaluations of the characters and instructed to "to try their absolute best" to avoid such bias. The task required less than 5 min.

### *Results*

A 2 (prime: pleasant vs. unpleasant) × 3 (duration: 100 vs. 750 vs. 2,000 ms) repeated measures ANOVA was conducted on the proportion of characters participants judged as pleasant. This analysis revealed a significant main effect of the prime, such that participants evaluated more characters as pleasant after positive primes than after negative primes, $F(1, 37) = 40.84$, $p < .01$, Cohen's $d = 1.38$. The main effect of the prime was qualified by

a Prime × Target Duration interaction, $F(2, 74) = 5.01$, $p < .01$ (see Figure 5). The priming effect was significant for each duration individually (all $F$s $> 17.03$, $p$s $< .001$). Cohen's effect sizes associated with the priming effect in the 100 ms, 750 ms, and 2,000 ms blocks were 1.71, 1.42, and 1.00, respectively. Individual comparisons indicated that the 100-ms and 750-ms blocks did not differ from each other in the magnitude of the priming effect, $F(1, 37) = 0.44$, *ns*. However, the priming effect in the 2,000 ms block was significantly smaller than the priming effect in both of the remaining blocks (both $p$s $< .04$).

Reliability analysis showed that across all trials, reliability was high ($\alpha = .90$). When reliability was estimated separately for each duration, coefficients were .80, .83, and .73 for the 100, 750, and 2,000 ms durations, respectively.

### *Discussion*

This study demonstrated a significant decrease in priming with long target presentations. Nonetheless, the priming effect remained strong and significant at longer presentations. This suggests that the misattribution of affect from the prime to the target can happen very quickly, within the first 100 ms that the prime is presented. Yet the effect appears to last at least 2,000 ms. The sparse and ambiguous qualities of the target pictographs may be key here, because even with plenty of time, there is little that participants can glean from these items as a basis for evaluation.

Experiments 1–4 demonstrated a replicable pattern of evaluative priming with a large effect size, high reliability, and resistance to warnings. However, because the primes were consensually positive or negative, it remains to be seen how well the AMP can assess individual differences in attitudes or predict behavior. To this end, we selected a topic with considerable social and political importance: attitudes toward the candidates in the 2004 United States presidential election. We tested whether the AMP could be used to predict self-reported attitudes and intentions to vote for one candidate over the other. Because we were interested only in unintentionally expressed attitudes from the AMP, we gave a blatant warning to all participants in Experiment 5.

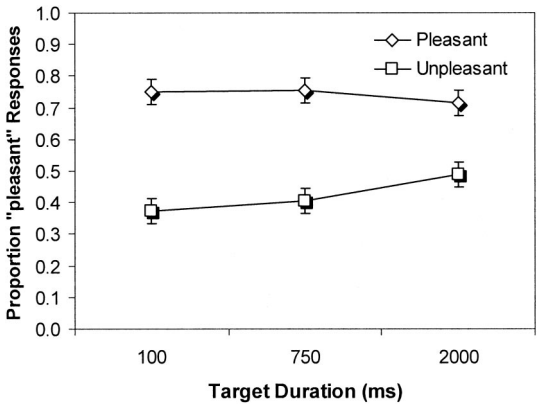


*Figure 5.* Proportion of "pleasant" responses as a function of prime pleasantness and target duration, Experiment 4. Error bars represent one standard error.

## Experiment 5

When applied to political attitudes, we expected the AMP to correlate with participants' explicit attitudes and their intentions to vote for one candidate over the other. This study was conducted several weeks before the 2004 United States presidential election. To measure attitudes toward presidential candidates George Bush and John Kerry, we selected a set of photographs of the two candidates from Internet media sources. We first collected 25 photographs of each candidate. From that set we selected pictures so that the two candidates were matched on several criteria: whether the candidate was (a) smiling, (b) speaking, (c) gesturing, or (d) facing toward or away from the camera and (e) whether the photograph included an American flag. We selected 12 photos of each candidate that were matched on these features, so that on average each feature was present equally in both sets. All of the photographs included the candidate's head and shoulders and, in some cases, upper torso. In all photographs, the candidate was dressed in a suit or other business attire.

### *Method*

#### *Participants*

Forty undergraduate students (24 women, 16 men) participated in return for partial course credit. The data from 1 participant who spoke Chinese were removed.

#### *Design*

The design of this experiment was a 3 (prime: Bush, neutral, Kerry) × 2 (vote preference: Bush, Kerry) factorial. Participants were given the same instructions as in Experiment 1, with the following exceptions. Participants were informed that, "The real-life photographs will be of familiar people. They are pictures of President George W. Bush and Democratic primary candidate John Kerry." All participants were warned against being influenced by the prime photos. The text of the warning read:

> It is important to note that the real-life image can sometimes bias people's judgments of the drawings. Because we are interested in how people can avoid being biased, **please try your absolute best not to let the real-life images bias your judgment of the drawings!** Give us an honest assessment of the drawings, regardless of the images that precede them.

A second change to the AMP was that the number of trials was increased to maximize reliability. Seventy-two unique pictographs were paired randomly with each of the 36 prime pictures (12 Bush, 12 Kerry, 12 gray squares) twice, for a total of 72 trials. This task lasted approximately 5 min.

Following the AMP, we assessed participants' explicit attitudes using seven self-report questions regarding each candidate. These items were (a) "Are your FEELINGS toward [George Bush/John Kerry] generally warm and favorable or cold and unfavorable?" (b) "Overall, would you say you generally LIKE or DISLIKE [George Bush/John Kerry]?" (c) "To what extent do you think [George Bush/John Kerry] is a COMPETENT leader?" (d) "To what extent do you think [George Bush/John Kerry] is a MORAL leader?" (e) "To what extent do you SHARE [George Bush/John Kerry]'s positions on political issues?" (f) "To what extent do you SUPPORT [George Bush/John Kerry]'s policies?" (g) "To what extent do you SUPPORT [George Bush/John Kerry] in the upcoming presidential election?" Each item was answered on a 5-point scale anchored with *not at all* and *extremely*.

Next, participants' behavioral intentions were assessed by asking them to indicate which candidate they would vote for if the election were held

today. Finally, participants were asked some questions about the AMP, including a question about their attempts to correct their judgments for the influence of the primes. This item read, "How much did you attempt to CORRECT your evaluations for the influence of the candidate photos?" Answers were rated on a 5-point scale. Following completing the questionnaire measures participants were debriefed.

### *Results*

Results were analyzed using a 3 (prime) × 2 (voting preference) ANOVA. The sample was evenly divided in their preferences, with 19 Kerry voters and 20 Bush voters. Results are displayed in Figure 6. As is clear from this figure, the only significant effect was the predicted Prime × Voting Preference interaction, $F(2, 74) = 11.13$, $p < .01$. Kerry voters judged the pictographs more pleasant when primed with Kerry than Bush. In contrast, Bush voters judged the pictographs more pleasant when primed with Bush than Kerry. Simple effects tests showed that the main effect of prime was significant for both Kerry voters, $F(2, 36) = 4.07$, $p < .05$, and for Bush voters, $F(2, 38) = 9.42$, $p < .01$, in opposite directions as predicted.

Further follow-up analyses showed that responses on Bush-prime trials were significantly more positive for Bush voters than Kerry voters, $F(1, 37) = 5.38$, $p < .05$. Also, responses on Kerry-prime trials were significantly more positive for Kerry voters than Bush voters, $F(1, 37) = 11.06$, $p < .01$. Responses on neutral-prime trials did not differ significantly between voter groups, $F(1, 37) = 1.02$, $p = .32$.

The size of the attitude-consistent priming effect can be estimated by the effect size of the Prime × Vote interaction. Using Cohen's effect size, we found this effect size was 0.81, large by conventional standards. A second way to gauge the size of the effect is to compute the point-biserial correlation between the priming effect and voting preference. A priming score was created for each participant by subtracting the proportion of "pleasant" judgments on Kerry prime trials from the proportion on Bush prime trials. This difference score represents the degree to which participants were more positive toward Bush than Kerry. The correlation between priming and voting intention was .58 ($p < .01$). Finally, we computed the reliability of the AMP using the same procedure as in the previous studies. The measure again showed high internal consistency (Cronbach's $\alpha = .90$).

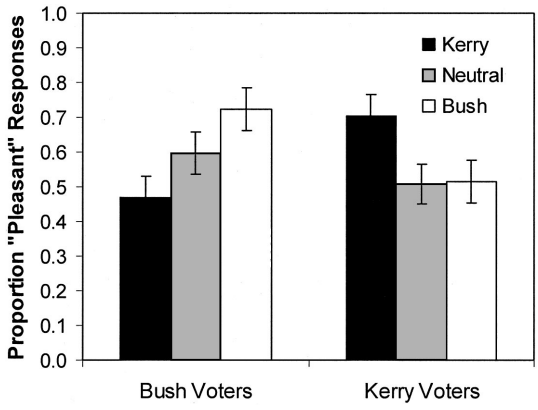


*Figure 6.* Proportion of "pleasant" responses as a function of prime and voting intention, Experiment 5. Error bars represent one standard error.

We next report the relationship between the priming score and participants' self-reported attitudes toward the candidates. We created a comparative attitude index by taking the difference between responses to each Bush item and each corresponding Kerry item, with higher values representing more positive attitudes toward Bush. Table 1 displays the correlations between the AMP score and each question. As can be seen, the relationship was strong and positive for each item. When the seven items were averaged into a single scale ($\alpha = .98$), the size of the relationship between implicit and explicit attitude estimates was $r = .65$ ($p < .01$).

In the preceding analyses, we reported comparative measures using the difference between evaluations of Bush and Kerry. However, in many cases it is useful to estimate attitudes without comparison to a contrast category. We investigated whether the AMP could be used for absolute rather than relative estimates in the following analyses. The tendencies to respond "pleasant" on Bush trials and on Kerry trials were surprisingly independent ($r = -.02$). However, tendencies to respond "pleasant" on neutral trials were correlated with responses on both Bush trials ($r = .40$, $p < .01$) and Kerry trials ($r = .30$, $p = .07$). Therefore, in the following analyses we report partial correlations, holding constant performance on neutral trials to correct for general tendencies to respond "pleasant." For the self-reported attitude items, Bush items were averaged together into an Attitudes Toward Bush Scale and Kerry items were averaged into an Attitudes Toward Kerry Scale.

Controlling for performance on neutral prime trials, we found the proportion of pleasant responses on Bush trials was positively related to attitudes toward Bush ($pr = .42$, $p < .01$) and negatively related to attitudes toward Kerry ($pr = -.38$, $p < .05$). Similarly, the proportion of pleasant responses on Kerry trials was positively correlated with attitudes toward Kerry ($pr = .53$, $p < .01$) and negatively related to attitudes toward Bush ($pr = -.55$, $p < .01$). These results suggest that AMP performance can be meaningfully separated into independent estimates of attitudes toward multiple targets and need not be confined to relative comparisons.

As a final analysis, AMP performance was compared with self-reported effort at correction. The correction question asked participants to rate the degree to which they tried to correct their judgments for the influence of the prime pictures. AMP performance was converted to an absolute value such that higher values represent a larger priming effect in either a pro-Bush or pro-Kerry direction. This analysis showed that attempts to correct judgments were uncorrelated with actual bias in performance ($r = .24$, $p = .16$). The mean rating of correction attempts was 2.60 ($SD = 1.09$), near the midpoint of the 5-point scale. Participants used the entire range of the scale, but that variance was not related to actual performance. It is interesting that the positive direction of this relationship suggests that, if anything, participants who tried more to correct their judgments showed a stronger attitude-consistent bias.

### *Discussion*

The results of this study suggest that the AMP provided a reliable and valid individual difference measure of political attitudes. The measure predicted behavioral intentions to vote. It also correlated strongly with explicit attitude measures. Critically, these

Table 1
*Correlations Between Affect Misattribution Procedure Scores and Self-Reported Attitudes Toward Candidates*

| Item | *r* |
|---|---|
| Feeling thermometer | .60* |
| Liking for candidate | .57* |
| Candidate's competence | .59* |
| Candidate's morality | .67* |
| Share candidate's values | .61* |
| Support candidate's policies | .63* |
| Support as a candidate | .62* |
| Overall *M* | .65* |

* $p < .01$.

results were found even with all participants strongly warned against allowing the primes to influence their judgments.

We chose the topic of political attitudes to validate the AMP because participants are not motivated to conceal their political attitudes. Consequently, it was possible to observe strong correlations between implicit and explicit measures of attitudes. However, a principal strength of implicit measures is their potential for measuring attitudes in more socially sensitive domains. To the extent that implicit measures can circumvent participants' monitoring and control, they can reveal evaluative tendencies that participants may be unwilling or unable to report (cf. Greenwald & Banaji, 1995). Our next step was to test whether the AMP could detect attitude biases in the much more socially sensitive domain of racial attitudes.

## Experiment 6

To measure racial attitudes with the AMP, we replaced the pictures of George Bush and John Kerry with photographs of 12 young Black men and 12 young White men. We recruited a sample of Black and White participants to complete the measure. They also completed an explicit measure of racial attitudes. Previous research has shown that the relationship between implicit and explicit racial attitude measures depends on participants' motivations. Among participants who are highly motivated to avoid acting prejudiced, there is usually little or no correspondence. However, among participants who are not particularly motivated to be unprejudiced, correspondence is more often observed (Banse & Gawronski, 2003; Dunton & Fazio, 1997; Fazio et al., 1995; Payne, 2001). This contingency is usually interpreted to mean that highly motivated people conceal their attitudes on explicit measures, whereas less motivated people express their attitudes more freely. On the basis of these previous findings, we included two measures of motivation to avoid acting with prejudice.

On the basis of previous research and the preceding studies with the AMP, we tested two main predictions. The first was that White participants would judge pictographs as more pleasant when primed with Whites than with Blacks. Because conflicting findings have been reported concerning the performance of Black participants on various implicit measures of race bias, we had no specific predictions for how they would perform. The second hypothesis was that AMP performance and explicit attitudes would be associated, but only for individuals low in the motivation to control prejudice. Critically, we did not expect highly motivated participants to show less bias on the AMP. Instead, we expected this pattern to be driven by highly motivated participants expressing more positive sentiments on self-report measures, but not the AMP.

### Method

#### Participants

Fifty-five participants (36 White and 19 Black) took part in return for course credit. Thirty-four were women and 21 were men. Initial analyses showed no sex differences, and so the analyses reported collapsed over this factor.

#### Design

Photographs of White and Black young men replaced the political candidates of the previous experiment. These photographs were matched on attractiveness using ratings from a pilot study. Only the face was shown in each photo, and each model had a neutral expression. We manipulated the warning between subjects because we expected participants to be sensitive to the potential for race bias and wanted to see whether they were responsive to a blatant warning in the domain of race. Participants were randomly assigned to the warning or no-warning conditions. In the warning condition, the warning read:

> It is important to note that the real-life image can sometimes bias people's judgments of the drawings. Specifically, people may judge a drawing as unpleasant after a picture of a Black person, and judge a drawing as pleasant after a picture of a White person. Because we are interested in studying how people can avoid being biased, **please try your absolute best not to let the real-life images bias your judgment of the drawings!** Give us an honest assessment of the drawings, regardless of the images that precede them.

We chose to word the warning against anti-Black bias because on the basis of previous implicit bias research, this was the pattern we expected from most participants. Previous research is mixed on the extent to which Black Americans can be expected to show pro- or anti-Black implicit bias. In any case, we reasoned that Black participants would not be surprised to read a general warning that anti-Black bias could occur. Therefore, we chose this wording as a plausible warning for all participants. The design of this experiment was 3 (prime race: Black, White, neutral) × 2 (warning, no-warning) design. In addition to the AMP, participants completed an explicit measure of race attitudes and two motivational questionnaires, described below.

#### Questionnaires

Participants' explicit racial attitudes were measured using a "feeling thermometer" type of rating scale. The questions asked participants to rate their feelings toward four different groups: Blacks, Whites, Asians, and Hispanics. Ratings were made on an 11-point scale anchored by the labels 0 (*cold and unfavorable*) and 10 (*warm and favorable*). Feelings toward Blacks and Whites were of primary interest.

Participants' racial motivations were measured using the Internal and External Motivation to Respond Without Prejudice Scale (Plant & Devine, 1998) and the Motivation to Control Prejudiced Responses Scale (Dunton & Fazio, 1997). Both of these scales include two subscales—one that captures a feeling of external social pressure and one that captures a more internalized motivation to be unprejudiced. For Plant and Devine's measure, these subscales are labeled the External Motivation Scale (External) and Internal Motivation Scale (Internal), respectively. For the Dunton and Fazio measures, these are labeled Restraint to Avoid Dispute (Restraint) and Concern With Acting Prejudiced (Concern) Scales. Both measures

have demonstrated evidence of reliability and validity (Dunton & Fazio, 1997; Plant & Devine, 1998).

### Procedure

Participants were first asked to complete the AMP under the same instructions as described in Experiment 5, with the race-specific warning manipulation described above. This procedure took 4–5 min. Following the AMP, participants were asked to take part in an ostensibly separate and unrelated study. During this second phase they completed the questionnaire measures and demographic information about themselves. Following this participants were debriefed.

## Results

### Mean Performance

We first report mean results for the AMP and explicit racial attitudes. Individual difference correlations are reported following these analyses. AMP performance was analyzed using a 3 (prime) × 2 (warning) × 2 (participant race) ANOVA. Figure 7 displays the priming results. This analysis revealed a significant main effect of prime race, $F(2, 102) = 4.34$, $p = .02$. However, this main effect was driven by the fact that participants evaluated the pictographs more pleasantly after a neutral prime compared with a face prime of either race—the main effect was significant only as a quadratic trend, $F(1, 51) = 6.83$, $p = .02$, but not as a linear trend, $F(1, 51) = 0.14$, $p = .71$.

Of more theoretical interest, there was a marginally significant Prime Race × Participant Race interaction, $F(2, 102) = 2.94$, $p = .058$. Whereas White participants evaluated the pictographs more favorable after a White prime than a Black prime, Black participants showed the opposite pattern. Interestingly, there was no effect of the warning manipulation. The main effect, $F(1, 51) = 0.00$, $p = .99$, the Prime × Warning interaction, $F(2, 102) = 1.43$, $p = .24$, and the Prime × Warning × Participant Race interaction, $F(2, 102) = .65$, $p = .52$, did not show any significant effects.

When the neutral primes were removed from the analysis because they were of less theoretical interest, the Prime Race × Participant Race interaction became significant, $F(1, 51) = 5.63$, $p = .02$. Simple effects tests showed that White participants made significantly more "pleasant" responses after a White prime than a Black prime, $F(1, 34) = 5.59$, $p = .02$. The difference between primes for Black participants was not significant, $F(1, 17) = 1.43$, $p = .25$. This is likely because the lower number of Black participants provides less statistical power.

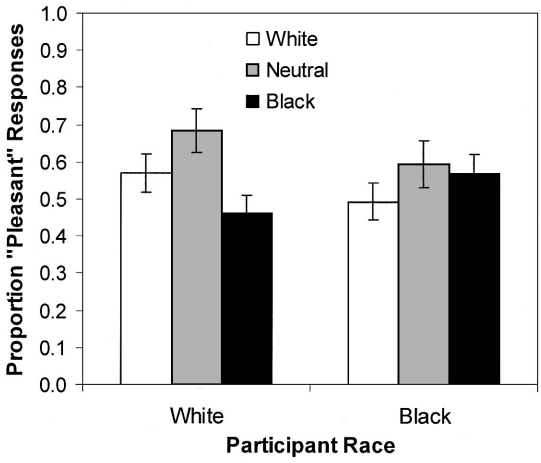


*Figure 7.* Proportion of "pleasant" responses as a function of prime race and participants' race, Experiment 6. Error bars represent one standard error.

Because White and Black participants performed differently on the task, effect sizes were calculated separately for the two groups. White participants showed a medium effect size (Cohen's $d = 0.50$). Black participants showed a smaller effect, although still within the range conventionally considered "medium" (Cohen's $d = 0.28$). Reliability analysis showed a high internal consistency across trials (Cronbach's $\alpha = .85$).

Finally, participants' degree of actual bias was compared with their self-reported efforts to correct their judgments ($M = 2.5$, $SD = 1.05$). AMP performance was computed as an absolute value, such that higher values reflect greater bias in either a pro-White or a pro-Black direction. This analysis showed that correction attempts were unrelated to actual degree of bias ($r = .13$, *ns*). However, correction attempts were significantly correlated with the Concern subscale of the Motivation to Respond Without Prejudice Scale ($r = .29$, $p < .05$), as well as weakly correlated with greater motivation on the Restraint subscale ($r = .17$, $p = .20$), the External Motivation scale ($r = .22$, $p = .11$), and more positive explicitly reported attitudes toward Blacks ($r = .18$, $p = .19$). This pattern of correlations suggests that individuals more motivated to avoid prejudice had a slight tendency to invest more effort in correction attempts during the AMP, although that effort did not translate into lower actual bias on the AMP.

In sum, participants showed a pattern of in-group bias. Whites showed a significant preference for White primes, whereas Blacks showed a nonsignificant preference for Black primes. Providing a blatant warning did not qualify any of these effects, and they were independent of self-reported efforts to correct judgments. The following analyses report the mean effects on explicitly reported attitudes.

Participants' explicit attitude reports were analyzed using a 2 (target group: White, Black) × 2 (participant race: White, Black) ANOVA. This analysis showed a pattern of in-group favoritism that mirrored the priming results. White participants reported more favorable feelings for Whites than for Blacks, ($M$s = 8.92 and 7.44, respectively), $F(1, 35) = 13.87$, $p < .01$. In contrast, Black participants reported more favorable feelings for Blacks than Whites ($M$s = 8.63 and 7.37, respectively), $F(1, 18) = 4.34$, $p = .05$. The results thus far show parallel patterns of in-group bias on both implicit and explicit measures. The next analyses tested whether the implicit and explicit measures were related to each other.

### Individual Differences

*Implicit–explicit correlations.* On the basis of previous research (Fazio et al., 1995; Payne, 2001), we expected some relationship between implicit and explicit measures, but we expected this relation to be driven by those participants who were unmotivated to avoid appearing prejudiced. Among these individuals, we expected the same positive or negative affective reactions that are detected by the AMP to be reported on the self-report measure.

The relationship between motivation to control prejudice and implicit measures is important for understanding when implicit versus explicit measures are more useful. Motivated participants are likely to express positive racial attitudes on self-report scales. However, they may not be able to control how they perform on implicit measures. As a result, implicit and explicit measures may convey very different information about people's attitudes in situations or among individuals with strong motivation to avoid prejudice.

To examine individual differences, we calculated an AMP score and a feeling thermometer score for each participant. The AMP score was created by subtracting the proportion of pleasant responses after White primes from the pleasant responses after Black primes. Higher values on this score reflect greater positivity toward Blacks. Similarly, a single self-report score was created by subtracting feelings toward Whites from feelings toward Blacks. Higher scores again reflect positivity toward Blacks.

Implicit and explicit estimates of racial attitudes were strongly related ($r = .58$, $p < .01$). The correlation was significant in both the warning ($r = .58$) and no-warning ($r = .58$) conditions. It was also significant for both White participants ($r = .42$) and Black participants ($r = .66$) separately (these correlations were not significantly different, $z = 1.03$, $p = .13$).

To investigate the correlations using absolute rather than comparative estimates, we conducted a second set of analyses. Responses on Black trials and White trials were positively correlated ($r = .35$, $p < .01$). Therefore the following analysis controlled for White and neutral trials when Black trials were analyzed, and controlled for Black and neutral trials when White trials were analyzed. These partial correlations showed that positivity on Black trials was positively related to feelings toward Blacks ($pr = .43$, $p < .01$) and negatively correlated to feelings toward Whites ($pr = -.30$, $p < .05$). Positivity on White trials was positively correlated with feelings toward Whites ($pr = .36$, $p < .01$) but uncorrelated with feelings toward Blacks ($pr = -.16$, *ns)*. These results show some specificity, in that misattributions for each race related more strongly to feelings toward that same race than toward the other race.

*Motivations to control prejudice.* Given this significant correlation between implicit and explicit measures, it is important to test whether the relationship was moderated by participants' racial motivations. We expected the correlation to be stronger for participants who were not motivated to avoid prejudice. Critically, we expected this difference *not* to be driven by highly motivated participants showing less racial bias on the AMP, which appears very difficult to control. Instead, we expected motivations to cause participants to express different attitudes on the explicit measure than their AMP scores would suggest, lowering the correspondence for implicit and explicit measures among the highly motivated. In other words, we predicted no correlation between motivations to control prejudice and AMP scores. However, we predicted a Motivation to Control Prejudice × AMP Score interaction when predicting explicit attitude scores. We tested these predictions first by examining the zero-order correlations between motivation scales and implicit and explicit attitude measures. We then conducted a multiple regression analysis to examine the interactions between attitudes and motivations.

For the analyses involving motivations, we report only the results for White participants. This is because there is little reason to expect that motivations to act without prejudice toward Black individuals would play the same role for Black participants as Whites. For instance, Black participants probably do not react the same way as White participants to the item "Because of today's PC standards, I try to appear nonprejudiced toward Black people." Moreover, the scale items were written specifically targeting prejudice toward Blacks. A series of correlation analyses showed that participants' AMP scores were not significantly related to any of the motivation to control prejudice scales (correlations between $-.15$ and $+.22$, *ns*, $p > .19$). The independence between AMP and motivation measures is consistent with the prediction that highly motivated participants would be unable to modify their AMP scores.

Multiple regression was used to test whether the implicit–explicit relationship was moderated by motivations. In each analysis, the independent and dependent variables were standardized, and the self-report rating was the dependent variable. Because all variables are standardized, beta weights can be interpreted as standardized beta weights (i.e., on a scale from $-1$ to $+1$, much like a partial correlation). The main effects were entered on the first step, followed by the two-way interaction. The relevant statistics are displayed in Table 2.

In a first analysis, we tested the role of Plant and Devine's (1998) Internal Motivation scale. This analysis showed that internal motivation was significantly related to explicit ratings. Critically, the AMP × Internal interaction was significant. Figure 8A shows the relationship between implicit and explicit measures plotted at one standard deviation above and below the mean of internal motivation. The implicit–explicit relationship was strong at one standard deviation below the mean of internal motivation ($B = .74$), but absent at one standard deviation above the mean ($B = -.04$).

When the External Motivation Scale was tested (Figure 8B), it did not significantly moderate the implicit–explicit relationship. The relationships between AMP and explicit scores were slightly stronger for low External ($B = .50$) than high External ($B = .30$), although the AMP × External interaction did not reach significance.

A parallel analysis was run using Dunton and Fazio's (1997) Motivation to Control Prejudiced Responses Scale. First, the Con-

Table 2
*Regression Results Predicting Self-Report Racial Attitudes From Affect Misattribution Procedure (AMP) and Motivations to Control Prejudice, Experiment 6*

| Independent variable | *B* | *SE* | *t*(34) | $p <$ |
|---|---|---|---|---|
| Internal motivation | .64 | .12 | 5.55 | .01 |
| AMP | .35 | .12 | 2.85 | .01 |
| Internal × AMP | −.39 | .13 | 3.03 | .01 |
| External motivation | −.06 | .16 | 0.39 | .70 |
| AMP | .41 | .16 | 2.51 | .02 |
| External × AMP | −.10 | .16 | 0.66 | .52 |
| Concern | .21 | .15 | 1.37 | .18 |
| AMP | .43 | .17 | 2.63 | .01 |
| Concern × AMP | −.42 | .14 | 2.99 | .01 |
| Restraint | −.08 | .16 | 0.51 | .62 |
| AMP | .42 | .16 | 2.66 | .01 |
| Restraint × AMP | −.37 | .15 | 2.43 | .02 |

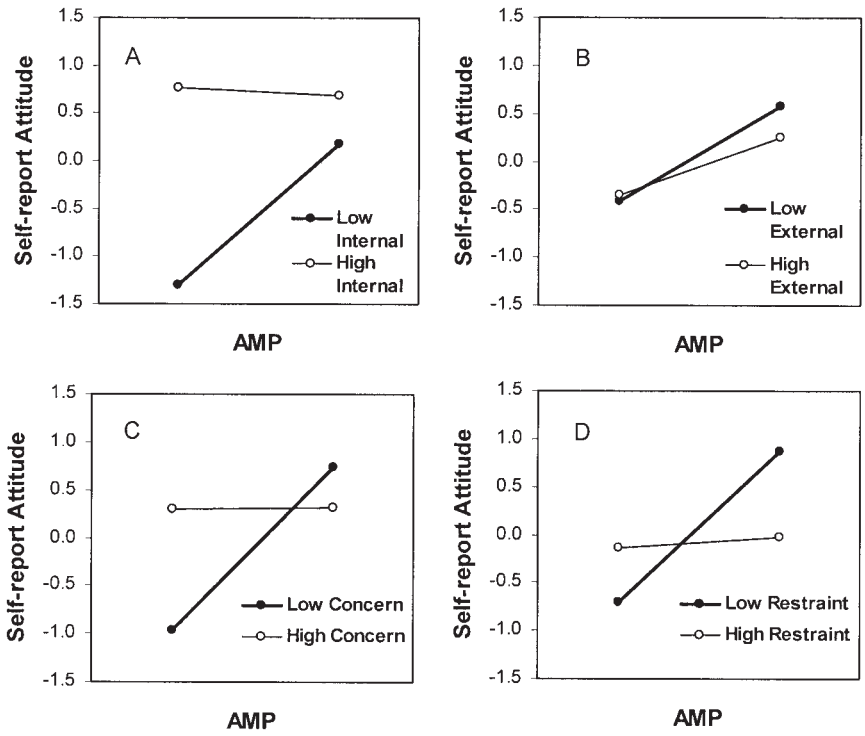


*Figure 8.* Correlations between racial attitudes as measured by AMP and self-reported racial attitudes, plotted separately for individuals high versus low in motivations to control prejudice. Values are plotted at one standard deviation above and below the mean of each scale. All values represent standardized scores. Internal = Internal Motivation to Control Prejudice; External = External Motivation to Control Prejudice; Concern = Concern With Acting Prejudiced; Restraint = Restraint to Avoid Dispute.

cern With Acting Prejudiced Scale was tested. These results converged nicely with the results from the Internal Scale, as illustrated in Figure 8C. The implicit–explicit relationship was strong for those low in Concern, ($B = .85$) but absent for those high in Concern ($B = .01$). The AMP × Concern interaction was significant, as shown in Table 2.

Finally, the Restraint to Avoid Dispute Scale was examined, as illustrated in Figure 8D. The implicit–explicit relationship was significantly moderated by restraint. Individuals low in restraint showed strong correspondence between implicit and explicit measures ($B = .79$); those high in restraint did not ($B = .05$).

### *Discussion*

Participants' responses showed systematic differences in racial evaluations. White participants showed an anti-Black bias, whereas Black participants showed a tendency toward anti-White bias. Even in the controversial domain of race, in which responses evoke social norms and personal standards, performance was not significantly affected by a blatant warning.

The relationship between AMP-estimated attitudes and self-reports was moderated by people's motivations to control prejudice. The relationship observed between the AMP and self-reported attitudes was driven almost entirely by individuals who were unmotivated to control prejudice. Across different motivation scales, the average relationship between AMP and explicit measures was a strong $B = .72$. However, among the highly motivated, that relationship averaged only $B = .10$.

Participants who were more motivated to control prejudice did not perform any differently on the AMP. Instead, they responded differently on the self-report measure. These highly motivated participants showed a lack of correspondence between implicit and explicit measurements, in agreement with dissociations frequently reported in racial attitude studies.

This pattern of moderation by motivations is important for interpreting and using the AMP as an implicit measure. Given the strong implicit–explicit correspondence found with this measure, one might wonder whether the AMP provides new information above and beyond explicit measures. In other words, does this measure provide "incremental validity"? The results of the racial attitudes study suggest that it does, in particular for individuals who are highly motivated to act without prejudice. For these individuals, the AMP and self-report measures were virtually independent of each other, providing completely nonoverlapping information. Of course, it is precisely among individuals or in situations where motivational pressures are high that implicit measures are the most valuable for their ability to circumvent response strategies. The AMP may be more revealing than explicit measures in such situations.

## General Discussion

In these studies, we developed a method for assessing attitudes using the misattributions people make for their own evaluations. This method has an intuitive appeal reminiscent of classic projective tests. Much like inkblots, participants imbued ambiguous symbols with liking or disliking from their own reactions. Unlike classic projective methods, the AMP is scored in an objective, reliable, and quantitative way that can be easily interpreted. It is hoped that by combining the logic of projective tests with the precision and rigor of computer-based priming methods, the AMP will provide a powerful measurement tool.

The evaluations revealed by the AMP appear to be automatic by several different criteria. First, they occurred regardless of people's intentions, whether those intentions are understood as the warnings we manipulated or the reports of correction we measured. Second, priming appeared to be efficient, in the sense that it occurred very rapidly. A third criterion related to automaticity is conscious awareness. As we discuss below, we suspect that participants were unaware, from one trial to the next, when their judgments were being influenced. However, this supposition will need to be directly tested in future research. A separate but related issue is whether participants were aware that they possessed the dispositions revealed by the task. We do not yet know the degree to which the AMP reveals unconsciously held attitudes versus attitudes that are sometimes consciously concealed (see Fazio & Olson, 2003). Nonetheless, the present studies have answered a number of questions about how the AMP performs against the important measurement criteria of validity, reliability, and sensitivity.

### *Measurement Criteria*

The AMP demonstrated evidence of validity in several ways. First, it was sensitive to normatively evaluated items. Second, it predicted intended voting behavior and explicit attitudes toward political candidates. Third, a more complicated relationship emerged in racial attitudes. Overall, the AMP was significantly related to self-reported racial attitudes, but this relation was driven by those individuals who were unmotivated to avoid race bias. Those who were more motivated showed the same degree of race bias on the AMP, but this did not correspond to their self-reported attitudes, which were more positive toward Blacks. An implication of this relationship is that for individuals low in motivation to avoid prejudice, AMP and self-report measures were largely redundant. However, for individuals high in motivation, they were almost completely independent. Critically, in this case the AMP revealed something about participants over and above what self-reports could yield.

Just as important as validity is reliability. Across the six studies, the AMP showed an average internal consistency of $\alpha = .88$. Cunningham et al. (2001) recently argued that the IAT and priming measures of race bias are more strongly related to each other and to explicit measures once they are corrected statistically for unreliability. The AMP showed substantial relationships with other measures without the need to correct for unreliability using sophisticated statistical procedures.

Having provided evidence for validity and reliability, it is worth highlighting that the sensitivity of a measure can be reflected in its effect size. Across studies, the (weighted) average effect size of the AMP was 1.25 (which translates into an $r$ of .53). This large effect increases the power to detect real differences with fewer observations, and to replicate findings.[1]

An additional advantage of the affect misattribution procedure is ease of administration. A typical version of the AMP can be completed in less than 5 min. The program is relatively straightforward to implement and the materials may be obtained by contacting B. Keith Payne via e-mail.

### *Potential Mechanisms*

The fact that participants were influenced by the primes despite blatant warnings is perhaps the most striking aspect of these studies. Fast primes, targets, and SOAs can explain some of this, but not most of it. These procedural details are only a small part of the story.

One potential account is that participants did not correct their judgments because they felt that the primes provided a "hint" of how the experimenters wanted them to evaluate the pictographs. However, our data are inconsistent with this experimenter demand account. First, in the studies of political candidates and racial bias, different people showed different biases. The demand explanation would have to assume that pro-Bush voters inferred a pro-Bush hint from the experimenters, whereas pro-Kerry voters inferred the opposite. Further, this account would have to hold that White participants in the race bias experiment inferred an anti-Black hint from the experimenters.

We find these assumptions unlikely. However, even if it were true that participants' inferences about the demands of the experimenters were guided by their own attitudes, this would still mean that the AMP was indirectly measuring participants' attitudes. The sequence would be that attitudes influenced inferences about what the experimenters were hinting, which in turn influenced responses. In either case participants would project their own atti-

[1] One interesting potential explanation suggested by a reviewer for the large effect sizes is that residual effects from previous trials may influence a current trial. As a result, when the prime on a previous trial was incongruent with the prime on a current trial, participants may feel a large swing in affect, resulting in larger judgmental effects, compared with when primes are congruent from trial to trial. To test this hypothesis, we recoded each randomly ordered trial in Experiment 2 (the experiment with the largest effect size) by whether the previous trial contained a pleasant, neutral, or unpleasant trial. We then analyzed responses to pleasant, unpleasant, and neutral trials as a function of whether the previous trial was congruent or incongruent. When the current prime was positive, there was no effect of the previous prime's valence, $F(2, 82) = 0.41$, $p = .66$. Likewise, when the current prime was neutral, there was no effect of the previous trial, $F(2, 76) = 1.50$, $p = .23$. When the current prime was negative, the previous trial exerted a marginally significant effect, $F(2, 82) = 2.59$, $p = .08$. However, this effect was not entirely in the direction predicted by this account. Responses were least likely to be pleasant when the previous prime was pleasant ($M = 0.21$, $SD = 0.30$), but were most likely to be pleasant when the previous trial was neutral ($M = 0.31$, $SD = 0.30$), and pleasant responses when the previous trial was unpleasant were intermediate ($M = 0.25$, $SD = 0.31$). For all types of previous trial primes, the effect size of the AMP remained very large (all $d$s $> 1.84$). Thus, the congruence with previous trials does not appear to account well for the effect sizes observed.

tudes—in one case, onto the pictographs, in the other, onto the experimenters.

A second, more direct reason also shows that our data are inconsistent with a demand account. Whatever subtle demands participants inferred from the primes would have been directly contradicted by the blatant warnings that "we are interested in studying how people can avoid being biased" and our explicit requests to "please try your absolute best not to let the real-life images bias your judgment of the drawings!" In using these warnings, we pitted demand characteristics against the priming effect, but it remained nonetheless.

A different account based on participants' perceptions of bias is more consistent with our results. We suggest that the resistance to correction stems, at least in part, from participants not feeling that they are being influenced (see Pronin, Gilovich, & Ross, 2004, for a review of related phenomena). This may be because participants are aware of the outcome of their attributions, but not the process (Nisbett & Wilson, 1977). That is, once an evaluation of a prime has been attributed to the target, participants come to perceive the target differently. Regardless of the original source of affect, once the attribution has occurred participants may actually like or dislike the pictograph itself.

Consider the example of the swastika. Because of its historical usage by the Third Reich, many people feel revulsion at this symbol. However the Nazis did not invent this symbol; it is an ancient symbol that can be seen in artifacts from cultures all over the world. Nonetheless, it would not make much difference to ask a person to separate historical associations from esthetics. Given that it is already imbued with a particular historical meaning for many people, the symbol itself has become noxious. The same may be true for pictographs in our studies. Once a transfer of affect has occurred, the target itself may become evidence of its pleasantness or unpleasantness. This attribution appears to happen very quickly, as most of our studies presented primes for 75 ms and targets for 100 ms. At this speed, the transfer may be a mental slight of hand that is invisible to participants. The rapid transfer of affect from prime to target may explain why participants are ineffective at correcting because by the time participants consider whether they are being biased on any particular trial, the pictograph itself is a source of affect. We suspect that *if* participants recognized that their judgment on any given trial was being influenced by the prime, they would be able to correct by simply giving the opposite response. However, if they felt inclined to evaluate a pictograph positively because the pictograph actually seemed pleasant, then an adjustment would seem inappropriate.

This account would explain why the priming effect occurred in the race bias study even among individuals highly motivated to avoid prejudice. It would also explain why the effect occurred among individuals who reported trying a great deal to correct their judgments for bias. Finally, it would explain why the effect occurred even at slow presentation rates, when participants had ample opportunity to implement whatever strategy they wished. This monitoring failure account is consistent with research showing that affective priming in the Murphy and Zajonc (1993) paradigm is not reduced by attribution interventions that portray affective reactions as nondiagnostic for judgments (Winkielman et al., 1997). In short, we suggest that the misattribution effect was difficult to control because participants did not believe they were experiencing it.

### *Unanswered Questions and Future Directions*

This theoretical account will, of course, need to be directly tested. To that end, the subjective experience of the person as he or she makes each evaluation is of interest. What do participants think is the source of their evaluation? A common finding in studies of brain-lesioned patients is that when participants execute a behavior without full knowledge of why they did so, they create a plausible story that explains their behavior. In the affect misattribution procedure, would participants "confabulate" by finding features of the pictographs to justify their evaluations? Along with this account, several additional questions deserve further attention. One question is whether the same effects would be found if verbal materials rather than pictures were used as primes. A second is whether affect is unique in eliciting these misattributions (Murphy & Zajonc, 1993). Can the same misattribution occur with concepts, such as traits and stereotypes? Or even more complex propositions? Finally, we suspect that the target pictographs were effective in eliciting misattributions primarily because they contain very little evaluative information. This sparseness renders them very ambiguous, and therefore easily categorized on the basis of other accessible information. There are probably any number of other items that would produce the same result (e.g., ink blots, polygons), so long as they provide little basis for evaluation on their own. We are currently investigating many of these questions.

The premise behind projective measurement has captured the imagination of psychologists since Freud (1911) proposed the concept of projection and Rorschach (1921) proposed a way to measure it. Although many questions remain to be answered, the preliminary evidence for the affect misattribution procedure appears promising. The AMP offers a tool that is at once projective and implicit, avoiding some of the limitations of these approaches while preserving their strengths.

## References

Banse, R., & Gawronski, B. (2003). Die Skala Motivation zu vorurteilsfreiem Verhalten: Psychometrische Eigenschaften und Validitat [The scale Motivation to Act Without Prejudice: Psychometric properties and validity]. *Diagnostica, 49,* 4–43.

Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79,* 631–643.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin, 115,* 401–423.

Campbell, D. T. (1950). The indirect assessment of social attitudes. *Psychological Bulletin, 47,* 15–38.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12,* 163–170.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality & Social Psychology Bulletin, 23,* 316–326.

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.

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. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology, 54,* 297–327.

Freud, S. (1911). *General psychological theory: Papers on metapsychology.* New York: Collier Books.

Frijda, N. H. (1999). Emotions and hedonic experience. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being* (pp. 190–210). New York: Russell Sage Foundation.

Gazzaniga, M. S. (1985). *The social brain.* New York: Basic Books.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 192,* 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.

Greenwald, A. G., & Nosek, B. A. (2001). Health of the implicit association test at age 3. *Zeitschrift Fuer Experimentelle Psychologie, 48,* 85–93.

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition & Emotion, 15,* 143–165.

Hoffman, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (in press). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality & Social Psychology Bulletin.*

Ito, T. A., & Cacioppo, J. T. (2000). Electrophysiological evidence of implicit and explicit categorization processes. *Journal of Experimental Social Psychology, 36,* 660–676.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Erlbaum.

Kawakami, K., & Dovidio, J. F. (2001). The reliability of implicit stereotyping. *Personality & Social Psychology Bulletin, 27,* 212–225.

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

Lilienfield, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. *Psychological Science in the Public Interest, 1,* 27–66.

Martin, L. L., Seta, J. J., & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. *Journal of Personality and Social Psychology, 59,* 27–37.

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

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General, 106,* 226–254.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Palfai, T. P., & Ostafin, B. D. (2003). Alcohol-related motivational tendencies in hazardous drinkers: Assessing implicit response tendencies using the modified IAT. *Behaviour Research & Therapy, 41,* 1149–1162.

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.

Phelps, E. A., O'Conner, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12,* 729–738.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75,* 811–832.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review, 111,* 781–799.

Rorschach, H. (1921). *Psychodiagnostics: A diagnostic test based on perception.* New York: Grune & Stratton.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110,* 145–172.

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.

Sherman, S. J., Rose, J. S., Koch, K., Presson, C. C., & Chassin, L. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. *Journal of Social & Clinical Psychology, 22,* 13–39.

Strack, F., Schwarz, N., Bless, H., Kubler, A., & Wanke, M. (1993). Awareness of the influence as a determinant of assimilation versus contrast. *European Journal of Social Psychology, 23,* 53–62.

Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 141–208). Mahwah, NJ: Erlbaum.

Wegner, D. M. (2003). The mind's best trick. *Trends in Cognitive Sciences, 72,* 65–69.

Wells, G. L., & Loftus, E. F. (2003). Eyewitness memory for people and events. In A. M. Goldstein (Ed.), *Handbook of psychology: Forensic psychology* (pp. 149–160). New York: Wiley.

Wiers, R. W., Van Woerden, N., Smulders, F. T. Y., & De Jong, P. J. (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology, 111,* 648–658.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin, 116,* 117–142.

Winkielman, P., Zajonc, R. B., & Schwarz, N. (1997). Subliminal affective priming resists attributional interventions. *Cognition & Emotion, 11,* 433–465.

Wittenbrink, B., Judd, C., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72,* 262–274.

Zanna, M. P., & Rempel, J. K. (1988). Attitudes: A new look at an old concept. In D. Bar-Tal & A. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 315–334). New York: Cambridge University Press.

## Appendix

### Picture Stimuli From the International Affective Picture System (IAPS) Used in Experiments 1–4

| Pleasant images | | Unpleasant images | |
|---|---|---|---|
| IAPS Number | IAPS Name | IAPS Number | IAPS Name |
| 1440 | Seal | 1111 | Snakes |
| 1460 | Baby tiger* | 1220 | Spider |
| 1610 | Bunny* | 1275 | Roaches |
| 1750 | Bunnies | 1301 | Dog |
| 1710 | Puppies | 2110 | Angry face |
| 1811 | Monkeys | 2722 | Jail |
| 1920 | Porpoise | 2900 | Crying boy |
| 1999 | Mickey | 3160 | Eye infection* |
| 2057 | Baby (1)* | 3181 | Battered woman* |
| 2070 | Baby (2) | 3300 | Sick child* |
| 2209 | Woman & girl* | 3301 | Bandaged boy* |
| 2340 | Family | 6190 | Gun & hand* |
| 2540 | Mom & child* | 6610 | Gun |
| 2550 | Couple* | 8230 | Bloody boxer* |
| 2655 | Kid & dog* | 9230 | Oil fire |
| 5470 | Astronaut | 9300 | Dirty toilet* |
| 5760 | Lake* | 9340 | Garbage (1) |
| 5830 | Ocean* | 9390 | Garbage (2) |
| 5831 | Seagulls | 9470 | Ruins |
| 7325 | Watermelon* | 9561 | Injured kitty* |
| 7330 | Sunday* | 9584 | Dental work* |
| 7502 | Castle | 9594 | Needle* |
| 8370 | Rafting* | 9810 | KKK* |
| 8501 | Money | 9920 | Burnt car* |

*Note.* Items with an asterisk were used only in Experiment 4.

Received August 30, 2004
Revision received February 16, 2005
Accepted April 8, 2005 ■