Journal of Experimental Social Psychology **37,** 244–252 (2001)
doi:10.1006/jesp.2000.1456, available online at http://www.idealibrary.com on IDEAL®

Evaluative versus Conceptual Judgments in Automatic Stereotyping and Prejudice

Bernd Wittenbrink

University of Chicago

and

Charles M. Judd and Bernadette Park

University of Colorado

Received February 14, 2000; revised June 29, 2000; accepted July 2, 2000; published online April 3, 2001

The goal of the study reported in this article was to examine whether automatic processes in stereotype and prejudice activation are sensitive to task characteristics of the assessment procedure and whether these influences may account for existing inconsistencies that have recently been reported in the literature on automatic racial prejudice. Using a sequential priming paradigm with subliminal primes ("BLACK" and "WHITE") to examine automatic prejudice, the study varied the judgment task in which the priming procedure was presented. Whereas half of the participants were asked to perform a lexical decision task (word/nonword), the remaining participants made evaluative judgments (good/bad). Results showed reliable influences of the judgment task on the observed pattern of priming effects. Moreover, the priming effects found in both conditions replicated the respective results reported in previous research that had used either evaluative or conceptual judgment tasks (Fazio, Jackson, Dunton, & Williams, 1995; Wittenbrink, Judd, & Park, 1997). In addition, the response time measure also showed different relationships with explicit measures of racial prejudice, depending on the judgment condition. In addition to their implications for the assessment of automatic stereotyping and prejudice these results suggest that automatic responses are not as invariant as it is sometimes posited. © 2001 Academic Press

Much recent research has documented that group attitudes and stereotypes may be activated spontaneously from memory, without the perceiver's intent, merely triggered by exposure to a relevant stimulus cue in the environment (e.g., Banaji & Greenwald, 1995; Banaji & Hardin, 1996; Blair & Banaji, 1996; Chen & Bargh, 1997; Devine, 1989; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz 1998; Kawakami, Dion, & Dovidio, 1998; Lepore & Brown, 1997; Macrae, Bodenhausen, Milne, Thorn, & Castelli, 1997; Macrae, Stangor, & Milne, 1994; Wittenbrink, Judd, & Park, 1997). Such automatic activation occurs within a few hundred milliseconds after stimulus exposure and requires only very limited cognitive resources (Shiffrin & Schneider, 1977). It is generally assumed that the perceiver has only very limited control over the activation and often remains unaware of it and its potential influences on subsequent behaviors.

One reason why this work is of interest is because of its potential to provide new measures of prejudice and stereotyping, relatively freed from the normative constraints and social demand characteristics that contaminate traditional questionnaire measures. If one can assess prejudice at the automatic level, without respondents being aware of their responses, then these procedures offer promise as a "bona fide pipeline" to respondents "true" attitudes and beliefs (Fazio et al., 1995).

Because of this potential, considerable debate has emerged concerning how measures of automatic stereotyping and prejudice relate to traditional explicit questionnaire measures. Whereas a number of studies indicate that indi-

This research was supported by funding from the University of Chicago Graduate School of Business to the first author and by NIMH Grant R01 MH45049 to the second and third authors.

Address correspondence and reprint requests to Bernd Wittenbrink, Center for Decision Research, Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, IL 60637. E-mail: bernd.wittenbrink@gsb.uchicago.edu.

0022-1031/01 $35.00
Copyright © 2001 by Academic Press
All rights of reproduction in any form reserved.

vidual differences in spontaneous prejudice are unrelated to differences on explicit measures (Banaji & Hardin, 1996; Devine, 1989; Dovidio et al., 1997, Studies 1 and 3; Fazio et al., 1995; Greenwald et al., 1998), other studies have found moderately strong relationships (Dovidio et al., 1997, Study 2; Lepore & Brown, 1997; Kawakami et al., 1998; Wittenbrink et al., 1997).

One obvious explanation for these discrepancies is the limited validity of explicit self-report measures, due to, most notably, their sensitivity to normative pressures as well as other context factors such as question order and wording. (e.g., Strack & Martin, 1987; Sudman, Bradburn, & Schwarz, 1996). Undoubtedly, explicit self-report measures have their limitations. In the present article, however, we focus on an alternative explanation for these discrepancies, namely on the possibility that the different automatic response measures used in the research to date have themselves measured somewhat different things. We focus in particular on two lines of work that have offered especially divergent conclusions: the evaluative priming procedure developed by Fazio and his colleagues (Fazio et al., 1995) and the concept priming procedure that we have used in our own research (Wittenbrink et al., 1997).

The particular procedures used by both Fazio et al. (1995) and Wittenbrink et al. (1997) involve response time assessments in a sequential priming paradigm. Participants are presented with a priming stimulus (e.g., a group reference or exemplar), followed by a target stimulus that requires a response. The time it takes participants to respond to this stimulus serves as the primary dependent measure. These times are taken as indicative of the associative strength between prime and target stimuli in memory (Neely, 1977; Posner & Snyder, 1975). In addition, the nature of the stimulus presentation insures that the observed priming effects are attributable to automatic activation of the target concept and not influenced by controlled strategies of the participant.

Although both lines of research share these general characteristics, the actual judgment task that participants perform differ substantially. Specifically, in our own work (Wittenbrink et al., 1997), we have used a variation of Meyer and Schvaneveldt's (1971) lexical decision task in which group primes (BLACK and WHITE) are followed by target items that consist of words and nonwords (e.g., SMART and RAMST). The participants' task is to decide whether the target item represents an actual word, thus requiring concept identification. In contrast, the studies by Fazio and his colleagues (1995) ask participants to determine the evaluative implications of the target items (is this good or bad?). Having to determine a stimulus' identity on the one hand and its evaluative implications on the other hand are tasks which may activate somewhat different associations from long-term memory. Stimulus attributes that are relevant to one are not necessarily important for the other, and, thus, the memory contents primed under these different judgment contexts may differ in important ways. Indeed, prior work on the role of processing goals in knowledge activation suggests that different judgment goals in priming tasks are likely to yield different priming effects (see Gollwitzer & Moskowitz, 1996; Macrae et al., 1997; Smith & Lerner, 1986).

The primary purpose of the present research therefore is to clarify whether the seemingly inconsistent findings obtained by the two lines of research are in part explained by the differences in judgment task. Before presenting the results of a study that tested this hypothesis, we review the main differences in findings reported by the two lines of research.

The concept priming procedure used by Wittenbrink et al. involves presentation of primes that represent the relevant ethnic groups (e.g., BLACK and WHITE) or a neutral prime, followed by target letter strings for which participants have to decide whether the strings form a word. On the critical trials, four types of target words are used, positively and negatively valenced attributes that are either stereotypic of White Americans (e.g., educated and greedy) or stereotypic of African Americans (e.g., athletic and poor). The primes are presented outside of participants' conscious awareness to preclude controlled processes.

Wittenbrink et al. found that for White American participants the BLACK prime facilitated responses to negatively valenced attributes stereotypic of African Americans and the WHITE prime facilitated responses to positively valenced attributes stereotypic of White Americans. There was no evidence for more generalized facilitation, either to similarly valenced words that were not stereotypic of the primed group or to stereotypic words that were of the opposite valence (e.g., positively valenced words stereotypic of African Americans). Additionally, individual differences in this pattern of automatic "stereotypic prejudice" were relatively highly correlated with a number of explicit racial attitudes measures, gathered in a supposedly unrelated experimental session [e.g., the correlation with McConahay, Hardee, & Batts' (1981) Modern Racism Scale was .40; $p < .001$].

The experimental paradigm developed by Fazio and his colleagues differs from our procedure in several respects. Participants are presented, supposedly as part of a memory task, with photographs of either African American and White American targets which are followed by either positively or negatively valenced adjectives. As mentioned above, the participant's task is to determine the evaluative connotation of the target adjectives. Different from the concept priming procedure described above, priming stimuli are clearly visible. Exclusion of controlled processing is instead sought by means of a short stimulus onset asynchrony (SOA) between priming and target stimuli. Moreover, the target adjectives, while having clear evaluative

connotations, are not chosen to be stereotypic of either target group (e.g., attractive or disgusting).

With White American participants, Fazio et al. found evidence for automatic prejudice, such that negative adjectives showed stronger facilitation when preceded by an outgroup prime than when preceded by an ingroup prime. Unlike the Wittenbrink et al. results, this facilitation did not depend on whether the adjectives were stereotypic of the target group since no stereotypic words were included. Thus these facilitation results suggest a more generalized pattern of prejudice, not dependent on the stereotypic match between the target adjective and the prime. Importantly, Fazio et al. reported that individual differences in this form of automatic prejudice were uncorrelated with MRS scores.

In an attempt to understand the differences between these two sets of results, we conducted a study that followed in large part the procedure used in our previous research (Wittenbrink et al., 1997); however, we manipulated the nature of the reaction time task. Half of the participants were asked to make conceptual judgments (word/nonword) of the target stimuli, whereas the remaining participants judged the stimuli for their evaluative connotation (good/bad). Moreover, the reaction time task also included additional, nonstereotypic, but valenced, target items similar to those used by Fazio et al.

METHOD

Participants

The participants were 161 students recruited from the University of Chicago campus and paid $10 for their participation. Data from 11 participants who identified themselves as African American were excluded from the analyses.

Procedures

The study consisted of three ostensibly unrelated experiments, one on "judgmental accuracy," one on "word comprehension," and one involving a questionnaire. Upon arrival, participants received an introduction to all three experiments that emphasized their seemingly different nature.

In the first experiment, participants were asked to identify the ethnicity of individuals based on their first names. As in our previous research, the purpose of this task was to strengthen the association between the relevant ethnic groups and the lexical labels "BLACK" and "WHITE" that would subsequently serve as group primes. Participants judged 20 first names, half of which were stereotypically African American (e.g., Lamont) and half of which were stereotypically White American (e.g., Mark). Participants indicated whether each individual was likely to be Black or White.

Next, participants were told they would work on a "word comprehension" experiment. For this task, they would go through a large number of computer-based trials. On each trial they would first see a string of X's, followed by a letter sequence. This letter sequence would require their response which they should give as quickly and accurately as possible. Depending on the experimental condition to which they had been randomly assigned, they were given one of two judgment instructions. In the "Conceptual Judgment" condition, participants were given standard instructions for a lexical decision task, asking them to judge whether the target sequence constituted a word or a nonword. The response keys were labeled "YES" and "NO." In the "Evaluative Judgment" condition, participants received instructions following Fazio et al. Specifically, they were asked to indicate whether the target item made them think of something positive and good or negative and bad. Participants in this condition were instructed that some of the target letter sequences would actually not form correct English words. Nevertheless, they should respond according to their first inclination. Response keys in this condition were labeled "GOOD" and "BAD." Aside from these judgment instructions the reaction time procedure was identical in both conditions.

Once participants had completed the reaction time task, they were handed a questionnaire containing six different explicit measures of racial prejudice. Specifically, the questionnaire included a measure commonly used to assess feelings toward social groups, the "feeling thermometer" rating scale, as well as a set of five belief-based prejudice measures: the Modern Racism Scale by McConahay et al. (1981), the Pro-Black and Anti-Black scales of Katz and Hass (1988), the Diversity and the Discrimination Scales (both Wittenbrink et al., 1997).

Reaction Time Task Stimuli

Presentation of experimental stimuli and data collection was controlled by the PSYSCOPE software package (Version 1.2.2, Cohen, MacWhinney, Flatt, & Provost, 1993) on Apple PowerMacintosh 7200/120 computers equipped with PSYSCOPE button boxes and 14-in. monitors. The monitors were set to a resolution of 860 $\times$ 640 pixels. All stimuli were presented in 18 point Times Macintosh font.

Each trial of the reaction time task, started with a fixation point ("+") in the center of the computer screen. The fixation point appeared for 1000 ms and was immediately followed by the prime. After 15 ms, the prime was replaced by a masking stimulus ("XXXXX"), which remained on the screen for 250 ms. Following the masking stimulus, the target letter sequence appeared for another 250 ms. The computer then paused until the participant had responded.

Thus, the stimulus presentation followed the procedure used in Wittenbrink et al. However, we used a different set of target items, including not only items stereotypic of the

TABLE 1
Target Items Used in Reaction Time Task

Items stereotypic of African Americans	Nonstereotypic items	Items stereotypic of White Americans
charming	appealing	intelligent
religious	delightful	successful
merry	desirable	ambitious
cheerful	fabulous	industrious
athletic	favorable	educated
expressive	likable	responsible
streetwise	pleasant	wealthy
musical	wonderful	ethical
poor	awful	exploitative
dishonest	disturbing	materialistic
complaining	horrible	stuffy
violent	irritating	boring
shiftless	offensive	callous
superstitious	repulsive	uptight
lazy	rotten	greedy
threatening	upsetting	selfish

two target groups, White and African Americans, but also valenced nonstereotypic adjectives (e.g., appealing).[1] These were chosen to be similar to the Fazio et al. stimulus materials, but to not include adjectives that could potentially be seen as stereotypic. Table 1 presents the full set of target adjectives. Given the nature of the judgment task in the conceptual judgment condition (i.e., word/nonword), target items also included 16 nonword fillers.

In total, the set of trials fully crossed three independent factors, prime ("BLACK," "WHITE," and neutral), target item (stereotypic of African American, stereotypic of White American, nonstereotypic, and nonword), and valence (positive, and negative),[2] resulting in a total of 192 experimental trials, which were individually randomized for each participant. Additionally, 10 practice trials were presented at the start of the reaction time task.

Stimulus Timing

In the Wittenbrink et al. study, pretesting showed that participants were not aware of the primes given their 15-ms presentation. We conducted a similar pretest, recruiting an additional 10 participants. They first took part in the name identification "experiment," just as participants did in the main study. They then completed a random subset of 48 trials from the main reaction time procedure identical to those used in the experiment proper, except that we told participants that words would briefly appear on the screen prior to the stimulus mask and that they should attempt to identify them. Despite these instructions, pretest participants could not identify the primes correctly on any of the 480 trials.

[1] Items stereotypic of the two target groups were based on Judd et al. (1995).

[2] The nonword target items did not vary systematically in valence. Instead the design included an equal number of nonword target items for the two cells (positive/negative). These items, of course, only served as fillers to make the word/nonword judgment task meaningful and were not included in the analyses.

RESULTS

We first present the results for each of the two judgment conditions separately, reporting both analyses of mean reaction times as well as correlations with the explicit attitudes measures. Next, we then present a comparison of results between conditions.

To examine response facilitation, we examined two focused contrasts of theoretical interest (following Wittenbrink et al., defined in Table 2). The first contrast (I) captures what we have previously called stereotypic prejudice: facilitation in response to valenced stereotypic items. More specifically, this contrast examines whether negatively valenced items stereotypic of African Americans are facilitated by the "BLACK" prime and positively valenced items stereotypic of White Americans are facilitated by the "WHITE" prime.

The second contrast (II) captures what we have called generalized prejudice, that negatively valenced items, regardless of their stereotypicality, are more facilitated by the "BLACK" prime than are positively valenced items and the reverse is true for the "WHITE" prime. We have argued that the facilitation results of Fazio et al. are of this more generalized form, since they did not reflect facilitation on stereotypic items. Accordingly, a variant on this generalized prejudice contrast (IIA) examines whether facilitation is consistent with generalized prejudice when omitting the stereotypic items.

Prior to the analyses, we examined the distributions for the response latency data. As is common for this kind of measure (Ratcliff, 1993), the data showed positive skew and included small numbers of outliers. To deal with these problems, response latencies faster than 150 ms and slower than 2 standard deviations above the individual's mean response time were deleted. This resulted in an exclusion rate of 3%, similar to that applied by other researchers. Additionally, analyses of the latency data were conducted following an inverse transformation. For ease of interpretation, however, we report mean values in milliseconds.

Conceptual Judgment Condition

Response facilitation. To examine response facilitation due to the two group primes, the response latency for each target item following each of the group primes was subtracted from the latency for that same target item following the neutral prime. Accordingly, more positive values indicate greater response facilitation due to a group prime. These facilitation scores were analyzed as a function of

TABLE 2
Contrast Weights

	Prime type: BLACK — item stereotypicality			Prime type: WHITE — item stereotypicality		
	AA	NON	WA	AA	NON	WA
I. Stereotypic Prejudice						
Item Valence						
Positive	−1	0	0	0	0	+1
Negative	+1	0	0	0	0	−1
II. Generalized Prejudice						
Item Valence						
Positive	−1	−1	−1	+1	+1	+1
Negative	+1	+1	+1	−1	−1	−1
IIA. Generalized Prejudice (excluding stereotypic items)						
Item Valence						
Positive	0	−1	−1	+1	+1	0
Negative	0	+1	+1	−1	−1	0

Note. AA = Target items stereotypic of African Americans; NON = Target items stereotypic of neither group; WA = Target items stereotypic of White Americans.

group prime, item stereotypicality, and item valence, with all factors varying within subjects. Mean facilitation values for the conceptual judgment condition are given in Table 3.

Considering first the conditions that were included in our previous research, that is, looking only at items that are stereotypic of one group or the other—omitting the nonstereotypic items, the pattern of results obtained for this judgment task replicates our previous results. The data show that stereotypic prime/target combinations (BLACK/AA items and WHITE/WA items) yield overall stronger facilitation than do counterstereotypic combinations [M's = 9.37 vs 0.49; $F(1,76) = 5.49$; $p = .022$). The means make clear, however, that this overall stereotyping effect is attributable solely to the facilitation observed in two cells of the design: BLACK primes/negative AA items and WHITE primes/positive WA items. In other words, as in Wittenbrink et al., the stereotyping effect was again qualified by an outgroup valence bias. BLACK primes facilitated especially those African American target items that were negative (M = 19.70) rather than those that were positive (M = 1.59). In contrast, WHITE primes increased response speed for positive stereotypic items (M = 17.43), while barely affecting responses to negative stereotypic items ($M = -1.21$). The corresponding stereotypic prejudice contrast (Contrast I) is significant, again replicating our previous results [$F(1,76) = 6.79$; $p = .011$].

TABLE 3
Mean Response Facilitation for Conceptual Judgment Condition (in Milliseconds)

	Prime type: BLACK — item stereotypicality			Prime type: WHITE — item stereotypicality		
	AA	NON	WA	AA	NON	WA
Item valence						
Positive	1.59	−1.18	1.93	−0.66	−1.69	17.43*
Negative	19.70*	−18.30	0.02	0.69	−1.10	−1.21

Note. AA = Target items stereotypic of African Americans; NON = Target items stereotypic of neither group; WA = Target items stereotypic of White Americans.

The cell means suggest that a more generalized form of prejudice, involving facilitation for nonstereotypic items, is not found with this conceptual judgment task. And indeed, a test of the generalized prejudice contrast (II) shows it to be nonsignificant [$F(1,76) = 1.98$; $p = .168$]. Leaving out the stereotypic items, and only testing for generalized prejudice on the non-stereotypic and counterstereotypic items, confirms the absence of this effect (Contrast IIA: $F < 1$).

In addition to tests of these theoretically motivated contrasts, we were also interested in whether facilitation in an absolute sense occurred. Accordingly, we tested whether cell means differed significantly from zero. Again replicating our earlier results, only two means were significant: when negatively valenced African American stereotypic items followed the BLACK prime and when positively valenced White American stereotypic items followed the WHITE prime (indicated with asterisks in Table 3).

In sum, replicating our earlier results, we find evidence for stereotypic prejudice but no generalization of these priming effects to nonstereotypic or counterstereotypic items.

Relationships with explicit measures. Separate scores on each of the five explicit belief-based measures were

TABLE 4
Explicit Prejudice Measures: Internal Consistency and Intercorrelation (across Judgment Conditions)

Explicit measures	A	B	C	D	E	F
A. Modern Racism $\alpha = .81$	1.00					
B. Pro-Black $\alpha = .76$	.62***	1.00				
C. Anti-Black $\alpha = .85$	.42***	.20*	1.00			
D. Diversity $\alpha = .64$	.59***	.52***	.38***	1.00		
E. Discrimination $\alpha = .89$	.75***	.46***	.56***	.60***	1.00	
F. Thermometer	.48***	.28**	.22**	.36***	.40***	1.00

* $p < .05$.
** $p < .01$.
*** $p < .0001$.

computed. Additionally, a score on the thermometer measure was computed by subtracting the rating given to Blacks from that given to Whites. Table 4 presents the intercorrelations among all six explicit measures (higher scores on all indicate more explicit prejudice). All five belief-based attitude scales are intercorrelated substantially and in the expected direction. The thermometer measure is also reliably, although more weakly, correlated with the other explicit measures.[3]

To examine the relationships between these explicit prejudice scores and the patterns of facilitation from the response time procedure, we computed contrast scores for each individual on the two primary contrasts: stereotypic prejudice and generalized prejudice. Higher scores on these contrasts indicate that an individual participant is showing greater automatic prejudice. These scores were correlated with the six explicit prejudice measures, with the results presented in the left portion of Table 5. These correlations replicate what we found earlier. The stereotype specific contrast of automatic prejudice consistently correlates more highly with these explicit measures than does the generalized contrast. Indeed, there are significant correlations between the stereotypic prejudice contrast and the MRS, Diversity, and Discrimination scales. Only the Discrimination scale correlates significantly with generalized prejudice. The thermometer measure fails to correlate significantly with either contrast.

[3] Given that we did not expect differences for participants' explicit responses as a result of the judgment conditions (nor were they observed, F values for all mean differences < 1), we list these scale indices independent of condition.

Evaluative Judgment Condition

Response facilitation. Table 6 presents the mean facilitation scores for participants in the evaluative judgment condition. A quick inspection of these scores reveals that

TABLE 5
Correlations between RT Contrasts and Explicit Measures by Judgment Condition

	Judgment condition			
	Conceptual judgment		Evaluative judgment	
	contrast		contrast	
	Stereotypic prejudice	Generalized prejudice	Stereotypic prejudice	Generalized prejudice
Explicit measures				
Modern Racism	.40***	.20	.18	.11
Pro-Black	.17	.09	.13	.02
Anti-Black	.12	.13	−.14	−.12
Diversity	.26*	.23*	.14	.14
Discrimination	.24*	.14	.14	.14
Thermometer	.19	.07	.16	.29*

* $p < .05$.
*** $p < .0001$.

TABLE 6
Mean Response Facilitation for Evaluative Judgment Condition (in Milliseconds)

	Prime type					
	BLACK			WHITE		
	item stereotypicality			item stereotypicality		
	AA	NON	WA	AA	NON	WA
Item valence						
Positive	1.45	−0.01	1.93	10.11*	11.27*	14.65*
Negative	20.94*	9.08*	14.38*	−0.40	−0.84	0.52

Note. AA = Target items stereotypic of African Americans; NON = Target items stereotypic of neither group; WA = Target items stereotypic of White Americans.

patterns of facilitation seem quite different in this condition. Whereas before significant facilitation was found only for negatively valenced African American items following the BLACK prime and positively valenced White American items following the WHITE prime, in this condition significant facilitation (indicated by the asterisks in Table 6) is found for all negatively valenced items, regardless of stereotypicality, following the BLACK prime, and all positively valenced items, regardless of stereotypicality, following the WHITE prime, suggesting the more generalized form of automatic prejudice.

Tests of both stereotypic prejudice and generalized prejudice contrasts are significant in this condition [Contrast I: $F(1, 72) = 9.05$; $p = .004$; Contrast II: $F(1, 72) = 20.24$; $p < .0001$]. In sharp contrast, however, to the results from the conceptual task condition, strong evidence for generalized prejudice is found here when excluding the stereotypic items [contrast IIA: $F(1, 72) = 11.75$; $p = .001$]. And in fact, when only nonstereotypic items are included (thus only using items like those used by Fazio et al.), we find significant generalized prejudice [$F(1,72) = 4.49$; $p = .039$], replicating the results from Fazio et al.

Relationship with explicit measures. Correlation coefficients for the within-subject contrasts from this judgment condition and the explicit prejudice measures are listed on the right side of Table 5. Whereas in the conceptual judgment condition, we found systematic and reliable relationships between the explicit prejudice measures and participants' tendency to display automatic stereotypic prejudice, the valence judgment condition yields very different results. First, with regard to the belief-based measures, correlations with the two prejudice contrasts are generally positive but relatively weak and nonsignificant. However, the Thermometer measure of attitudes is significantly correlated with the generalized prejudice contrast. This measure, unlike the attitude scales which ask respondents to indicate their level of agreement with certain group-related beliefs, may be tapping a more affective or feeling-based form of explicit prejudice. It is perhaps therefore not surprising that it now correlates with generalized prejudice when the task is an evaluative one, whereas it was the belief-based measures that correlated with stereotypic prejudice from the conceptual task. Each task seems to manifest a characteristic form of automatic prejudice and these different forms then show correlations with explicit measures that seem to tap similar explicit components.

Comparison of the Judgment Conditions

We next looked at potential differences in the overall level of facilitation between the two conditions. As we already mentioned, the evaluative judgment condition resulted in broader priming effects, with 6 instead of 2 of the 12 cells of the design yielding significant facilitation. However, the overall facilitation means do not differ by condition [conceptual: $M = 1.43$; evaluative: $M = 6.92$; $F(1,149) = 1.96$; $p = .164$].

To compare the influences of priming and target item factors in the two conditions, we conducted analyses with task as a between-participants factor. Of the effects involving the task factor, the only significant one was the interaction between task and the generalized prejudice contrast excluding the stereotypic items (IIA) [$F(1,149) = 5.58$; $p = .019$]. This interaction captures what we have already described as the primary difference in the facilitation patterns from the two judgment conditions: automatic prejudice is found only for stereotypic items for the conceptual judgment condition, whereas it generalizes to nonstereotypic items in the other condition.

DISCUSSION

Our goal was to determine whether some of the apparently conflicting results in past research on automatic prejudice may be due to differences in the sorts of tasks that have been used. More specifically, it seemed to us that a conceptual task, such as the lexical task that we have used previously, and an evaluative task, like that developed by Fazio et al., may tap into different memory contents, with consequences for the pattern of correlations manifested with explicit prejudice measures.

The results we obtained nicely replicated our own earlier results, in the conceptual task condition, and those of Fazio et al., in the evaluative task condition. Specifically, we found that the conceptual task resulted in a response pattern that we have called stereotypic prejudice, with group primes facilitating responses to valenced items only if they are stereotypic of the primed group (negatively valenced items for the out-group; positively valenced ones for the in-group). Additionally, individual differences in the strength of this automatic response pattern were reliably correlated with scores on explicit racial attitude scales that tap beliefs about racial inequities in our society. When the task was an evaluative one, however, a more generalized form of auto-

matic prejudice was observed, with the out-group prime facilitating responses to negatively valenced items regardless of whether they were stereotypic of the primed group and in-group primes facilitating responses to positively valenced items regardless of their stereotypicality. Individual differences in this pattern of automatic prejudice were not correlated significantly with the belief-based explicit measures of racial attitudes, although there was a significant correlation with a more affective-based explicit measure, i.e., the feeling thermometer ratings.

Implications for the Assessment of Group Attitudes and Stereotyping

An immediate implication of the present results concerns the assessment of attitudes. As mentioned above, one particular reason why automatic attitudes have received such attention in the past several years is that they seem to promise bona fide measures of people's true attitudes. As much as we share the hope for veridical measures of social constructs, we do believe that the search for the single, unbiased measure of a person's "true" sentiments ought to be elusive and that such automatic measures will result at best in a true assessment of only some *aspect* or *component* of the underlying attitude.

The present results then provide further support for the notion that attitudes are not necessarily based on a single, homogenous, representation in memory, but instead are based on multiple, potentially diverse and discrepant, memory contents. Therefore, different measures may reveal different "attitudes," depending on what aspects of the underlying representation the measures make salient. With regard to explicitly measured attitudes and beliefs, such effects of assessment context are widely recognized (Schwarz & Strack, 1991; Tesser, 1978; Tourangeau & Rasinski, 1988; Wilson & Hodges, 1992; Zanna & Rempel, 1988). The present results suggest that these effects are not limited to explicit measures and that therefore—even at the level of automatic activation—there is not just one single measure of a person's attitude toward a given attitude object.

With regard to the particular contents that the two judgment conditions tap in the present experiment, the three-component view of attitudes offers one possible explanation. That is, there is widespread support for the assumption that attitudes include an affective basis, as well as cognitive components and behavioral predispositions, and that these three components need not always be entirely consistent (Eagly & Chaiken, 1998). Explicit measures clearly vary in which of these components they most strongly tap. Likewise, it seems perfectly reasonable in light of this component view of attitudes that our more conceptual task seems to tap more directly the cognitive or belief-based component of the attitude, while the evaluative task used by Fazio et al. is more closely picking up the more affective aspects of the automatic attitude. Our results are certainly consistent with more general propositions concerning the relative independence of affective and cognitive systems (e.g., Cacioppo, Gardner, & Berntson, 1999; LeDoux, 1996; Murphy & Zajonc, 1993; Zajonc, 1980), although the present study was certainly not designed to explicitly tests these propositions. More pragmatically, the present data suggest that future research can benefit from a careful differentiation of these aspects of prejudice.

Automaticity in Social Cognitive Functioning

Beyond their immediate relevance for intergroup attitudes and stereotypes, we believe the present results have broader implications for social psychology's conceptualization of automaticity. Automatic responses have commonly been considered obligatory responses that are inevitably triggered when a certain stimulus is encountered in the environment. The present results are at odds with such a view. In the present experiment, activation of group stereotypes and attitudes was "automatic" in the sense that it occurred involuntarily, in response to stimuli of which participants were unaware, and in the sense that it was observed within a time frame too short to be affected by participants' active control. Nevertheless, these automatic responses were not invariant and were not obligatory, reflexive reactions to a certain stimulus. Rather, identical priming stimuli led to quite different automatic responses in the two judgment conditions.

Our present argument that automatic responses are indeed malleable rather than fixed, reflexive stimulus responses is to some extent consistent with Bargh's (1994) notion of "conditional automaticity." With this concept, Bargh contrasts automatic cognitive processes that are triggered invariantly, whenever the presence of the triggering stimulus in the environment is registered, from those automatic processes that occur only if certain preconditions are met. Such preconditions, according to Bargh, may include recent relevant controlled thought processes ("postconscious automaticity") or prior activation of relevant goal states ("goal-dependent automaticity"). In particular, then the present findings are consistent with the notion of goal-dependent automaticity. In this study, participants' automatic activation of group stereotypes and attitudes depended on their objective when performing the judgment task. Trying to determine the evaluative implications of stimuli appearing on the computer screen led to different spontaneous responses than trying to determine their identity.

CONCLUSION

Social scientific research of the past 40 years has documented that racial attitudes among White Americans are complex and multidetermined. If we are to understand the responses of White Americans to racial minorities in U.S. society, then it is clear that we need to understand racial beliefs and values in all their complexity. Knowing the associations that are stored in long-term memory and being able to assess

these in an "unfiltered" and "pure" manner is undoubtedly useful in this endeavor. But just as explicit measures of racial attitudes may tap different aspects and components of those attitudes, so too may different automatic assessment procedures.

REFERENCES

Banaji, M. R., & Greenwald, A. G. (1995). Implicit gender stereotyping in judgments of fame. *Journal of Personality and Social Psychology,* **68,** 181–198.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science,* **7,** 136–141.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 1–40). Hillsdale, NJ: Erlbaum.

Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology,* **70,** 1142–1163.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology,* **76,** 839–855.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology,* **33,** 541–560.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphical system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers,* **25,** 257–271.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology,* **56,** 5–18.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology,* **33,** 510–540.

Eagly, A. H., & Chaiken, S. (1998). Attitude structure and function. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology,* (4th ed., Vol. 1, pp. 269–322). Boston: McGraw–Hill.

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.

Gollwitzer, P. M., & Moskowitz, G. B. (1996). Goal effects on action and cognition. In E.T. Higgins, & A.W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 361–399) New York: Guilford.

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.

Judd, C. M., Park, B., Ryan, C. S., Brauer, M., & Kraus, S. (1995). Stereotypes and ethnocentrism: Diverging interethnic perceptions of African American and White American youth. *Journal of Personality and Social Psychology,* **69,** 460–481.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology,* **55,** 893–905.

Kawakami, K., Dion, K. L., & Dovidio, J. F. (1998). Racial prejudice and stereotype activation. *Personality and Social Psychology Bulletin,* **24,** 407–416.

LeDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life.* New York: Simon & Schuster.

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology,* **72,** 275–287.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., Thorn, T. M. J., & Castelli, L. (1997). On the activation of social stereotypes: The moderating role of processing objectives. *Journal of Experimental Social Psychology,* **33,** 471–489.

Macrae, C. N., Stangor, C., & Milne, A. B. (1994). Activating social stereotypes: A functional analysis. *Journal of Experimental Social Psychology,* **30,** 370–389.

McConahay, J. B., Hardee, B. B., & Batts, V. (1981). Has racism declined in America? It depends on who is asking and what is asked. *Journal of Conflict Resolution,* **25,** 563–579.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology,* **90,** 227–234.

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.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R.L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Ratcliff R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin,* **114,** 510–532.

Schwarz, N., & Strack, F. (1991). Context effects in attitude surveys: Applying cognitive theory to social research. In W. Stroebe, & M. Hewstone (Eds.), *European review of social psychology* (Vol. 2, pp. 31–49). London: Wiley.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review,* **84,** 127–190.

Smith, E. R., & Lerner, M. (1986). Development of automatism of social judgments. *Journal of Personality and Social Psychology,* **50,** 246–259.

Strack, F., & Martin, L. (1987). Thinking, judging, and communicating: A process account of context effects in attitude surveys. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 123–148). New York: Springer-Verlag.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology.* San Francisco, CA: Jossey–Bass.

Tesser, A. (1978). Self-generated attitude change. In L. Berkowitz (Ed.), *Advances in experimental social psychology,* (Vol. 11, pp. 289–338). New York: Academic Press.

Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin,* **103,** 299–314.

Wilson, T. D., & Hodges, S. D. (1992). Attitudes as temporary constructions. In L. L. Martin, & A. Tesser (Eds.), *The construction of social judgements* (pp. 37–66). Hillsdale, NJ: Erlbaum.

Wittenbrink, B., Judd, C. M., & 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.

Zajonc, R. B. (1980). Feeling and Thinking. Preferences need no inferences. *American Psychologist,* **35,** 151–175.

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