Swiss Journal of Psychology 58 (4), 1999, 241–250

Implicit stereotyping and prejudice and the primed Stroop task

Kerry Kawakami[1], Kenneth L. Dion[2] and John F. Dovidio[3]

[1] University of Nijmegen
[2] University of Toronto
[3] Colgate University

In the present study, automatic stereotype activation related to racial categories was examined utilizing a primed Stroop task. The speed of participants' ink-color naming of stereotypic and nonstereotypic target words following Black and White category primes were compared; slower naming times are presumed to reflect interference from automatic activation. The results provide support for automatic activation of implicit prejudice and stereotypes. With respect to prejudice, naming latencies tended to be slower for positive words following White than Black primes and slower for negative words following Black than White primes. With regard to stereotypes, participants demonstrated slower naming latencies for Black stereotypes, primarily those that were negatively valenced, following Black than White category primes. These findings provide further evidence of the automatic activation of stereotypes and prejudice that occurs without intention.

Key words: Stroop task, automatic activation, racial stereotypes, implicit stereotyping

In recent years, stereotype theorizing has been dominated by the social cognitive approach (Park & Hastie, 1987; Schneider, 1991). This viewpoint has emphasized the importance of social categorization to the process of stereotyping and its researchers have attempted to understand not only the antecedents and consequences of categorization but also the link between categorization and stereotyping (Hamilton & Trolier, 1986; Tajfel & Turner, 1986; Taylor, 1981; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). In general, categories are described as having a close and immediate tie with what people see and how they judge it (Allport, 1954). Once a person is recognized as belonging to a specific category, perceivers may infer that the person has many of the qualities shared by other category members. According to Taylor (1981), "Social categories are used as a means of organizing incoming person information. Stereotypes can be thought of as attributes that are tagged to category labels (e.g., race, sex) and imputed to individuals as a function of their being placed in that category" (p. 110).

Recently, Greenwald and Banaji (1995; Banaji & Greenwald, 1995) have emphasized the importance of distinguishing between explicit and implicit indices of stereotyping. Explicit measures of stereotypes operate in a conscious mode and are exemplified by traditional measures of these constructs (e.g., Katz & Braly, 1933). Implicit stereotypes, in contrast, operate in an unconscious fashion. Implicit stereotypes are "introspectively unidentified (or inaccurately identified) traces of past experience that mediate attributions of qualities to members of a social category" (Greenwald & Banaji, 1995, p. 15).

Response latency procedures and other techniques, often borrowed from cognitive psychology, have been frequently used in social psychology to assess the content of implicit stereotypes (Banaji & Greenwald, 1995; Dovidio, Evans, & Tyler, 1986; Gaertner & McLaughlin, 1983; Hense, Penner, & Nelson, 1995), as well as implicit attitudes (Dovidio & Fazio, 1991; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Fazio, Sanbonmatsu, Powell, &

Kardes, 1986). These techniques potentially assess automatic activation and offer a conceptually and empirically different perspective on both stereotypes and attitudes than traditional self-report measures.

Research using a range of response latency procedures has demonstrated that stereotypes may operate like other semantically related concepts (e.g., doctor-nurse) to facilitate responses and decision making. Gaertner and McLaughlin (1983), for example, used a lexical decision task in which participants were asked to make judgments about whether pairs of letter strings presented simultaneously were both words (Meyer & Schvaneveldt, 1971). They found that both high and low prejudiced individuals made their decisions about letter strings faster when the words "Blacks" (or "Negroes") and "Whites" were paired with stereotype consistent than with stereotype inconsistent words. Studies by Dovidio et al. (1986), Baker and Devine (1988), and Zarate and Smith (1990) using different paradigms offer generally convergent results.

Although the latter studies demonstrate that people can respond faster to semantically related social stimuli (i.e., stereotypes) than to semantically unrelated social stimuli, they do not necessarily demonstrate preconscious or automatic processes (Bargh, 1994; Greenwald & Banaji, 1995; Kihlstrom, 1990). In automatic activation, the simple perception of a diagnostic group physical feature (for example, through the mere presentation of an actual category member, a photograph of a category member, or a written category label) is all that is necessary to activate stereotypes (Bargh, 1996). In each of these studies, there were opportunities for more controlled processes to operate. For instance, participants were made aware that the study focussed on judgments about racial categories, were aware of the potential relationships between the stimuli, or the task-related parameters (stimulus onset asynchronies of 2500 ms) may have permitted conscious processing that could allow intentional activation of negative stereotyping within the paradigm (see Judd, Park, Ryan, Brauer, & Kraus, 1995). Thus, these findings may represent "controlled" rather than "automatic" processing (Posner & Snyder, 1975).

Qualities of Automaticity

But what does it mean that a process is automatic? In accordance with current findings in cognitive psychology, Bargh (1994) proposes that most complex mental processes are not exclusively automatic or controlled but are composed of some automatic and some controlled features. Instead of categorizing a process as "automatic" and thereby inferring that it has all the qualities of automaticity, he recommends that researchers abandon the idea that automaticity is all-or-none, and advises them to detail the particular qualities of automaticity which they study.

Accordingly, Bargh (1994; 1996) defines automaticity with respect to at least three specific qualities – intentionality, awareness, and efficiency. If any one of these qualities is present, automatic processing is involved to some degree; all three qualities do not have to be present. Intentionality refers to an individuals' goals during the instigation of the process. To the extent that individuals do not intend to initiate a process – if the process is triggered simply by the stimulus – it is considered to be unintentional. An absence of awareness is another quality of automaticity. When people are unaware of the stimulus itself, or when people are aware of the stimulus but not aware of the effect that this event has on their thoughts and actions (Nisbett & Wilson, 1977), researchers can assume that the obtained effects of the stimulus are due to automatic processing (Lombardi, Higgins, & Bargh, 1987; Martin, Seta, & Crelia, 1990). Efficiency is a third important dimension. To the extent that a process requires few attentional resources, the process is considered to be highly efficient and thus automatic. If cognitive capacity and time are necessary for the process to occur, the process is not efficient.

Because each of these qualities relates to the degree of automaticity involved, Bargh and others (Anderson, Spielman, & Bargh, 1992; Bargh, 1989; Bargh, 1994) have recently argued that researchers should be more specific in their claims and support for the unintentionality, unawareness, or high efficiency of social processes. Just as explicit measures of stereotypes may be influenced by the context in which they are elicited (Strack & Martin, 1987), the expression of implicit stereotypes may also be shaped by techniques which differentially reflect each of these qualities. The primary goal of the present research was to examine the importance of one specific quality of automaticity, the intentionality of the process of stereotype activation.

Intentionality

Although recent research has demonstrated automatic stereotypic activation for a range of social groups, including Blacks (Devine, 1989; Lepore & Brown, 1997; Kawakami, Dion, & Dovidio, 1998; Wittenbrink, Judd, & Park, 1997), women and men (Blair & Banaji, 1996; Banaji & Hardin, 1996; Banaji, Hardin, & Rothman, 1993), elderly people (Hense et al., 1995; Perdue & Gurtman, 1990), Asians (Macrae, Bodenhausen, & Milne, 1995), skinheads (Kawakami, Dovidio, Moll, Hermsen, & Russin, in press), child abusers, soccer hooligans, (Macrae, Stangor, & Milne, 1994), and professors (Dijksterhuis & van Knippenberg, 1996), most of these studies have uti-

lized paradigms specifically intended to examine the efficiency or awareness of the stereotyping process. For example, by limiting the stimulus onset asynchrony (SOA) between the presentation of the prime and the presentation of the target word, researchers can demonstrate the efficiency and speed of a process (Blair & Banaji, 1996; Kawakami et al., 1998). By presenting the category prime either subliminally (Devine, 1989) or by separating the priming session from the measurement of stereotype activation into ostensibly separate studies (Macrae et al., 1994; 1995), researchers can demonstrate that a process occurs without the awareness of the prime itself or awareness of the influence of the prime on the activation process. Few studies, however, have specifically focussed on the influence of intention to the stereotype activation (Fiske, 1989; see, for example, Blair & Banaji, 1996). To examine the impact of intention on the automatic activation of racial stereotypes and to provide potentially convergent evidence already obtained with paradigms that focus on the awareness (Wittenbrink et al., 1997) and efficiency (Kawakami et al., 1998) of this process, the present research employed a relatively new technique, a primed Stroop task.

In general, participants in the standard Stroop task are presented with a series of words in a variety of colors and are asked to name the color in which the word is presented as quickly as possible (Stroop, 1935). When the letters of the word spell a color name different than the color in which the word is printed, participants are slower at naming the color. Research demonstrates that as the word's semantic association to the concept of color increases, so does its potential to interfere with speed of naming ink-colors (Klein, 1964). For example, when the words "blue," "sky," and "XXX" are printed in red ink, participants are slowest at naming the color of the word "blue," followed by the word "sky," and finally the string of letters "XXX."

In a primed Stroop task, participants are first primed with a "distractor" word, which they are simply asked to read. This prime is then followed by the target word printed in various colors. The participants' task is to indicate the color of the target word. Research with the primed Stroop task has demonstrated that even target words unrelated to color can interfere with naming ink-colors when previously activated. Warren (1972; 1974), for example, found strong interference when the priming word and the stimulus word were highly associated and systematically weaker interference when the associative strength between the stimulus pair decreased. In general, cognitive psychologists have found that the degree of interference in naming ink-colors is a function of the activation state of the word's semantic meaning – the greater the activation level, the greater the amount of processing resources needed to inhibit it which in turn results in slower responses in color naming (Bargh & Pratto, 1986; Klein, 1964; MacLeod, 1991).

The Stroop task is a particularly powerful method to examine intentionality because it pits controlled inhibition and automatic activation processes against one another. Specifically, if stereotype activation is unintentional, participants in this task will be unable to ignore the semantic stereotype-related target words, which will then interfere with the color-naming responses for these words. Stereotype activation related to the primed Stroop task can be considered unintentional, and thus automatic, because it occurs despite the participants' intention to instigate this process. Even when participants are attempting to ignore the semantic meaning of stereotypes by focussing their attention away from the content of these words and towards their primary task of naming of the ink-colors, this process still occurs (Bargh, 1994; Higgins, van Hook, & Dorfman, 1988; Pratto & John, 1991). Specifically, to perform effectively on the Stroop task, participants' best strategy is *not* to activate stereotypic words or their meanings. If participants nonetheless activate stereotypes, this effect occurs *despite* the participants' intentions.

In summary, the primary aim of the present study was to examine the role of intention in the hypothesized automatic activation of racial stereotypes utilizing a primed Stroop task. Specifically, the speed of participants' naming of ink-colors of stereotypic and nonstereotypic target words following category primes were compared. The category primes of interest consisted of Black and White representations. Positive and negative stereotypes and nonstereotypes were utilized as target stimulus. Stereotypic words comprised traits associated more strongly with Black than White categories. Nonstereotypic words were traits not differentially associated with either Black or White categories. A Category Prime × Stereotype Trait interaction effect was predicted in which participants were expected to respond slower at naming the ink-color of Black stereotypic words following Black than White category primes. No difference in color naming latencies were anticipated for nonstereotypic words.

Method

Participants and Design

The final sample consisted of 27 (12 female and 15 male) visitors to the Ontario Science Center. Only visitors from Canada or the United States and who ranged from 18 to approximately 50 years old were selected to participate. Data from one color-blind participant, one non-white participant, and one tourist from outside of North America were excluded from consideration in the analyses. The

study included three independent variables in a 2 (Black vs. White Prime) × 2 (Positive vs. Negative Trait) × 2 (Stereotypical vs. Nonstereotypical Trait) within-subject design.

Materials and Procedure

The primary aim of the experimental procedure was to compare the time needed to name the color of stereotypic and nonstereotypic traits following Black and White primes. Participants were informed that each trial consisted of an initial asterisk and two words that would appear in sequence on a computer screen. They were told that their task was to read the first word silently and to state the color of the second word into the microphone. They were further instructed to ignore the meaning of the second word but to name the color as quickly and as accurately as possible. Finally, to ensure that participants interpreted the primes as racial groups and not as color names, they were specifically informed that the prime words consisted of a human social category, an animal category, or a string of letters.

Two main types of category primes were utilized: a target category (i.e., Blacks) and a comparison category (i.e., Whites). Although in accordance with previous research our primary comparison was between Black and White category primes (Dovidio et al., 1986; Dovidio et al., 1997; Kawakami et al., 1998), a third CCCCCC prime was included to examine subsequently whether the category priming effects were due to facilitation versus inhibition processes (Fazio et al., 1986).

The target arrays included stereotypic and nonstereotypic traits selected on the basis of results from a pilot study. In this study a separate group of undergraduate students from the University of Toronto ($N = 84$) completed a diagnostic ratio for Black and White target groups (McCauley & Stitt, 1978; McCauley, Stitt, & Segal, 1980). Eight traits that significantly differentiated ($p < .05$) between Black and White categories on the basis of paired *t*-tests and were attributed more to Blacks were categorized as stereotypes. Eight traits that did not significantly differentiate between the two groups were categorized as nonstereotypes.

It is important to note that the primary statistical comparison was the effect of Black versus White primes on color-naming of stereotypes and nonstereotypes. Because of this focus on within-subject comparisons of category primes, minor variations in the stimulus material and individual differences between participants such as age and speed of responding are conceptually less important. Nonetheless, attempts were made to match the set of positive and negative stereotypes and nonstereotypes on word length and word frequency (Kucera & Francis, 1967). This degree of experimental control reduced the number of words to eight stereotypes (e.g., four positive and negative) and eight nonstereotypes (e.g., four positive and negative) which is similar to the number of traits utilized by other researchers (Dovidio et al., 1986; Gaertner & McLaughlin, 1983; Macrae et al., 1995).

A 2 (Positive vs. Negative Trait) × 2 (Stereotypical vs. Nonstereotypical Trait) analysis of variance performed separately on difference scores for attributions to White and Blacks, word length, and word frequency, demonstrated only one significant main effect. As expected, the difference between trait attributions to Black and Whites by participants in the pilot study was larger for stereotypes ($M = 10.30$) than nonstereotypes ($M = 2.50$), $F(1,3) = 47.16$, $p < .001$. No other significant main effects or interactions were found, demonstrating that the positive and negative stereotypic and nonstereotypic traits did not differ on word length or word frequency. Please refer to Table 1 for a complete list of the experimental target words.

Participants were tested individually with a 386DX personal computer with a high resolution (SVGA) quality color monitor and a display refresh rate of 60 Hz. This equipment is comparable to other computers used in this type of research on automatic activation (Fazio et al., 1995), with the same presentation and monitor refresh (60 Hz) rates (see Lepore & Brown, 1997; Wittenbrink et al., 1997). The primed Stroop task was programmed in Turbo Pascal IDE. Timing events were within millisecond resolution utilizing TPTIMER version 1.0. Specifically, an

Table 1: Experimental target words

	Difference*	Word length	Word frequency
Positive stereotypes			
cool	15	4	62
funny	6	5	41
musical	13	7	85
athletic	13	8	18
Negative stereotypes			
poor	14	4	113
angry	9	5	18
tough	4	5	22
criminal	8	8	24
Positive nonstereotypes			
glad	–3	4	38
loyal	–1	5	18
reliable	–2	8	22
friendly	3	8	61
Negative nonstereotypes			
odd	–2	3	44
selfish	–4	7	8
cunning	–3	7	5
confused	–2	8	44

* Differential trait attributions to Black versus White categories

asterisk was presented in the center of the computer screen for 300 ms to prepare participants, followed by a blank screen for 500 ms. Next, the prime appeared for 950 ms, followed by a blank screen for 50 ms before the onset of the target word. The target array was presented until the voice key was activated and was followed by a blank screen for 1000 ms before the next trial.

In total, four blocks of trials were presented which consisted of 48 trials in which the White, Black, and neutral primes were presented 16 times with each traits. Within a given block, each category was presented with all four positive and negative stereotypical and nonstereotypical traits but in different colors (i.e., red, blue, green, and yellow). Across blocks the same traits were once more presented but the word color – category combination varied. Over all trials each category prime was, therefore, presented 48 times and each adjective trait was presented 12 times. Trials and blocks were randomized independently for each participant. Participants were given breaks at the end of each block and instructed to press the appropriate key when they were ready to continue the experiment.

Errors were recorded by the experimenter who was present throughout the experiment and were defined as stutters, mispronunciations, and stating the wrong color. Trials were also classified as errors if the voice key was inappropriately triggered (e.g., when participants spoke too softly or exhaled heavily into the microphone before naming the word color). Before the experimental trials, participants were presented with a practice block of 24 trials in which three prime categories (cat, dog, and XXX) were presented with positive and negative stereotypic and nonstereotypic associations for the category "cat."

To specifically examine the participants' intentions, 28 undergraduate students not utilized in the main study were presented with a primed Stroop task and extensively debriefed in a separate study. Specifically, after 192 color naming trials similar to those in the present study (but related to male and female stereotypes), participants were asked what their goals and intentions were during the task. As expected, no subject mentioned the importance of category primes or the relationship between category primes and target words in describing the purpose of the study or their intentions. Instead, participants uniformly stated that, in order to indicate the color of the target word as fast as possible and to be correct in their color-naming, they attempted to focus on the color of the target word and to ignore its meaning. When specifically asked about the effect of the category primes on the color naming of the target words, none of the subjects were aware of the influence of these primes on their naming latencies (Bargh, 1994; Nisbett & Wilson, 1977). Similarly, during debriefing of participants in the main study of the present research, none of the participants identified the relationship between the category primes and the target words, and they all indicated their intention to say the color of the target word as quickly as possible. These reports are consistent with the assumption that interference on the Stroop task occurs without intent (Bargh, 1994; Higgins et al., 1988; Pratto & John, 1991).

Results

In accordance with recommendations by Ratcliff (1993), response latencies exceeding a 1000 ms cutoff were identified as outliers.[1] Response latencies related to errors (1.6%) and outliers (3.5%) were removed and replaced with missing values. For each participant, the mean naming latencies of ink-colors for each of the four conditions (positive vs. negative traits X stereotypical vs. nonstereotypical traits) were computed for the main priming categories (Black, White). This was accomplished by taking the mean of the four color words for each of the four trait dimensions in each condition.

To examine automatic stereotype activation a 2 (Black vs. White Prime) × 2 (Positive vs. Negative Trait) × 2 (Stereotypical vs. Nonstereotypical Trait) repeated measures analysis of variance was performed on the response latencies. The predicted Prime × Stereotype interaction was significant, $F(1, 26) = 5.99$, $p < .05$. As expected, color-naming responses were somewhat slower to Black stereotypic words following a Black prime ($M = 672$) than a White prime ($M = 666$), $t(26) = 1.57$, $p < .13$. Responses were also somewhat slower to nonstereotypic words following a White prime, ($M = 671$) than a Black prime ($M = 664$), $t(26) = 1.90$, $p < .07$.

Consistent with previous evidence of implicit negative racial attitudes (Dovidio et al., 1997; Fazio et al., 1995), a marginally significant Prime × Valence interaction was also obtained, $F(1, 26) = 3.05$, $p < .09$. While participants responded slower to positively valenced words following a White prime ($M = 673$) than a Black prime ($M = 668$), they responded slower to negatively valenced words following a Black prime ($M = 669$) than a White prime ($M =$

1 Supplementary analyses using alternative measures of identifying outliers produced a similar pattern of results. For example, for one set of analyses, response latencies exceeding three or more standard deviations beyond each participant's mean response times were identified as outliers. The mean percentage of outliers using this cut-off was 1.4%. For these analyses, the Prime × Stereotype interaction was marginally significant, $F(1, 26) = 3.76$, $p = .06$; the Prime × Valence interaction was significant, $F(1, 26) = 4.36$, $p < .05$; and the Prime × Stereotype × Valence effect was also significant, $F(1, 26) = 7.18$, $p < .01$.

Table 2: Mean naming latencies of ink-colors in milliseconds

	Stereotypes		Nonstereotypes	
	Positive	Negative	Positive	Negative
Black prime	668 (63)[a]	677 (64)	667 (61)	660 (63)
White prime	672 (67)	660 (69)	673 (64)	668 (61)

[a] The standard deviation for each mean is given within parentheses

664). Neither of these two specific pairwise comparisons, however, was statistically significant.

The above two-way interactions were qualified by a significant Prime × Stereotype × Valence effect, $F(1, 26) = 4.30$, $p < .05$. See Table 2. Consistent with previous findings on automatic racial stereotyping (Wittenbrink et al., 1997), simple effects analysis demonstrated the Prime × Stereotype interaction was significant for negative words, $F(1, 26) = 9.24$, $p < .01$, but not for positive words, $F(1, 26) = .03$, $p = .86$. For negative words, color-naming responses were slower to Black stereotypes following a Black prime ($M = 677$) than a White prime ($M = 660$), $t(26) = 2.81$, $p < .01$, but responses were not significantly different as a function of Black primes ($M = 660$) vs. White primes ($M = 668$) for nonstereotypes, $t(26) = 1.46$, $p = .16$. In contrast, for positive words there was no effect for prime for either the Black stereotypes or the nonstereotypes.

Although the CCCCCC prime condition was included initially to examine facilitation and inhibition effects following Black and White primes, in accordance with a number of recent studies (Bargh, Chaiken, Raymond, & Hymes, 1996; Blair & Banaji, 1996; Kawakami et al., 1998), we chose to dispense with this method. Due to conceptual differences between a letter prime and social category prime (e.g., semantic meaning), the validity of the CCCCCC prime to function as a neutral prime condition has been seriously questioned (Bargh, Chaiken, Govender, & Pratto, 1992; Fazio et al., 1995; Jonides & Mack, 1984). It is important to note here, however, that even when facilitation and inhibition scores are computed (i.e., difference scores from the CCCCCC prime), the significance levels of the results do not change. Because the same scores are subtracted from the response times following both Black and White primes, the relative differences between conditions is maintained. In examining the response latencies related to the CCCCCC prime, the results related to positive and negative stereotypes and nonstereotypes (*Ms* = 686, 668, 666, and 675, respectively) provide as much evidence for facilitation effects, which are difficult to explain theoretically in a Stroop paradigm, as interference effects. These findings, therefore, provide further evidence that letter strings may be both conceptually and empirically different from category representations (Jonides & Mack, 1984).

To examine baseline levels, the present study included a second more valid indicator – responses to nonstereotypic words which were not differentially associated with Black and White racial groups (see Banaji & Hardin, 1996; Blair & Banaji, 1996). Consistent with the hypothesized interference effect of automatic stereotype activation elicited by Black primes, the longest color-naming latency was related to negative Black stereotypes following Black primes. Specifically, following Black primes, participants were significantly slower at responding to negative Black stereotypes ($M = 677$) than negative nonstereotypes ($M = 660$), $t(26) = 2.53$, $p < .05$. Furthermore, no difference between color-naming latencies of negative Black stereotypes ($M = 660$) and negative nonstereotypes ($M = 668$) was found following White primes, $t(26) = 1.43$, $p = .16$.

Discussion

The primary aim of the present study was to examine further the automaticity of stereotype activation related to racial categories. To achieve this goal, a primed Stroop task was utilized to investigate the importance of intentionality to stereotype activation. If as predicted, stereotyping is automatic, category primes would be expected to facilitate the activation of associated stereotypes regardless of the intentions of the participant to instigate, to not instigate, or even to inhibit this process.

In particular, it was assumed that the predicted effects of priming in a Stroop task paradigm would be due to a "spreading activation" between items in an organized semantic category (Higgins et al., 1988; MacLeod, 1991). With regard to Black stereotype activation, presentation of the Black prime was expected to activate any strong associations (evaluations, stereotypes, exemplars) to that category. Or as succinctly summarized by Allport (1954, p. 21), "A person with dark brown skin will activate whatever concept of Negro is dominant in our mind." According to this theoretical account, when the prime and target word were associated, exposure to the prime automatically increases the accessibility of the target word's meaning, which makes it more difficult for the participant to attend to the ink-color of the word and ignore its meaning. Evidence for this process of response interference would be reflected by longer color-naming latencies in conditions in which automatic activation was hypothesized to occur.

In general, the results are consistent with these predictions and provide support for the concept of facilitated accessibility to associated characteristics. Specifically, the present study offers further evidence that racial stereo-

types and attitudes can be activated without intention. Consistent with the hypothesis that priming racial categories automatically activates associated evaluations and characteristics (stereotypic traits) and consequently produces longer color-naming latencies, Prime × Valence and Prime × Stereotype interactions were obtained. In terms of implicit attitudes, response times were longer to positive words following a White prime than a Black prime, and slower to negative words following a Black prime than a White prime. The generally symmetric pattern of this pair of differences suggests that the racial attitudes of participants were characterized as much by anti-Black sentiment as by pro-White bias.

With respect to implicit stereotypes, color-naming latencies for Black stereotypic words were longer following Black primes than White primes. This difference did not occur for nonstereotypic words, which were not more associated with either Blacks or Whites. Although not predicted, this effect was qualified by the valence of the word. The Prime × Stereotype interaction was significant for negative but not positive traits. Only traits that were negative and stereotypic were facilitated following Black primes.

In accordance with theorizing by Bargh (1999) and Kawakami et al. (in press), these results suggest that stereotype activation occurs regardless of the intentions of the participant to instigate or not instigate the process. In particular, results from the primed Stroop task indicate that this activation occurs even though participants attempted to direct their attention to naming the ink-color of the target words. Although participants were specifically instructed before presentation of the Stroop task to ignore the semantic meaning of the stereotypes, and the results from the pilot study demonstrate that this was indeed their intention, participants were unable to fulfill these goals and not activate stereotypic traits. Thus even though they did not intend to read the target word, and the results from the pilot study indicate that they were unaware of the effect of the prime on their delayed responding to stereotypes, category priming interfered with their ability to name the color of these associated words.

One alternative explanation for the effect of the category prime on the response latencies is that, even though participants were explicitly informed that the primes referred to social categories, the word "BLACK" itself is negative and so the effects are not necessarily related to stereotype activation but to differentially valenced color associations independent of race. The results related to the nonstereotypic words, however, does not support this explanation.[2] Specifically, if the findings in the present study were simply due to valence or connotations related to the color "BLACK" and not the cultural associations of the social category "BLACK," then interference would also be expected for negative nonstereotypic words. The effects of the category primes, however, were specific to negative Black *stereotypes.* Naming latencies related to positive stereotypes, negative nonstereotypes, and positive nonstereotypes did not differ as a function of Black and White category primes. Furthermore, because the primary comparisons focussed on within subject differences related to response latencies to Black versus White primes, discrepancies between stereotypic and nonstereotypic target words and individual differences between subjects can also not account for these effects. Nonetheless, because of the relatively small number of target words per condition and the small sample size in the present study, further research is recommended to replicate the current findings utilizing new target categories with different stereotypes and larger stimulus sets. The compatibility of our results with those of Wittenbrink et al. (1997), who used a broader range of stereotypic traits, a sample of U.S. college students, and a subliminal priming technique, however, suggests the reliability and generalizability of our findings.

Although we did not predict that automatic activation would be specific to negative stereotypic associations, these findings are in accordance with previous results by Wittenbrink and his colleagues (Wittenbrink et al., 1997: Wittenbrink, Judd, Park, & Stone, 1998). One possible explanation for this particular pattern could be methodological. The negative stereotypic words used in the present study might have been more strongly associated with the Black category representations than the positive stereotypic words. This explanation, however, is inconsistent with the results of the pilot test utilized for selecting the stimuli. Mean differences in the attribution of these traits to Blacks and Whites were comparable for positive (M = 12) and negative (M = 9) stereotypic traits, $t(8) = 1.36$, $p = .21$.

An alternative explanation for the differential valence effects is that self-reports may systematically underestimate the extent to which negative stereotypic words are identified with Blacks (Dovidio & Fazio, 1991). Bogus pipeline procedures (e.g. Sigall & Page, 1971; Roese & Jamieson, 1993), for example, demonstrate that people often avoid making negative stereotypic attributions (which they actually believe) in order to appear unbiased. Thus, it is possible that negative stereotypic associations were substantially stronger than our pilot test revealed. In addition, to the extent that automatic activation is the result of cultural stereotypes related to Blacks, which have traditionally been largely negative (Devine, 1989), it is likely that only negative stereotypic associations are so strongly associated as to permit automatic activation (Fazio et

2 The authors would like to thank an anonymous reviewer for this helpful comment.

al., 1986). Thus, despite the apparent decrease in negative stereotyping and the increase in expression of positive stereotypes of Blacks over the past sixty years (Devine & Elliot, 1995; Dovidio & Fazio, 1991), negative stereotypes still represent the primary automatic response of Whites to Blacks. Nevertheless, the fact that some studies (e.g., Kawakami et al., 1998) find automatic stereotype activation across positive and negative qualities suggests that further attention should be devoted to this issue – focussing both on how the stimuli are developed (e.g., factors that might influence the validity of respondents' ratings in the pilot work that forms the basis for trait selection) and on similarities and differences in the nature of the techniques used for assessing implicit stereotypes and attitudes.

Conclusions

Although the Stroop task has in the past been utilized to examine a number of social cognitive processes – the chronic accessibility of personality constructs (Bargh & Pratto, 1986), self-schema concepts (Higgins et al., 1988; Segal, 1988), and automatic vigilance for negative social information (Pratto & John, 1991) – stereotype activation has not been one of these processes (Horowitz, Malle, Knutson, Dryer, & Nelson, 1994; for a study related to stereotype application and the Stroop task see Locke, MacLeod, & Walker, 1994). The significance of the primed Stroop task for the present objective is its ability to track the stereotyping process. Specifically, it is able to provide information on the automaticity of categorization effects on stereotyping.

Overall, the findings of the present research reveal further evidence of implicit stereotypes and attitudes generally (Greenwald & Banaji, 1995) and converging evidence with respect to implicit racial stereotypes (e.g., Kawakami et al., 1998; Wittenbrink et al., 1997) and prejudice (Dovidio et al., 1997; Fazio et al., 1995) in particular. Bargh and Chartrand (1999) have recently argued that implicit evaluations, beliefs, and motives play fundamental roles in "psychological life": "automatic evaluation of the environment is a pervasive and continuous activity that individuals do not intend to engage in and of which they are largely unaware. It appears to have real and functional consequences … [that] keep us in touch with the realities of our world in a way that bypasses the limitations of conscious self-regulation capabilities" (pp. 475–476). With respect to race, specifically, research in this area is moving beyond the mere documentation of implicit stereotypes and attitudes to consider how these measures may predict behaviors and opinions. Dovidio et al. (1997), for example, have demonstrated that automatically activated evaluative racial biases predict spontaneous behaviors, such as nonverbal behavior in ways independent of explicit, self-reported attitudes (see also Dovidio & Fazio, 1991; Fazio et al., 1995). Because implicit attitudes and stereotypes are often disassociated from self-reported attitudes and beliefs, their manifestation in nonverbal behaviors may undermine more conscious and egalitarian expressions in interracial interactions. Thus, through a better understanding of these basic processes, we can better develop methods to combat unintended negative influences and subtle forms of prejudice and discrimination.

References

Allport, G. (1954). The nature of prejudice. Reading, MA: Addison-Wesley.

Anderson, S., Spielman, L. & Bargh, J. (1992). Future-event schemas and certainty about the future: Automaticity in depressives' future-event predictions. Journal of Personality and Social Psychology, 63, 711–723.

Baker, S. & Devine, P. (1988, April). Faces as primes for stereotype activation. Paper presented at the 60th annual meeting of the Midwestern Psychological Association, Chicago.

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

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

Banaji, M., Hardin, C. & Rothman, A. (1993). Implicit stereotyping in person judgment. Journal of Personality and Social Psychology, 65, 272–281.

Bargh, J. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. Uleman & J. Bargh (Eds.), Unintended thought (pp. 75–123). New York: Guilford.

Bargh, J. (1994). The four horsemen of automaticity: Awareness, intention, efficiency and control in social cognition. In R. Wyer Jr. & T. Srull (Eds.), The handbook of social cognition (2nd ed., pp. 1–40). Hillsdale, NJ: Erlbaum.

Bargh, J. (1996). Automaticity in social psychology. In E. Higgins & A. Kruglanski (Eds.), Social psychology: Handbook of basic principles (pp. 169–183). New York: Guilford Press.

Bargh, J. (1999). The cognitive monster. In S. Chaiken & Y. Trope (Eds.), Dual process theories in social psychology. New York: Guilford Press.

Bargh, J., Chaiken, S., Govender, R. & Pratto, F. (1992). The generality of the automatic attitude activation effect. Journal of Personality and Social Psychology, 62, 893–912.

Bargh, J., Chaiken, S., Raymond, P. & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation and a pronunciation task. Journal of Experimental Social Psychology, 32, 104–128.

Bargh, J. & Chartrand, T., (1999). The unbearable automaticity of being. American Psychologist, 54, 462–479.

Bargh, J. & Pratto, F. (1986). Individual construct accessibility and perceptual selection. Journal of Experimental Social Psychology, 22, 293–311.

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

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

Devine, P. & Elliot, A. (1995). Are racial stereotypes really fading? The Princeton Trilogy revisited. Personality and Social Psychology Bulletin, 21, 1139–1150.

Dijksterhuis, A. & van Knippenberg, A. (1996). The knife that cuts both ways: Facilitated and inhibited access to traits as a result of stereotype-activation. Journal of Experimental Social Psychology, 32, 271–288.

Dovidio, J., Evans, N. & Tyler, R. (1986). Racial stereotypes: The contents of their cognitive representation. Journal of Experimental Social Psychology, 22, 22–37.

Dovidio, J. & Fazio, R. (1991). New technologies for the direct and indirect assessment of attitudes. In J. Tanur (Ed.), Questions about survey questions: Meaning, memory, attitudes and social interaction (pp. 204–237). New York: Russell Sage Foundation.

Dovidio, J., 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.

Fazio, R., Jackson, J., Dunton, B. & Williams, C. (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., Sanbonmatsu, D., Powell, M. & Kardes, F. (1986). On the automatic activation of attitudes. Journal of Personality and Social Psychology, 50, 229–238.

Fiske, S. (1989). Examining the role of intent: Toward understanding its role in stereotyping and prejudice. In J. Uleman & J. Bargh (Eds.), Unintended thought (pp. 75–123). New York: Guilford.

Gaertner, S. & McLaughlin, J. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. Social Psychology Quarterly, 46, 3–30.

Greenwald, A. & Banaji, M. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. Psychological Review, 102, 4–27.

Hamilton, D. & Trolier, T. (1986). Stereotypes and stereotyping: An overview of the cognitive approach. In J. Dovidio & S. Gaertner (Eds.), Prejudice, discrimination and racism (pp. 1–34). Orlando, FL: Academic Press.

Hense, R., Penner, L. & Nelson, D. (1995). Implicit memory for age stereotypes. Social Cognition, 13, 399–415.

Higgins, E., van Hook, E. & Dorfman, D. (1988). Do self-attributes form a cognitive structure? Social Cognition, 6, 177–207.

Horowitz, L., Malle, B., Knutson, B., Dyer, T. & Nelson, P. (1994). The schema concept in personality, social psychology, and psychopathology: An integrative model. Unpublished manuscript.

Jonides, J. & Mack, R. (1984). On the cost and benefit of cost and benefit. Psychological Bulletin, 96, 29–44.

Judd, C., Park, B., Ryan, C., 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, D. & Braly, K. (1933). Racial stereotypes in one hundred college students. Journal of Abnormal and Social Psychology, 28, 280–290.

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

Kawakami, K., Dovidio, J., Moll, J., Hermsen, S. & Russin, A. (in press). Just say no (to stereotyping): Effects of training in negation of stereotypic associations on stereotype activation. Journal of Personality and Social Psychology.

Kihlstrom, J. (1990). The psychological unconscious. In L. A. Pervin (Ed.), Handbook of personality theory and research: Theory and research (pp. 445–464). New York: Guilford.

Klein, G. (1964). Semantic power measured through the interference of words with color-naming. American Journal of Psychology, 77, 576–588.

Kucera, H. & Francis, W. (1967). Computational analysis of present day American English. Providence, RI: Brown University Press.

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

Locke, V., MacLeod, C. & Walker, I. (1994). Automatic and controlled activation of stereotypes: Individual differences associated with prejudice. British Journal of Social Psychology, 33, 29–45.

Lombardi, W., Higgins, T. & Bargh, J. (1987). The role of consciousness in priming effects on categorization: Assimilation versus contrast as a function of the priming task. Journal of Personality and Social Psychology, 13, 411–428.

MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. Psychological Bulletin, 109, 163–203.

Macrae, C., Bodenhausen, G. & Milne, A. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. Journal of Personality and Social Psychology, 69, 397–407.

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

Martin, L., Seta, J. & Crelia, R. (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.

McCauley, C. & Stitt, C. (1978). An individual and quantitative measure of stereotypes. Journal of Personality and Social Psychology, 36, 929–940.

McCauley, C., Stitt, C. & Segal, M. (1980). Stereotyping: From prejudice to prediction. Psychological Bulletin, 87, 195–208.

Meyer, D. & Schvaneveldt, R. (1971). Facilitation in recognizing pairs of words: Evidence of dependence between re-

trieval operations. Journal of Experimental Psychology, 90, 227–234.

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

Park, B. & Hastie, R. (1987). The perception of variability in category development: Instance- versus abstraction-based stereotypes. Journal of Personality and Social Psychology, 53, 621–635.

Perdue, C. & Gurtman, M. (1990). Evidence for the automaticity of ageism. Journal of Experimental Social Psychology, 28, 199–216.

Posner, M. & Snyder, C. (1975). Attention and cognitive control. In R.L. Solo (Ed.), Information processing and cognition: The Loyola symposium, (pp. 55–82). Hillsdale, N.J.: Erlbaum.

Pratto, F. & John, O. (1991). Automatic vigilance: The attention-grabbing power of negative social information. Journal of Personality and Social Psychology, 61, 380–391.

Roese, N. & Jamieson, D. (1993), Twenty years of bogus pipeline research: A critical review and meta-analysis. Psychological Bulletin, 114, 363–375.

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

Schneider, D. (1991). Social cognition. Annual Review of Psychology, 42, 527–561.

Segal, Z. (1988). Appraisal of the self-schema construct in cognitive models of depression. Psychological Bulletin, 103, 147–162.

Sigall, H. & Page, R. (1971). Current stereotypes: A little fading, a little faking. Journal of Personality and Social Psychology, 18, 247–255.

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.

Stroop, J. (1935). Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 18, 643–662.

Tajfel, H. & Turner, J. (1986). The social identity of intergroup behavior. In S. Worchel & W. Austin (Eds.), Psychology of intergroup behavior (2nd ed., pp. 7–24). Chicago: Nelson-Hall.

Taylor, S. (1981). A cognitive approach to stereotyping. In D. Hamilton (Ed.), Cognitive processes in stereotyping and intergroup behavior (pp. 83–114). Hillsdale, NJ: Erlbaum.

Turner, J., Hogg, M., Oakes, P., Reicher, S. & Wetherell, M. (1987). Rediscovering the social group: A self-categorization theory. Oxford, England: Blackwell.

Warren, R. (1972). Stimulus encoding and memory. Journal of Experimental Psychology, 94, 90–100.

Warren, R. (1974). Association, directionality, and stimulus encoding. Journal of Experimental Psychology, 102, 151–158.

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.

Wittenbrink, B., Judd, C., Park, B. & Stone, M. (1998). The valenced content of racial stereotypes: Assessing prejudice at implicit and explicit levels. Unpublished manuscript.

Zarate, M. & Smith, E. (1990). Person categorization and stereotyping. Social Cognition, 8, 161–185.

Kerry Kawakami

Department of Social Psychology
University of Nijmegen
Postbus 9104
6500 HE Nijmegen
tel.: 011 31 (24) 361 2676
E-mail: kawakami@psych.kun.nl