Personality and Social Psychology Review
In Press

The Malleability of Automatic Stereotypes and Prejudice

Irene V. Blair
University of Colorado at Boulder

The present article reviews evidence for the malleability of automatic stereotypes and prejudice. In contrast to assumptions that such responses are fixed and inescapable, it is shown that automatic stereotypes and prejudice are influenced by, (a) self and social motives, (b) specific strategies, (c) the perceiver's focus of attention, and (d) the configuration of stimulus cues. In addition, group members' individual characteristics are shown to influence the extent to which (global) stereotypes and prejudice are automatically activated. This evidence has significant implications for conceptions of automaticity, models of stereotyping and prejudice, and attitude representation. The review concludes with the description of an initial model of early social information processing.

> Given a thimbleful of facts we rush to make generalizations as large as a tub…. Life is short, and the demands upon us for practical adjustments so great, that we cannot let our ignorance detain us in our daily transactions (Allport, 1954, p. 9).

Psychologists have long recognized the ease with which perceivers use category-based knowledge (Brewer, 1988; Fiske & Taylor, 1984; Tajfel, 1969). Although stereotypes and prejudice may be socially abhorrent, they appear to be cognitively beneficial by allowing perceivers to process information and make judgments efficiently (Bodenhausen, 1990; Macrae, Milne, & Bodenhausen, 1994; Macrae, Stangor, & Milne, 1994; Sherman, Lee, Bessenoff, & Frost, 1998). Research conducted over the past 15 years, however, has suggested that such efficiency goes beyond the perceiver's cognitive laziness or strategic attempts to manage a complex environment. People may often not be aware of what they are doing, they might even intend to be doing something else, and perhaps worst of all, the operation of stereotypes and prejudice may be outside of their control (Bargh, 1999).

In one of the first demonstrations of the automatic operation of stereotypes, Gaertner and McLaughlin (1983) showed that participants were faster to identify paired letter strings if they were consistent rather than inconsistent with the stereotype of Black Americans (e.g., Blacks – lazy vs. Blacks – ambitious). A few years later, Devine (1989) demonstrated that even subliminally presented cues could activate stereotypes, and furthermore, those activated stereotypes could influence interpersonal judgments. Today, over one hundred studies have documented that Whites have automatic negative associations with Blacks (or other non-White groups), young adults have automatic negative associations with the elderly, and both men and women automatically associate males and females – as well as a variety of occupational and societal groups – with stereotypic attributes (for partial reviews, see Banaji, 2001; Bargh, 1999; Greenwald & Banaji, 1995). The relative ease of revealing these automatic associations, the strength and prevalence of the effects, and growing evidence that such associations predict and influence behavior have had a profound influence on how researchers view stereotyping and prejudice.

One of the most significant consequences has been the idea that because they are automatic, these early biases are inevitable and their influence nearly impossible to avoid (Bargh, 1999; Devine, 1989). As noted by Fiske (1998, p. 363), "According to current wisdom, automatic categorization and automatic associations to categories are the major culprits in the endurance of bias." The primary reason that automatic

Work on this article was supported by NIH grant MH 63372-01. I thank Bernd Wittenbrink, Nilanjana Dasgupta, Anthony Greenwald, and researchers at the University of Colorado Stereotyping and Prejudice (CUSP) Laboratory for their insightful comments on an earlier draft of the article. I am especially grateful to the many researchers who shared their unpublished research with me.

All correspondence concerning this article should be addressed to Irene V. Blair, University of Colorado, Department of Psychology, Boulder, CO 80309-0345. Electronic mail may be sent to irene.blair@colorado.edu.

stereotypes and prejudice are believed to have such power is the assumption that automatic processes are inflexible and impervious to the perceiver's intentions and goals. In the words of several prominent researchers: "A crucial component of automatic processes is their inescapability; they occur despite deliberate attempts to bypass or ignore them" (Devine, 1989, p. 6); "Automatic processes are effortless and are initiated spontaneously and inescapably upon the individual's encountering appropriate stimulus conditions" (Dovidio & Fazio, 1992, p. 215); "An automatic mental phenomenon occurs reflexively whenever certain triggering conditions are in place…. it does not matter where the current focus of conscious attention is, what the individual was recently thinking, or what the individual's current intentions or goals are" (Bargh, 1997, p. 3). Although most theorists make it clear that perceivers do have ultimate control over their judgments and behavior, the cognitive and motivational conditions necessary for such control make it likely that stereotypes and prejudice will more often prevail (Bargh, 1999; Brewer, 1988; Devine, 1989; Fiske & Neuberg, 1990; Monteith, 1993).

The assumption that automatic processes are inflexible and uncontrollable has also lead to the conclusion that such processes reflect people's true attitudes: Attitudes that are deep-seated, resistant to external pressures and strategic processes, and stable across time and situation (Bargh, 1999; Dovidio & Fazio, 1992; Fazio, Jackson, Dunton, & Williams, 1995). For many years, researchers have struggled with the problem that people can easily mis-report their attitudes, and that they are especially likely to do so in socially sensitive domains (Crosby, Bromley, & Saxe, 1980; Webb, Campbell, Schwartz, & Sechrest, 1966). If automatic processes are impervious to short-term manipulation, then the attitudes they reveal are presumably undistorted by situational and normative pressures (Dovidio & Fazio, 1992; Fazio et al., 1995).

In summation, there is impressive evidence for the automatic operation of stereotypes and prejudice. Based on assumptions about the inflexible and unavoidable nature of automatic processes, that evidence has been used to suggest that the early influences of stereotypes and prejudice are inescapable, and as a consequence, biased judgment and behavior are very difficult to avoid. In addition, the belief that automatic associations are deep-seated and impervious to strategic efforts has contributed to the idea that such associations represent people's true attitudes.

Goals and Definitions

The purpose of the present review is to report on the many experiments that have directly tested the assumption that automatic stereotypes and prejudice are inflexible and impossible to control. In contrast to that assumption, the evidence will show that automatic processes can be influenced by the perceiver's motives and goals, and aspects of the situation. The experiments that are included in this review were conducted in the service of many different goals, making the studies heterogeneous in approach and method. It is that diversity that motivates the present review. Evidence for the malleability of automatic stereotypes and prejudice should not be accepted lightly, in consideration of the serious theoretical and practical implications that such evidence would have (Bargh, 1999). Moreover, any single test is sure to have some flaws or can be disregarded as a special case. Many and diverse tests, on the other hand, ought to be more convincing. More importantly, a review can highlight important issues and areas of interest for future research. At the conclusion of the review, a model of early social perception is presented as an initial attempt to incorporate the research findings into an understanding of stereotyping and prejudice (see Figure 1).

To facilitate the review, it is important to provide some initial definitions, with the most important being a definition for automaticity. Over the years, researchers have suggested different criteria to define an automatic process, with the most common being the absence of awareness or attention, a lack of intention, and uncontrollability (for reviews, see Bargh, 1989, 1994; Kihlstrom, 1990). However, it is rare for any psychological process to meet all of those criteria; it is not always clear which criteria have been met by any particular process; and furthermore, no clear differences have been shown for processes that meet different criteria (e.g., lack of awareness vs. lack of attention). Thus, a hard and fast definition is impractical. The approach that will be taken here is to "count" a stereotype or prejudice effect as automatic if that is how it has been consensually viewed by researchers. At a minimum, that usually means that the operation of the stereotype or prejudice can be presumed to be unintended by the research participants (i.e., not deliberate), either because they are unaware of certain critical aspects of the procedure or because they are operating under conditions that make it difficult to deliberately base responses on specific beliefs or evaluations (c.f., Kihlstrom, 1990).

For example, an automatic stereotype is assumed to be operating if participants provide more stereotypic Asian word completions in the presence of an Asian than a White assistant because the participants are unlikely to be aware that the assistant's ethnicity had any influence on their responses (Gilbert & Hixon, 1991). Similarly, in a sequential priming task the influence of a prime word on participants' speed of responding to a subsequent target word is presumed to

be unintended, either because the prime is presented subliminally, participants are unaware that millisecond response timing is the outcome of interest, or because the short amount of time between the presentation of the prime and the onset of the target (< 500 ms. stimulus onset asynchrony [SOA]) makes it difficult for participants to deliberately use the prime to respond to the target. Thus, faster responses to a target that is preceded by a stereotypically consistent prime compared to an inconsistent prime (e.g., Black – lazy vs. White – lazy) is presumed to reflect an automatic stereotype (Banaji & Hardin, 1996; Blair & Banaji, 1996; Fazio et al., 1995; Wittenbrink, Judd, & Park, 1997).

To ensure that a consensual definition of automaticity has been met, all of the effects included in the present review were measured in a manner that is conventionally accepted to reveal an automatic process, such as the Implicit Association Test (IAT, Greenwald, McGhee, & Schwartz, 1998), lexical decision task (LDT, Macrae, Bodenhausen, Milne, & Jetten, 1994; Macrae, Bodenhausen, & Milne, 1995), sequential priming task (Banaji & Hardin, 1996; Blair & Banaji, 1996; Fazio et al., 1995; Wittenbrink et al., 1997), or word completions following unobtrusive priming (Gilbert & Hixon, 1991; Spencer, Fein, Wolfe, Fong, & Dunn, 1998). To keep the paper focused, lengthy descriptions of specific procedures are avoided, with the assumption that most readers are familiar with the measures. Brief descriptions are provided in the Appendix to establish common terminology and as a reminder of what the measures involve. Readers are referred to the original sources for more details.

The second definitional issue that must be addressed is the conventional distinction between stereotypes and prejudice, with the former referring to the beliefs (semantic associations) people have about social groups and the latter referring to their evaluations of groups. Although it is possible and often desirable to make such a distinction, it is less important for the present review. Both stereotypes and prejudice have been shown to operate automatically, and such automaticity has been assumed, for both, to indicate high resistance to change and invariance across context and strategy. Thus, while the distinction between semantic and evaluative associations is maintained in discussing individual studies, the conclusions drawn from those studies are applied more generally.

Finally, the scope of the review must be delineated. As discussed above, automatic stereotypes and prejudice have been assumed to be impervious to situational manipulations and strategic processes. The most powerful tests of that assumption, and the focus of the present review, are short-term (single session) experimental manipulations of those variables.

Evidence for the Malleability of Automatic Stereotypes and Prejudice

Just a few years ago, there were only a handful of studies on the malleability of automatic stereotypes and prejudice (for a review, see Bargh, 1999). The situation today is quite different, with nearly 50 investigations of their flexibility and responsiveness to a wide range of strategic, social, and contextual influences. The present review is organized around five general classes of moderators: (1) Self and social motives, (2) specific strategies to counter stereotypes, (3) focus of attention, (4) the configuration of stimulus cues, and (5) characteristics of individual category members. The first three classes focus on manipulations of the perceiver's motivations, goals, and strategies while in the testing situation. The last two classes of effects focus on situational manipulations that are external to the perceiver, such as changes in the context surrounding the stimulus or variations in group members' attributes. Table 1 provides a catalog of the studies in each section.

Self and Social Motives

Self-image motives. Preservation of one's self-image is a powerful motivator, with a variety of consequences for judgment and behavior (Greenwald, 1980). Several years ago, Fein and Spencer (1997) argued that a self-image threat can motivate people to invoke negative stereotypes of others as a means of feeling better about oneself. Spencer et al. (1998) have demonstrated that such an effect can be found with automatic as well as more controlled responses. Specifically, participants in their studies were given either positive or negative feedback on an intelligence test, with the latter assumed to pose a significant self-image threat. Following the feedback, automatic stereotypes of Asians (Study 1) or Black Americans (Study 2) were assessed with a word fragment test, completed while the participants were kept cognitively busy. (As described below, automatic stereotypes do not typically operate under that constraint; see Gilbert & Hixon, 1990; Spencer et al, 1998, Study 2.) As expected, Spencer et al. found no evidence for automatic stereotypes among the participants who had received positive feedback. In contrast, those who had received negative feedback exhibited strong automatic stereotypes, despite the cognitive constraint.

In related work, Sinclair and Kunda (1999) argued that people can be motivated to inhibit as well as magnify stereotypes, depending on whether the stereotypes harm or benefit the self. In particular, stereotypes harm the self if they serve to discredit a person who affirms one's self image; by the same token, they can benefit the self if the person being

discredited threatens one's self-image. Importantly, Sinclair and Kunda proposed that such motivated processing ought to be evident in the automatic operation of stereotypes. Consistent with their argument, automatic stereotypes of Blacks were found to be weaker after a Black supervisor had delivered a positive evaluation of the participants, compared to a control condition. The stereotypes were stronger, however, when the evaluation was negative. As an indication that the effects were indeed motivated by self-concern, Sinclair and Kunda (Study 2) showed that the supervisor's evaluation had no effect on participants' automatic stereotypes when the participants merely observed the evaluation being given to someone else. In two additional studies, self-image motives were also shown to determine which automatic stereotypes were strengthened or inhibited. A Black doctor, for example, can be alternatively associated with negative race stereotypes or positive professional stereotypes. As evidence for the flexibility of automatic stereotypes, Sinclair and Kunda found that participants who had received negative feedback from such an individual exhibited more automatic race stereotypes but less automatic doctor stereotypes, whereas participants who received positive feedback displayed the opposite pattern of responses. Such self-serving activation and inhibition of alternative stereotypes appeared to be especially pronounced for high-prejudiced participants (Study 4).

Social motives. One of the classic findings in prejudice research is that responses often depend on who is asking the questions. White participants are known to express less prejudice to a Black interviewer, for example, than to a White interviewer (Kinder & Sanders, 1996; Schuman, Steeh, Bobo, & Krysan, 1997). Such results are often interpreted in terms of social desirability and the participants' motivation to hide their prejudice. If they are successful, the response is characterized as an invalid estimate of attitudes and the development of a new measure is called for (e.g., Fazio et al., 1995; McConahay, Hardee, & Batts, 1981).

A different interpretation has been promoted by Lowery, Hardin, and Sinclair (2001), who suggest that such changes may reflect sincere attempts by the participants to achieve common ground with the interviewer, by adjusting their perspective and communicative attempts to more closely match the interviewer's presumed knowledge and attitudes – a process Lowery et al. call "social tuning". If such is the case, even automatic responses ought to reveal sensitivity to the ethnicity (and presumed attitudes) of the other person. Lowery et al. tested this hypothesis across three studies in which the participants completed a measure of automatic race prejudice (IAT or a subliminal sequential priming task) in the presence of a Black or White experimenter. As predicted, White participants exhibited significantly less automatic negativity toward Blacks in the presence of a Black experimenter than in the presence of a White experimenter. Of some interest was the finding that Asian participants did not show the same social tuning effects, and instead produced equivalent levels of automatic prejudice regardless of the experimenter's race. Lowery et al. interpreted that result as due to the fact that racism against Black Americans is largely viewed as a "White problem" and the Asian participants may have been less motivated to adapt to the views of the experimenter. Lowery et al. (Study 3) then showed that when the Black experimenter explicitly instructed the participants to "be the least prejudiced you can", both White and Asian participants produced lower levels of automatic prejudice.

As further evidence that people are sensitive to social roles and expectations, even at an automatic level of processing, Richeson and Ambady (2001) demonstrated that the relative status of the perceiver can moderate the extent to which he or she exhibits automatic prejudice. In their study, White participants learned that they would be working with a Black American student. Some participants were told that their goal during the upcoming interaction would be to evaluate their partner's performance (superior role); other participants were instructed to get along with their partner (equal-status role); and still other participants were told to manage the impression they make on their partner who would be evaluating them (subordinate role). The participants then completed an IAT measure of automatic race prejudice. Richeson and Ambady found that the participants assigned to the superior role produced a higher level of automatic prejudice than the participants assigned to the equal-status role, and the participants assigned to the subordinate role exhibited the least amount of automatic prejudice.

As a final example of the influence of social relationships on automatic responses, Sechrist and Stangor (2001) showed that perceived consensus can alter one's automatic stereotypes. In their study, the participants completed a sequential priming task after learning that either many or few of their peers agreed with their racial stereotypes. As predicted, those who believed they were "out of step" with their peers exhibited significantly lower levels of automatic race stereotypes, compared to participants who believed that their racial beliefs were congruent with those of their peers.

Summary. The evidence reviewed in this section shows that highly motivated individuals can modify the automatic operation of stereotypes and prejudice.

People whose self-image has been threatened may automatically activate negative stereotypes to make themselves look better or discredit a disliked evaluator (Sinclair & Kunda, 1999; Spencer et al., 1998). Or, just as easily, they may automatically inhibit negative stereotypes and activate positive ones when doing so would be beneficial to their self-image (Sinclair & Kunda, 1999). Automatic stereotypes and prejudice are also responsive to the social demands of the situation and the nature of one's relationship with other individuals. Whites may temper their automatic prejudice during (or while anticipating) a social interaction with a Black person, especially if they are in a subordinate position (Lowery et al., 2001; Richeson & Ambady, 2001). They may also moderate automatic stereotypes if those stereotypes appear to be discrepant with social norms (Sechrist & Stangor, 2001).

One issue that this research raises is the potential influence that other motives may have on automatic processes. Fiske (1998) has enumerated several motives that may influence stereotyping: Belonging, understanding, controlling, self-enhancing, and trusting. As reviewed above, some of these motives have already proven to be important for understanding automatic stereotypes and prejudice. Additional possibilities are raised by considering the social identity motives that stem from relevant intergroup relationships. Pratto and Shih (2000) have provided some suggestive evidence in this regard. Specifically, they found that the level of automatic prejudice toward an unspecified outgroup ("them") did not vary as a function of participants' social dominance orientation (SDO). However, when the status of the ingroup was threatened (Study 2), high SDO participants exhibited significantly more automatic prejudice than those who were low in SDO.

Because the effects reviewed in this section were obtained without the provision of specialized strategies, they suggest that automatic attitudes may be quite responsive to perceivers' motives in everyday situations. This is important but it leaves open the question of how people do it. Do motivated perceivers attempt to suppress their stereotypes and prejudice, bring a different representation of the group to mind, focus their attention on different target cues, or use some other strategy? The degree to which specific strategies influence automatic stereotypes and prejudice is examined in the next two sections.

Strategies to Counter Stereotypes

In this section of the review we consider specific strategies that have been investigated as moderators of automatic stereotypes. These strategies fall roughly into two sub-categories: Stereotype suppression and the promotion of counterstereotypes.

Suppression. If told to reduce their use of stereotypes, many people would probably try to banish such thoughts from their minds. Although some researchers have found that suppression does not reduce automatic stereotypes – and in fact may amplify them (Galinsky & Moskowitz, 2000; Macrae, Bodenhausen, et al., 1994) – other studies indicate that suppression strategies can be successful. Gollwitzer and Schaal (1998) reported that a goal to suppress stereotypes works if it is accompanied by a specific implementation intention. In their study, participants were motivated to judge others in a fair and unbiased manner. In addition, some participants were instructed to form the specific implementation intention, "And whenever I see Ina, I will ignore her gender!" In a subsequent primed Stroop task, participants' automatic gender stereotypes were measured in response to the primes "Ina" and "Bea". As predicted, participants who had formed the implementation intention produced less automatic gender stereotypes in response to the specified person (e.g., Ina). The strategy was very specific in its influence, however, and had no effect on responses to the other group member.

Kawakami, Dovidio, Moll, Hermsen, and Russin (2000) also demonstrated that certain types of suppression can be successful in moderating automatic stereotypes. Specifically, participants who had been trained to say "no" to stereotypic events and "yes" to nonstereotypic events produced significantly lower levels of automatic stereotypes, compared to that produced by participants who had received no training or who had been trained to affirm the stereotypes. In addition, this "stereotype negation" training was successful in moderating automatic stereotypes of skinheads and automatic race stereotypes; its effects were observed on both a primed Stroop task and a sequential priming task; and the moderation persisted over a 24-hour period of time. The only drawback to this strategy was that it was not effective unless the participants had engaged in extensive practice.

The promotion of counterstereotypes. A different strategy aimed at reducing automatic stereotypes is to focus on counterstereotypes. That is, instead of attempting to suppress stereotypes, perceivers can work to promote opposing, counterstereotypic associations that could challenge the dominance of stereotypes in information processing.

In the first test of such a strategy, Blair and Banaji (1996) manipulated participants' expectancies during a sequential priming task. Half of the participants were instructed to expect stereotypic prime-target trials and the other participants were told to expect

counterstereotypic trials. In truth, all participants received both stereotypic and counterstereotypic trials, with the expected trial type occurring only 63% of the time. Furthermore, the participants completed two blocks of trials, one block with a 350-ms. SOA and one block with a 2000-ms. SOA. Responses on the former block of trials are of particular interest because the short amount of time provides greater certainty that the outcome is based on an automatic process (see above).

The results of this test showed that the expectancy strategies had a significant influence on the participants' automatic stereotypes. When the SOA was only 350-ms., the counterstereotype expectancy produced significantly lower levels of automatic stereotypes than the stereotype expectancy. Not surprisingly, this difference was substantially larger with the 2000-ms SOA, which allowed for the greater influence of controlled processes. A follow-up study examined the counterstereotype expectancy at both 250-ms. and 2000-ms. SOAs, and showed that (a) once again the expectancy was much more effective with the 2000-ms SOA, and (b) even with a 250-ms. SOA, the expectancy prevented the participants from producing a significant level of automatic stereotypes. Taken together, the results of these two experiments suggest that people may be able to moderate automatic stereotypes by intentionally activating counterstereotypes.

Blair, Ma, and Lenton (2001) have recently examined mental imagery as another strategy to promote counterstereotypes. Prior research has shown that mental imagery increases the accessibility of the imagined event (e.g., Carroll, 1978; Gregory, Cialdini, & Carpenter, 1982). By the same token, Blair et al. argued that counterstereotypic mental imagery ought to increase the accessibility of counterstereotypic associations, and thereby decrease automatic stereotypes. In four separate tests, the participants were asked to spend approximately five minutes creating a mental image of a (counterstereotypic) strong woman and then complete a measure of their automatic gender stereotypes. In each test, the participants who had engaged in the counterstereotypic mental imagery produced substantially weaker automatic stereotypes, compared to participants who, (a) engaged in neutral mental imagery, (b) did not engage in any imagery, (c) imagined a weak woman, (d) imagined a strong man, or (e) attempted to suppress their stereotypes during the task. Moreover, the moderating influence of the counterstereotypic mental imagery was demonstrated through response times on the IAT, word detection sensitivity (d') on the GNAT (for female participants), and recognition false-alarms in a false memory induction procedure. The consistency of the effects and the variety of outcomes that were moderated suggest that mental imagery can have a powerful influence on automatic processes.

Taking a somewhat different approach, Dasgupta and Greenwald (2001) showed that exposure to counterstereotypic group members can also alter automatic prejudice. In their research, participants were exposed either to admired Black Americans and disliked White Americans (e.g., Bill Cosby and Timothy McVeigh), disliked Black Americans and admired White Americans (e.g., O.J. Simpson and John F. Kennedy), or nonracial stimuli (control). Following that exposure, the participants completed an IAT measure of automatic racial prejudice both immediately and 24 hours later. Dasgupta and Greenwald found that participants exposed to positive Black group members produced less automatic prejudice toward Blacks, compared to participants who had been exposed to negative group members or to nonracial stimuli. Moreover, this moderation continued to be significant when tested 24 hours later. A second study replicated the effect for the moderation of automatic age prejudice.

Finally, it is worth noting that the moderation of automatic group attitudes is not restricted to laboratory manipulations. Specifically, Rudman, Ashmore, and Gary (2001) showed that participating in a semester-long diversity course can alter students' automatic associations. In two quasi-experimental studies, students enrolled in a "prejudice and conflict" seminar exhibited significant reductions across the semester in their automatic stereotypes and prejudice toward Blacks, whereas students enrolled in control courses (e.g., research methods) showed no such reduction.

Summary. In discussing evidence for automatic stereotypes, there have been suggestions that perceivers' specific goals and strategies have no influence on such processes (Bargh, 1999; Devine, 1989). The studies reviewed in this section provide strong evidence that such factors are not so inconsequential. Attempting to suppress a stereotype, expecting counterstereotypic events, or focusing on counterstereotypic group members have all been shown to have a significant influence on automatic stereotypes. Moreover, there is some evidence that such strategies can have longer-term effects (Dasgupta & Greenwald, 2001; Kawakami et al., 2000)

Although the evidence is compelling with regard to the possibility of moderating automatic stereotypes, the likelihood of such moderation in everyday social encounters is not yet known. For example, suppression is an highly intuitive control strategy, yet its success depends on the employment of a specific implementation intention (Gollwitzer & Schall, 1998) or extensive practice (Kawakami et al., 2000). In addition, other research has shown that suppression

(presumably without those additional features) can backfire and actually magnify automatic stereotypes (Galinsky & Moskowitz, 2000; Macrae et al., 1994). It is also important to acknowledge that any explicit strategy takes time and motivation to implement (Blair & Banaji, 1996), and its effects may not generalize beyond the specific context within which it is used. As discussed above, Gollwitzer and Schall (1998) found that suppression coupled with an implementation intention reduced automatic stereotypes associated with the targeted person but not another group member. And in a replication of the Blair et al. (2001) effects, Carpenter and Banaji (2001) found that counterstereotypic mental imagery moderated participants' automatic stereotypes but not their automatic evaluations of women.

It remains for future research to determine what strategies under what conditions are the most effective in moderating automatic stereotypes. An answer will surely depend on additional considerations, such as individual differences in motivation and skill. For example, Moskowitz, Gollwitzer, Wasel, & Schaal (1999) found that people who have a chronic goal of fairness exhibited less automatic stereotypes than non-chronics. Because the chronically motivated individuals had as much explicit knowledge about stereotypes as the non-chronics, Moskowitz et al. suggested that the former group may be better at automatically suppressing their stereotypes. Wasel and Gollwitzer (1997) also found that chronically motivated perceivers were able to significantly reduce automatic stereotypes when the stimuli were consciously perceptible (200-ms presentation), as they were in the Moskowitz et al. study. However, such moderation did not occur when the stimuli were subliminal (33-ms. presentation) and the participants were unaware that stereotypes might be operating.

It is probably the case that strategic efforts to moderate automatic stereotypes and prejudice require some awareness, motivation, skill, and resources to be successful (Bargh, 1992, 1999) – although this may be less true with practice (Kawakami et al., 2000; Monteith, 1993). Nonetheless, the evidence shows that just because a process is automatic, it cannot be assumed to be impervious to perceivers' goals and strategies.

Focus of Attention

The amount and type of attention a perceiver pays to another person has long been viewed as critical for predicting the impact of stereotypes on judgment and behavior. A well-supported finding is that stereotypes dominate unless the perceiver spends more time learning about the person's unique attributes (for a review, see Fiske, 1998). Much less has been said about the role of attention in automatic stereotypes and prejudice, with the assumption that attention is irrelevant for such processes (Bargh, 1997). Recent research, however, shows that perceivers' focus of attention has a significant influence on the automatic operation of stereotypes and prejudice.

In one of the first investigations of such an effect, Gilbert and Hixon (1991) examined the impact of cognitive "busyness" (attentional load) on automatic stereotypes, as measured by a word completion task in the presence of an Asian versus White assistant. Attentional load was manipulated by instructing some of the participants to rehearse a random string of digits while they attempted to complete the word fragments. Of great interest was the finding that those participants produced fewer Asian stereotypes in the presence of the Asian assistant, compared to participants who were not so cognitively busy (Gilbert & Hixon, 1991). Spencer et al. (1998, Study 2) have since replicated that effect for stereotypes of Black Americans. Specifically, participants who were cognitively busy produced fewer stereotypic word completions following subliminal exposure to Black American faces, compared to participants who did not have the added task of rehearsing digits.

Macrae, Bodenhausen and colleagues have also demonstrated that perceivers' focus of attention can moderate automatic stereotypes. In one set of studies, Macrae, Bodenhausen, Milne, Thorn, and Castelli (1997) investigated attentional focus in a sequential priming task, in which the primes were pictures of common inanimate objects or women, and the targets were stereotypic or counterstereotypic traits associated with women. In one condition, the participants were asked to decide whether each picture was of an animate object, whereas the participants in another condition were asked to decide whether a white dot was present. Although the tasks may not seem very different, they effectively manipulated attention to the (female) features of the photographed women, because such features would have only been useful for participants judging the animateness of the objects. Macrae et al. predicted and found that the pictures of women facilitated responses to the stereotypic traits only in the animate-nonanimate judgment task; automatic stereotypes did not appear when the participants' attention was focused on detecting the white dot.

In a second set of studies, Macrae, Bodenhausen, Milne, and Calvini (1999) examined whether social cues presented outside the perceiver's "attentional spotlight" would have less automatic influence than cues presented inside the spotlight. Using a flanker task, Macrae et al. instructed the participants to make a simple judgment about male and female target names that were presented in the center of the computer screen. On each trial, the target name was presented

with additional stimuli (flankers) that were supposed to be irrelevant to the task, but on some trials were actually names that were inconsistent with the gender of the target. Importantly, the flankers were presented either very close to the target or slightly further away. Macrae et al. found that the gender-inconsistent flankers only interfered with judgment (evidence for an automatic category association) if they were presented very close to the target. The more distant flankers had no such influence. In a subsequent task, Macrae et al. showed that the participants were able to categorize both the near and distant flankers faster than words they hadn't seen before, indicating that the participants had processed the flankers and the moderation was not the result of a failure to process the more distant flankers.

Other research has shown that attentional focus may also determine which category associations are automatically activated when the stimulus can be viewed in multiple ways. For example, Mitchell, Nosek, and Banaji (2001) examined changes in automatic evaluation depending on which social category was at the focus of attention. In their first study, Mitchell et al. had participants complete two IATs with the same stimuli appearing in both: Names that varied simultaneously in gender and race, and pleasant and unpleasant words. In one of the tests the participants were asked to categorize the names by gender (male vs. female), and in the other test they were asked to categorize them by race (Black vs. White). As predicted, automatic evaluations of Black females and White males were radically different, depending on whether the participants were attending to race or gender. When gender was salient, Black females were evaluated positively and White males were evaluated negatively; when race was salient, Black females were evaluated negatively whereas White males were evaluated positively. Because evaluation is consistent on both dimensions for White females (positive) and for Black males (negative), the same automatic attitudes were expressed toward those targets regardless of which category was salient. Mitchell et al. obtained similar results with IATs assessing automatic attitudes toward well-liked Black athletes (e.g., Michael Jordan) and disliked White politicians (e.g., Newt Gingrich): When the participants were focused on race, the Black athletes were automatically evaluated more negatively than the White politicians; when attention switched to occupation, the Black athletes were automatically evaluated more positively than the White politicians.

Mitchell et al. provided additional evidence for changes in automatic attitudes with a series of Go/No-go Association Tasks (GNAT, Studies 3 & 4) in which attention to race or gender was manipulated more subtly. Specifically, the GNAT requires participants to respond selectively to a series of stimuli, at a speed that makes considered responses impossible. In each block of trials, responses must be made to the target stimuli (e.g., Black female names & unpleasant words), and not to any others (distracters). Mitchell et al. manipulated participants' focus of attention by changing the composition of the distracters. For example, when the target category was "Black females", the use of White female distracters focused attention on race whereas Black male distracters focused attention on gender.[1] With this manipulation, Mitchell et al. again found that Black females and White males elicited very different automatic evaluations depending on the focal dimension.

Finally, Wittenbrink, Judd, and Park (2001a) have examined how attention to evaluation versus meaning can alter participants' automatic evaluations of Black Americans. In their study, all of the participants completed a sequential priming task in which the primes were category labels ("White" and "Black") and the targets were adjectives that were stereotypic, counterstereotypic, or unrelated to the categories. Half of the participants were instructed to judge each target according to whether it was good versus bad, thereby focusing attention on evaluation. The other participants received nonword targets intermixed with the word targets, and they were instructed to judge each target according to whether it was a word or a nonword, thereby focusing attention on semantic meaning. Wittenbrink et al. found that the participants who were focused on evaluation produced higher levels of generalized automatic prejudice (i.e., associations between Black – negative and White – positive with adjectives unrelated to the category) than participants who were focused on meaning. In contrast, the evaluation and meaning foci produced equivalent levels of automatic evaluative stereotypes (i.e., associations between Black – negative and White – positive with adjectives stereotypically related to the respective categories).

Summary. The evidence reviewed in this section leaves no doubt that the perceiver's focus of attention can influence the automatic operation of stereotypes and prejudice, as well as more controlled processes. People who are preoccupied with other matters may

[1] This manipulation may also be characterized as an alteration of the context within which the target category was embedded (see below). It is included in this section because it seems likely that the manipulation had the effect of changing the participants' explicit task focus from race to gender, and it was that change that produced the corresponding effects.

not automatically activate stereotypes associated with the target's social category (Gilbert & Hixon, 1991; Macrae et al., 1997; Spencer et al., 1998); social category cues that are outside of the perceiver's focus of attention may not automatically activate category information (Macrae et al., 1999); perceivers who attend to different target identities may automatically activate different group attitudes (Mitchell et al., 2001); and perceivers who focus on meaning versus evaluation may produce different types of automatic attitudes (Wittenbrink et al., 2001a).

These findings are of great interest for at least two reasons. First, as noted earlier, the absence of attention is often considered a defining feature of automatic processes, and demonstrations that stereotypes and prejudice operate under very minimal levels of attention (e.g., subliminal cues) have indicated that those processes meet that definition of automaticity (see Bargh, 1999). It may therefore seem odd that attentional manipulations would have the significant effects shown here. Second, many studies have found that stereotyping is more likely when perceivers cannot or do not pay attention to the target (see Fiske, 1998), with the common interpretion being that, (a) stereotypes are activated regardless of the perceiver's attentional state (i.e., automatically), (b) processes that counteract stereotypes can only operate with sufficient attention, and (c) therefore a lack of attention makes it highly likely that judgment and behavior will be biased by automatic stereotypes (Fiske, 1998; Fiske & Neuberg, 1990). The evidence reviewed here, however, suggest that a lack of attention can actually have the opposite effect of reducing automatic stereotypes and thereby decreasing the likelihood of stereotypic judgments.

A comprehensive theory of the role of attention in social information processing is beyond the scope of the present review. However, a couple of observations may help make sense of the evidence. First, it is important to point out that the fact that a process does not need the perceiver's attention to operate does not necessitate the conclusion that attention cannot influence that process (Logan, 1989). There are numerous demonstrations that processes that ordinarily occur with little attention (e.g., typing) can be altered when the person chooses to pay attention and change what he or she is doing (for reviews, see Logan, 1989; Logan & Cowan, 1984). Some of the most compelling evidence comes from research on the modification of the blink reflex, a response that is so automatic it occurs even in people who are brain dead. Although the blink reflex does not require attention, it is modified by changes in attention: Greater attention to the eliciting stimulus enhances the blink reflex, whereas focusing attention away from the stimulus attenuates the response (for reviews, see Anthony, 1985; Filion, Dawson, & Schell, 1998). Thus, evidence that perceivers can influence automatic stereotypes and prejudice by deliberately altering their focus of attention should not be viewed as fundamentally inconsistent with other evidence showing that stereotypes and prejudice often operate without much attention from the perceiver.

Second, although current stereotyping models characterize attention as a late-selection process (e.g., influencing deliberate choices and decisions, but not the automatic activation of stereotypes), there are alternatives that allow attention to influence early as well as later stages of processing (e.g., Cohen, Dunbar, & McClelland, 1990; Logan, 1988; see also Bodenhausen & Macrae, 1998). Cohen et al., for example, model automatic processes within a parallel distributed processing (PDP) framework, with the strength of processing pathways the primary determinant of their automaticity. Although very strong pathways in this model produce all of the characteristic features of an automatic process, attention still modulates processing on those pathways: The more attention is focused on a pathway (due to task goals and perceiver intentions), the more responsive that pathway is to stimulation.

People cannot perceive everything in their environment, and even if they could, it makes little sense for them to process information that has no relevance to their current goals and intentions. An attentional mechanism that aids in the early selection of social information would be beneficial for efficient processing, especially when the stimulus is capable of evoking conflicting automatic attitudes (e.g., Mitchell et al., 2001).

The Configuration of Stimulus Cues

The next class of phenomena concerns the context within which social category cues are embedded. Based on Gestalt principles, early social psychologists (e.g., Heider, Lewin, and Asch) believed that human behavior could only be understood by considering the entire stimulus field (see Read, Vanman, & Miller, 1997). A particular attribute could have one meaning in one context and a different meaning in another. In his classic studies on impression formation, Asch (1946) showed that people formed very different impressions based on a list of traits (e.g., intelligent, skillful, industrious), depending on whether the trait "warm" or "cold" was included. Moreover, the exact same traits produced different impressions when they were presented in different orders. Asch argued that the effects could not be understood as the simple addition or subtraction of attributes. Rather, "a given quality derives its full concrete content from its place within the system formed by the relations of the qualities….

Identical qualities in different structures may cease to be identical: the vectors out of which they grow may alter, with the consequence that their very content undergoes radical change" (p. 283). Research has begun to suggest that such fluidity in meaning may also be observed in automatic stereotypes and prejudice.[2]

For example, Macrae et al. (1995) demonstrated that a slight change in context can have a large effect on automatic stereotypes. In their studies, all of the participants were exposed to a Chinese woman, and automatic stereotypes of both Chinese and women were subsequently measured with a LDT. In one condition the Chinese woman was putting on makeup, whereas in another condition she was using chopsticks. As predicted, Macrae et al. found that compared to control participants, those who saw the person put on makeup were faster to respond to traits stereotypic of women and slower to respond to traits stereotypic of Chinese, whereas the participants who saw her use chopsticks produced the opposite pattern of responses. The stimulus person was the same, yet a small change in the context produced a dramatic change in the automatic stereotypes elicited by her presence.

In another study, Macrae, Hood, Milne, Rowe, and Mason (in press) presented participants with picture primes of men and women who varied across conditions in a very subtle attribute: Whether their eye gaze was direct, averted to the side, or absent (i.e., closed eyes). Macrae et al. argued that another person's eye gaze is an important early cue in social interaction, with direct eye gaze signalling that the person has intentions in regard to oneself and is therefore a potentially important object in the environment. Macrae et al. found, in accordance with their hypothesis, that the pictured men and women produced significantly stronger automatic stereotypes if they had a direct eye-gaze, rather than averted or closed eyes.

A third example comes from recent research by Wittenbrink, Judd, and Park (2001b), who examined how different social contexts can moderate automatic evaluations of a group. In one study, participants who had seen a video clip of Black Americans at an outdoor barbecue produced significantly less automatic negativity toward the group, as measured on a subsequent IAT, compared to participants who had seen a video clip of Black Americans in a gang-related setting. The group was the same, yet it elicited very different automatic evaluations, depending on the context within which it was embedded.

In a follow-up study, Wittenbrink et al. (2001b) used a sequential priming procedure to manipulate the context and category members orthogonally. On each trial, a Black or White face prime was shown briefly, followed by a positive or negative target adjective. Just before the face appeared, a context picture was presented. On some trials the picture was of a city street, whereas on other trials the picture was of the inside of a church. In both cases, the background remained on the screen when the face appeared, making it look as if the person had appeared in the scene. With even this subtle manipulation, Wittenbrink et al. found that automatic negativity toward Black Americans was significantly higher in the street context than in the church context. Indeed, in the latter context, significant automatic positivity toward Blacks was observed. Although the context also affected attitudes toward White Americans – with more positive automatic attitudes in the church context – the effect was weaker, primarily because the street context did not produce automatic negativity toward Whites.

Summary. The studies reviewed in this section provide clear evidence that automatic responses to category cues depend on the surrounding context. The same woman can evoke different automatic stereotypes depending on whether she has a make-up brush or chopsticks in her hand (Macrae et al., 1995); the same man or woman elicits significantly stronger automatic gender stereotypes if he or she appears to be looking at the perceiver rather than looking away (Macrae, Hood, et al., in press); the same Black person can elicit different automatic attitudes when he is on a city street versus inside a church (Wittenbrink et al., 2001b). As Asch (1946) argued many years ago, these effects cannot be explained by the mere addition or subtraction of independent qualities. A pair of chopsticks is a minor cue for Chinese. If its effect was merely additive, one would expect to see a small increase in automatic stereotypes of Chinese. Instead, the presence of chopsticks produced a pattern of automatic responses that was opposite that produced by the presence of make-up, suggesting that the stimulus (a Chinese woman) was automatically viewed in a qualitatively different manner. Similarly, a city street is typically viewed more negatively than a church interior, as is a Black face compared to a White face. But each of these qualities did not independently influence the automatic process. Instead, their effects were multiplicative, producing especially strong automatic negativity when the Black face was embedded in the street context. Finally, the direction of a person's eye gaze would seem to have nothing to do with the person's stereotypicality, yet someone who is

[2] Although not of direct relevance for the present review, another example of the application of Gestalt principles to automatic processes can be found in the Greenwald et al. (in press) unified theory of social cognition, in which naturally occurring attitudes and beliefs are shown to be constrained by principles of balance and "good form".

looking directly at oneself activates significantly more automatic stereotypes than if he or she is looking away. Together these results suggest that automatic perception depends on the integration of the stimulus components, with small changes capable of producing radically different outcomes.

Characteristics of Individual Category Members

In the previous four sections, manipulations of motive, strategy, and context have been shown to alter automatic attitudes in response to the same stimuli (i.e., the group members, names, or pictures used to elicit the attitude), suggesting that an automatic attitude is not impervious to perceiver's goals and intentions nor invariant across situations. In this final section of the review, we consider research that has examined a related issue: Do all category members elicit the same or different automatic attitudes? This question is important because the research goal in measuring automatic stereotypes and prejudice is usually to determine the perceiver's attitude toward the group as a whole, with the assumption that there is a global attitude that reveals how the perceiver is likely to respond (automatically) to individual members of the group. In practice, however, many studies measure responses to specific group members (e.g., Banaji & Greenwald, 1995; Banaji, Hardin, & Rothman, 1993; Blair & Banaji, 1996; Dasgupta, McGhee, Greenwald, & Banaji, 2000; Fazio et al., 1995; Greenwald et al., 1998). With adequate stimulus sampling, researchers can determine the automatic attitude in regard to the average group member. But such sampling is rare, and test stimuli are more typically selected on the basis of their clarity and convenience. If there is a global attitude that is evoked uniformly, then such selectivity would be inconsequential. Research suggests, however, that automatic attitudes differ according to the characteristics of the individual group members.

In a series of studies, Livingston and Brewer (2002) showed that automatic prejudice depends on the appearance of the specific group members to which the perceiver responds. In their study, photographs of Black Americans served as the primes in a sequential priming task, with each photograph followed by a positive or negative target word. Importantly, the photographed Black Americans varied in their facial features, with some possessing more "Negroid" features than others (e.g., darker skin, wider nose). In contrast to assumptions of a unitary attitude, Livingston and Brewer found that the Black Americans with more Negroid features evoked more automatic prejudice than those with less Negroid features, even though all of the photographs had been rated by pretest participants as members of the same group. Moreover, only the Black Americans with strong Negroid features elicited automatic evaluations that were significantly more negative than those made in response to photographs of White Americans, suggesting that prior reports of widespread automatic prejudice toward Black Americans may not apply to all members of the group.

Macrae, Mitchell, and Pendry (in press) also used a sequential priming procedure to investigate the effect of variations in category instances, in this case the effect of name familiarity on automatic stereotypes. Similar to Livingston and Brewer (2002), they found that members of the same group can elicit different automatic responses, depending on the familiarity of their names. Specifically, male and female names resulted in faster responses to stereotypic attributes (e.g., jeep and lingerie) if the names were familiar rather than unfamiliar (e.g., John and Sarah, vs. Isaac and Glenda).

Two additional sets of studies show that when the group members are well known, more abstract qualities, such as likeability, may also influence automatic attitudes. In one study, Mitchell, Nosek, and Banaji (1999) used two versions of the IAT to examine how differences among racial group members can affect automatic evaluations, even when the members are being explicitly categorized by race. The participants' task in both tests was to categorize a series of stimuli into four categories: Black or White people, pleasant or unpleasant words. In one of the tests, however, the stimuli representing the Black category were well-liked Black Americans, and the White category was represented by disliked White Americans (e.g., Martin Luther King vs. Dan Quayle); the other test contained disliked Black Americans and well-liked White Americans (e.g., Louis Farrakhan vs. John F. Kennedy). As predicted, Mitchell et al. found that the participants displayed less automatic negativity toward Black Americans when the group members were liked and the opposing category members were disliked, compared to the reverse situation. Although it may not seem surprising to find that well-liked individuals elicit a more positive response than those who are disliked, keep in mind that this effect occurred automatically and while the participants were explicitly focused on the individuals' race. As such, the results contradict the standard assumption that, "Once perceivers categorize the encountered individual, they automatically tend to feel, think, and behave toward that individual in the same way they tend to feel, think, and behave toward members of that social category more generally…" (Fiske, Lin, & Neuberg, 1999, p. 234).

Summary. The studies reviewed in this section suggest that automatic stereotypes and prejudice may not be elicited to the same degree by all members of a

group. Group members who have less of their group's distinctive physical features (but are still categorized as group members), who have less familiar names, or whose likeability does not correspond to the way in which the group is generally viewed, are all less likely to evoke the automatic stereotypes and prejudice associated with their group. This evidence does not disprove the existence of global group attitudes, but it does suggest that the automatic expression of those attitudes may be quite variable in response to specific members of the group.

Although these studies show that group members can elicit different automatic responses, they do not tell us why such effects occur. One possibility is that the individuals' unique characteristics influenced categorization, the process that is presumed to precede the activation of stereotypes and prejudice (Bodenhausen & Macrae, 1998; Brewer, 1988; Fiske & Neuberg, 1990). That is, social categorization is often treated as an all-or-none process (e.g., the person is or is not Black), but it may often occur in a more variable or probabilitic manner (Blair, Judd, Sadler, & Jenkins, in press). Two people may both be obviously Black American, but the one with less Negroid features may not activate the category representation to the same extent as the one with more Negroid features. Indeed, Livingston and Brewer (2002, Study 2) found that participants were slower to categorize Black American faces with less Negroid features, compared to faces with more Negroid features. Livingston and Brewer also found that differences in the automatic evaluation of group members (based on their appearance) disappeared when the participants were instructed to explicitly categorize each person by race (Study 4). On the other hand, Mitchell et al. (1999) obtained differences in automatic evaluation between liked and disliked group members, even though the participants in their study were also explicitly categorizing the individuals by race.

Another explanation for variability in automatic responses to group members is that some group members may have activated sub-categories or group subtypes instead of the more commonly studied superordinate category (Brewer, Dull & Liu, 1981; Devine & Baker, 1991; Deaux, Winton, Crowley, & Lewis, 1985). A light-skinned Black American may be viewed as a "Black businessman", whereas a dark-skinned Black American may be viewed as a "Black gangster". Although little is known about automatic subtyping, there is little reason to question its existence (see Kunda & Thagard, 1996). Positing the automatic activation of subtypes, however, raises additional issues. Researchers must both specify the subtype in a non-circular manner and determine that the stimulus person is automatically categorized into a particular subtype. Moreover, the multiplication of subtypes soon renders the idea of a global attitude meaningless (Bem, 1998).

A third explanation is that category associations may not be the only type of information that is capable of automatic activation. That is, some models specify that stereotypes and prejudice are automatically activated in the first moments of an interaction, whereas the processing and integration of other attributes requires more effort (e.g., Fiske & Neuberg, 1990). The evidence reviewed in this section, however, may be interpreted as showing that both category and individuating characteristics automatically activate associated information, resulting in a response that represents a combination of that information.

Based on the current evidence, we cannot rule out any of these explanations. Indeed, ruling out explanations may not even be desirable. Complexity in automatic attitudes may occur for a variety of (non-exclusive) reasons and each represents an exciting avenue for future research. What is clear, however, is that one cannot assume that automatic stereotypes and prejudice will be the same in response to different group members.

General Discussion

The goal of the present review was to report on the many studies that have directly tested the assumption that automatic stereotypes and prejudice are immutable and inescapable. In contrast to that assumption, the results of these tests show that automatic stereotypes and prejudice can be moderated by a wide variety of events, including, (a) perceivers' motivation to maintain a positive self-image or have positive relationships with others, (b) perceivers' strategic efforts to reduce stereotypes or promote counterstereotypes, (c) perceivers' focus of attention, and (d) contextual cues. In addition, the research shows that group members' individual characteristics can influence the extent to which (global) stereotypes and prejudice are automatically activated. The conclusion that automatic stereotypes and prejudice are not as inflexible as previously assumed is strengthened by the number and variety of demonstrations (nearly 50 in all), the fact that the tests were conducted in the service of many different goals, and by the similarity of findings across different measures. This evidence has several theoretical and practical implications.

First and most obviously, our conception of automaticity needs to be revisited. As noted in the Introduction, an automatic process has often been assumed to be uncontrollable and inalterable by the perceiver's goals and strategies, with the mere presence of a triggering cue sufficient for its inescapable operation (Bargh, 1997, 1999; Devine, 1989; Dovidio & Fazio, 1992). The evidence reviewed here, however,

shows quite clearly that automatic stereotypes and prejudice are controllable, and perceiver's goals and intentions can matter quite a bit (for evidence on the malleability of other automatic processes, see Anthony, 1985; Cohen et al., 1990; Filion et al., 1998; Logan, 1989; Logan & Cowan, 1984). What then is an automatic process? That question cannot be answered easily. We know that well-learned associations (e.g., stereotypes and prejudice) can operate very quickly and efficiently, and that people may not be aware of their influence nor deliberately intend for it to occur. Those features of automaticity – as long as they are defined narrowly – are not challenged by the present results. At the same time, however, those facts should not be used to deduce that an automatic process is "attentionless" or that the perceiver's current motives, goals, and intentions are irrelevant. Similarly, the fact that minimal exposure to specific isolated stimuli (e.g., Black American faces) can set automatic stereotypes and prejudice into motion should not lead to the conclusion that exposure to those stimuli, regardless of surrounding context, will always produce the same automatic process. Social norms, situational pressures, and social context are not insignificant just because the process of interest is automatic.

Second, current models of stereotyping and prejudice are also in need of some revision. As described earlier, many contemporary models hold that the mere perception of social category cues (e.g., dark skin, a female name) results in the (inescapable) automatic activation of stereotypes and prejudice, which introduces early bias into judgment and behavior (e.g., Bargh, 1999; Devine, 1989; Dovidio & Fazio, 1992; Fiske & Neuberg, 1990). Furthermore, that bias is difficult to overcome because unbiased processing of information and control over one's behavior "lag behind" and require considerable cognitive resources and motivation. "Once it is activated, the horse has left the barn, and shutting the barn door at that point does no good" (Bargh, 1999, p. 375). The evidence reviewed here provides a different perspective, suggesting that the bias introduced by automatic stereotypes and prejudice is not obligatory, and both motivation and attention may be important in the early as well as later stages of processing. As a consequence, automatic processes may not be the primary culprits for the endurance of stereotyping and prejudice. Moreover, educational efforts aimed at decreasing discrimination and bias need not be singularly focused on "post-activation" control strategies (Devine & Monteith, 1999).

The third issue raised by the present review concerns the idea that automatic attitudes are more valid than their "controlled" (self-report) counterparts, because the former cannot be strategically altered. To understand the significance of this issue, one must keep in mind that the essential power of the attitude construct has lain in its assumed stability and generality: "That a person's reactions to various members of an object class, observed at different times and in an array of different settings, might all be influenced – and predicted – by that person's global attitude toward that class of objects" (Lord & Lepper, 1999, p. 266). Knowing a person's attitude toward Black Americans, for example, ought to tell researchers what the person is likely to think, feel, and do in response to any particular member of the group in any particular setting (cf., Fiske et al., 1999, p. 234).

Unfortunately, researchers have long had trouble validating this type of attitude. In the domain of intergroup relations, for example, self-reported attitudes were shown to be easily influenced by a number of situational variables, including anonymity (Sigall & Page, 1971), salient social norms (see Gaertner & Dovidio, 1986), and interviewer race (Hatchett & Schuman, 1975). In addition, people often reported an attitude in one situation but did something quite different in another (see Wicker, 1969). In the face of those challenges, two important distinctions have been made. First, It is now widely accepted that all attitudes are not equal and only strong attitudes ought to be stable and consequential (see Fazio, 1989; Krosnick & Petty, 1995). Second, researchers distinguish between a person's attitude and the expression of that attitude, with an acknowledgement that there are many extra-attitudinal factors that can influence what a person says and does in regard to the attitude object. Inconsistent attitudes may be reported across situations or a particular attitude report may not correspond to behavior, not because the attitude is unstable but because its expression is subject to varying constraints and pressures (Campbell, 1963). As a consequence of those two distinctions, researchers interested in capturing the power of attitudes must find a way to reveal the strongest and most "uncontaminated" attitudes. Over the years, many have devoted considerable effort to solving that problem (see Crosby et al., 1980; Roese & Jamieson, 1993; Webb et al., 1966). The development of theory and methods to measure automatic attitudes appeared to provide a very good solution (Dovidio & Fazio, 1992; Fazio et al., 1995). That is, only the strongest associations have been believed capable of operating automatically, and external forces are not supposed to alter (contaminate) an automatic process.

The now bountiful evidence that automatic attitudes – like self-reported attitudes – are sensitive to personal, social and situational pressures, suggests that such attitudes do not provide a ready solution to the problem of attitude malleability. Although future research may yet reveal the perfect attitude, an alternative and perhaps more productive approach is to

cease viewing malleability as a problem. Over the years of "failures" in attitude research, there have been periodic calls for the adoption of a more flexible, situation-specific definition of attitudes (e.g., Lord & Lepper, 1999; Tesser, 1978; Tourangeau & Rasinski, 1988; Wilson & Hodges, 1992). As stated by Tesser (1978), "An attitude at a particular point in time is the result of a constructive process…. And, there is not a single attitude toward an object but, rather any number of attitudes depending on the number of schemas available for thinking about the objects" (p. 297-298). Although the position may sound quite radical, there are stronger and weaker interpretations of it. The weaker version of the argument – and the one intended by Tesser – is that people can view the same object in different ways. A Black woman, for example, may evoke attitudes associated with Blacks or attitudes associated with women, depending on the perceiver's focus of attention (Mitchell et al., 2001). It is the (perceived) attitude object that is changing, not the attitude itself. Although this form of the argument makes the researcher's life more difficult (i.e., one must know how a perceiver views a particular object, in addition to knowing the perceiver's attitude(s)), it preserves the traditional view of attitudes. Nonetheless, the power of the construct is greatly weakened unless there is a relatively small number of potential attitudes that could be elicited by any particular object. To date, researchers have focused on only one variable at a time (e.g., attention to race vs. gender). As such, it is plausible in any one study that either attitude A or attitude B is being elicited. Demonstrations of the many variables that influence automatic attitudes, however, raises the possibility that they exert simultaneous and interactive effects, with each combination revealing a different attitude. It remains for future research to determine what constraints might govern such effects.

The stronger version of the argument is much more radical and it comes from a conceptualization of attitudes (including stereotypes and prejudice) as "states" rather than "things" (Smith, 1998). That is, an attitude is traditionally viewed as a representation (thing) that is stored and then retrieved in much the same form at a later time (Abelson & Prentice, 1989). An alternative view – based on exemplar and connectionist models – is that an attitude is a temporary construction that is formed in response to the current situation (Kunda & Thagard, 1996; Read et al., 1997; Smith 1998; Smith & Zárate, 1990; see also Cohen et al., 1990). This construction may have a stable component that gives the attitude some coherence from situation to situation, but importantly it also reflects the unique attributes of the specific situation, including social and motivational factors. From this perspective, an attitude, whether automatic or more controlled, is inherently flexible and sensitive to the immediate context. Although there may be reasons to prefer one form of the argument over the other (see Smith, 1998), the available evidence does not permit a determination of which is more applicable in regard to the demonstrated malleability of automatic stereotypes and prejudice. Indeed, it is not immediately clear what critical test would differentiate the two. What is clear, however, is that any model of social information processing must address the complexity and variability with which stereotypes and prejudice can be represented, and how that complexity may influence automatic processes.

By demonstrating that automatic stereotypes and prejudice are influenced by many factors, the present review has shown that these processes do not reveal the type of attitude that researchers have presumed. Some may view this evidence as deterrence for their further study. That would be an unfortunate conclusion. Even if they do not have all of the characteristics ascribed to them, automatic responses provide researchers with a unique view of human behavior. People can reason and make deliberate responses; they can also respond for reasons that remain outside of awareness and without their conscious intent. These are distinct processes and they are both worthy of study. In addition, despite their capacity for flexibility and change, automatic attitudes do influence and predict behavior. Perhaps even more importantly, such attitudes appear to predict different types of behavior than their more controlled counterparts (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio et al., 1995; McConnell & Leibold, 2001; Rudman & Glick, 2001). The evidence presented here does not alter the important role that automatic processes may play in social information processing.

Conclusions

Based on the findings reported in this article, an initial model of early social information processing is presented in Figure 1. This model is one attempt to bring coherence to the diverse findings that have been reviewed. In light of the many unanswered questions, the model is admittedly very general and likely incomplete. Nonetheless, it is provided to facilitate continuing discussion of automatic social cognition.

As shown in the Figure, the perceiver is presumed to have broad motives, which he or she brings to the situation (see Fiske, 1998). Some of these motives may be chronically accessible and others may become accessible in the situation. At present, it seems most reasonable to presume that these motives alter the automatic activation of information indirectly, by influencing the perceiver's approach to the situation (e.g., expectancy, focus of attention). The perceiver's

approach, in turn, may directly modulate automatic processes (Bodenhausen & Macrae, 1998; Cohen et al., 1990), or influence which cues (external or internal) are given priority. For example, the motive to maintain self-esteem may influence the extent to which the perceiver focuses attention on cues related to the other person's race or profession (Sinclair & Kunda, 1999); alternatively, the perceiver may focus on cues that draw attention toward or away from a particular social category (Macrae et al., 1997), or the focus may be on internally represented stereotypic or counterstereotypic exemplars (Blair et al., 2001). In accordance with current theories of automaticity (e.g., Bargh, 1996, 1997), those cues are considered to be the most proximal source of influence on the automatic activation of associated constructs, with the important caveat that both the perceiver's approach and the surrounding context (Macrae et al., 1995; Macrae, Hood, et al., in press; Wittenbrink et al., 2001b) can modulate that influence. The social context is also assumed capable of exerting a more indirect influence on automatic processes by increasing the accessibility of a particular motive or by altering the perceiver's approach. For example, a perceiver standing inside a church versus on an inner-city street may have different motives and approaches to an interaction with a Black American. In addition, the same Black American (i.e., the same cue) may automatically arouse different associations depending on the context within which he or she is encountered.

The present review makes a strong case for the malleability of automatic stereotypes and prejudice in response to the perceiver's motives and strategies, and to variations in the situation. This review was not conducted to determine which manipulations are most powerful or which automatic responses are the most likely to resist them. Those questions are interesting, but the heterogeneity of the findings makes such comparisons inadvisable until additional research is conducted. It is that very diversity, however, that suggests the malleability of automatic processes be considered seriously. As such, this evidence stands in stark contrast to assertions that automatic processes are immutable and inescapable, and it highlights the need to pay greater attention to the complexities involved in the early as well as later stages of social information processing.

References

Abelson, R. P., & Prentice, D. A. (1989). Beliefs as possessions: A functional perspective. In A. R. Pratkanis & S. J. Breckler & A. G. Greenwald (Eds.), Attitude strength and function (pp. 361 - 382). Hillsdale, NJ: Erlbaum.

Allport, G. W. (1954). The nature of prejudice. New York: Addison-Wesley.

Anthony, B. (1985). In the blink of an eye: Implications of reflex modification for information processing. Advances in Psychophysiology, 1, 167-218.

Asch, S. E. (1946). Forming impressions of personality. Journal of Abnormal and Social Psychology, 41, 258-290.

Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roediger, III & J. S. Nairne (Eds.), The nature of remembering: Essays in honor of Robert G. Crowder (pp. 117-150). Washington, DC, US: American Psychological Association.

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.

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

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence on social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), Unintended thought (pp. 3-51). New York: Guilford.

Bargh, J. A. (1992). Does subliminality matter to social psychology? Awareness of the stimulus versus awareness of its influence. In R. F. Bornstein & T. S. Pittman (Eds.), Perception without awareness (pp. 236-255). New York: Guilford Press.

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, 2nd edition (pp. 1-40). Hillsdale, NJ: Erlbaum.

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

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer (Ed.), Advances in Social Cognition (Vol. 10). Mahwah, NJ: Erlbaum.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), Dual process theories in social psychology. New York: Guilford.

Bem, S. L. (1995). Dismantling gender polarization and compulsory heterosexuality: Should we turn the volume down or up? Journal of Sex Research, 32, 329 - 334.

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

Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. Journal of Personality & Social Psychology, 81, 828-841.

Bodenhausen, G. V. (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. Psychological Science, 1, 319-322.

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In J. R. Wyer (Ed.), Advances in Social Cognition (Vol. 11, pp. 1-52). Mahwah, NJ: Erlbaum.

Brewer, M. B. (1988). A dual process model of impression formation. Advances in Social Cognition, 1, 1-36.

Brewer, M. B., Dull, V., & Lui, L. (1981). Perceptions of the elderly: Stereotypes as prototypes. Journal of Personality and Social Psychology, 41, 656-670.

Campbell, D. T. (1963). Social attitudes and other acquired behavioral dispositions. In S. Koch (Ed.), Psychology: A study of a science (Vol. 6). New York: McGraw-Hill.

Carpenter, S., & Banaji, M. R. (2001). Malleability in implicit stereotypes and attitudes, Paper presented at the annual meeting of the Society for Personality and Social Psychology.

Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. Journal of Experimental Social Psychology, 14, 88-96.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. Psychological Review, 97, 332 - 361.

Crosby, F., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of black and white discrimination and prejudice: A literature review. Psychological Bulletin, 87, 546-563.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. Journal of Personality and Social Psychology, 81, 800-814.

Dasgupta, N., McGhee, D. E., Greenwald, A. G., & Banaji, M. R. (2000). Automatic preference for White Americans: Eliminating the familiarity explanation. Journal of Experimental Social Psychology, 36(3), 316-328.

Deaux, K., Winton, W., Crowley, M., & Lewis, L. L. (1985). Level of categorization and content of gender stereotypes. Social Cognition, 3, 145-167.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. Journal of Personality and Social Psychology, 56, 680-690.

Devine, P. G., & Baker, S. M. (1991). Measurement of racial stereotype subtyping. Personality and Social Psychology, 17, 44-50.

Dovidio, J. F., & Fazio, R. H. (1992). New technologies for the direct and indirect assessment of attitudes. In J. Tanur (Ed.), Questions about questions: Inquiries into the cognitive bases of surveys (pp. 204-237). New York: Russell Sage Foundation.

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.

Fazio, R. H. (1989). On the power and functionality of attitudes: The role of attitude accessibility. In A. R. Pratkanis & S. J. Breckler & A. G. Greenwald (Eds.), Attitude strength and function (pp. 153 - 179). Hillsdale, NJ: Erlbaum.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? Journal of Personality and Social Psychology, 69, 1013-1027.

Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. Journal of Personality and Social Psychology, 73, 31-44.

Filion, D. L., Dawson, M. E., & Schell, A. M. (1998). The psychological significance of human startle eyeblink modification: A review. Biological Psychology, 47, 1 - 43.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. Gilbert & S. Fiske & G. Lindzey (Eds.), The handbook of social psychology (fourth ed., Vol. 2, pp. 357-411). New York: McGraw-Hill.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. Advances in Experimental Social Psychology, 23, 1-74.

Fiske, S. T., & Taylor, S. E. (1984). Social cognition . New York: McGraw-Hill.

Fiske, S. T., Lin, M., & Neuberg, S. L. (1999). The continuum model: Ten years later. In S. Chaiken & Y. Trope (Eds.), Dual-process theories in social psychology (pp. 231-254). New York: Guilford Press.

Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), Prejudice, discrimination, and racism (pp. 61-89). New York: Academic Press.

Gaertner, S. L., & McLaughlin, J. P. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. Social Psychology Quarterly, 46, 23-30.

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. Journal of Personality & Social Psychology, 78(4), 708-724.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. Journal of Personality and Social Psychology, 60, 509-517.

Gollwitzer, P. M., & Schaal, B. (1998). Metacognition in action: The importance of implementation intentions. Personality and Social Psychology Review, 2, 124-136.

Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. American Psychologist, 35(7), 603-618.

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

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. Journal of Personality and Social Psychology, 74, 1464-1480.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellot, D. S. (in press). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. Psychological Review.

Gregory, W. L., Cialdini, R. B., & Carpenter, K. M. (1982). Self-relevant scenarios as mediators of likelihood estimates and compliance: Does imagining make it so? Journal of Personality and Social Psychology, 43, 89-99.

Hatchett, S., & Schuman, H. (1975). White respondents and race-of-interviewer effects. Public Opinion Quarterly, 39, 523-528.

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotype associations on stereotype activation. Journal of Personality and Social Psychology, 78, 871-888.

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

Kinder, D. R., & Sanders, L. M. (1996). Divided by color. Chicago: University of Chicago Press.

Krosnick, J. A., & Petty, R. E. (1995). Attitude strength: An overview. In R. E. Petty & J. A. Krosnick (Eds.), Attitude strength. Antecedents and consequences (pp. 1-24). Hillsdale: Erlbaum.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. Psychological Review, 103(2), 284-308.

Livingston, R. W., & Brewer, M. B. (2002). What are we really priming?: Cue-based versus category-based processing of facial stimuli. Journal of Personality & Social Psychology, 82, 5-18.

Logan, G. D. (1988). Toward an instance theory of automatization. Psychological Review, 95, 492 - 527.

Logan, G. D. (1989). Automaticity and cognitive control. In J. A. Bargh & J. S. Uleman (Eds.), Unintended thought (pp. 52 - 74). New York: Guilford Press.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. Psychological Review, 91, 295 - 327.

Lord, C. G., & Lepper, M. R. (1999). Attitude representation theory. In M. Zanna (Ed.), Advances in Experimental Social Psychology (Vol. 31). New York: Academic Press.

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. Journal of Personality and Social Psychology, 81, 842-855.

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

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Calvini, G. (1999). Seeing more than we can know: Visual attention and category activation. Journal of Experimental Social Psychology, 35, 590-602.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. Journal of Personality and Social Psychology, 67, 808-817.

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., Hood, B. M., Milne, A. B., Rowe, A. C., & Mason, M. F. (in press). Are you looking at me? Eye gaze and person perception. Psychological Science.

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. Journal of Personality and Social Psychology, 66, 37-47.

Macrae, C. N., Mitchell, J. P., & Pendry, L. F. (in press). What's in a forename? Cue familiarity and stereotypical thinking. Journal of Experimental Social Psychology.

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.

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. Journal of Experimental Social Psychology, 37, 435 - 442.

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (1999). Dissociated implicit attitudes: Examples from race, gender, and profession. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2001). Contextual variations in implicit evaluation. Manuscript under review.

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice-reduction efforts. Journal of Personality and Social Psychology, 65, 469-485.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. Journal of Personality and Social Psychology, 77, 167-184.

Nosek, B. A., & Banaji, M. R. (in press). The go/no-go association task. Social Cognition.

Pratto, F., & Shih, M. (2000). Social dominance orientation and group context in implicit group prejudice. Psychological Science, 11, 515 - 518.

Read, S. J., Vanman, E. J., & Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and Gestalt principles: (Re)Introducing cognitive dynamics to social psychology. Personality and Social Psychology Review, 1, 26-53.

Richeson, J. A., & Ambady, N. (2001). Automatic prejudice in context: The moderating effect of situational roles, Manuscript under review.

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

Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. Journal of Personality & Social Psychology, 81, 856-868.

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. Journal of Social Issues, 57, 743-762.

Schuman, H., Steeh, S., Bobo, L., & Krysan, M. (1997). Racial attitudes in America: Trends and interpretations. (Revised ed.). Cambridge, MA: Harvard University Press.

Sechrist, G. B., & Stangor, C. (2001). Perceived consensus influences intergroup behavior and stereotype accessibility. Journal of Personality & Social Psychology, 80, 645-654.

Sherman, J. W., Lee, A., Bessenoff, G., & Frost, L. (1998). Stereotype efficiency reconsidered: Encoding flexibility under cognitive load. Journal of Personality and Social Psychology, 75, 589-606.

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

Sinclair, L., & Kunda, Z. (1999). Reactions to a Black professional: Motivated inhibition and activation of conflicting stereotypes. Journal of Personality and Social Psychology, 77, 885-904.

Smith, E. R. (1998). Mental representation and memory. In D. T. Gilbert & S. T. Fiske & G. Lindzey (Eds.), The handbook of social psychology (fourth ed., Vol. 1, pp. 391-445). New York: McGraw-Hill.

Smith, E. R., & Zárate, M. A. (1990). Exemplar and prototype use in social categorization. Social Cognition, 8(3), 243 - 262.

Spencer, S. J., Fein, S., Wolfe, C. T., Fong, C., & Dunn, M. A. (1998). Automatic activation of stereotypes: The role of self-image threat. Personality and Social Psychology Bulletin, 24, 1139-1152.

Tajfel, H. (1969). Cognitive aspects of prejudice. Journal of Social Issues, 25, 79-97.

Tesser, A. (1978). Self-generated attitude change. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology (Vol. 11, pp. 289-338): Academic Press.

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

Wasel, W., & Gollwitzer, P. M. (1997). Willentliche Kontrolle der <<automatischen>> stereotypaktivierung: Die rolle subliminaler vs. supraliminaler stimulusdarbietung. Sprache and Kognition, 3-4, 198-210.

Webb, E. J. N. U., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). Unobtrusive measures: Nonreactive research in the social sciences. New York: Rand McNally.

Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. Journal of Social Issues, 25, 41 - 78.

Wilson, T. D., & Hodges, S. D. (1992). Attitudes as temporary constructions. In L. L. Martin & A. Tesser (Eds.), The construction of social judgments (pp. 37-65). 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.

Wittenbrink, B., Judd, C. M., & Park, B. (2001a). Evaluative versus conceptual judgments in automatic stereotyping and prejudice. Journal of Experimental Social Psychology Special Issue:, 37(3), 244-252.

Wittenbrink, B., Judd, C. M., & Park, B. (2001b). Spontaneous prejudice in context: Variability in automatically activated attitudes. Journal of Personality & Social Psychology, 81, 815-827.

Appendix

Go/ No-go Association Task (GNAT; Nosek & Banaji, in press): Participants are presented with a series of test stimuli in rapid succession (e.g., 150-ms. ITI), and they must decide whether to respond to each stimulus within approximately 500 ms. of its onset. On each block of trials, the participants are given two classes of target stimuli to which they must respond (e.g., Black names and negative words), with the additional stimuli serving as distracters. Signal detection analysis is applied to the data and an automatic stereotype or prejudice is defined as greater sensitivity (d') to classes of stimuli that are consistent vs. inconsistent with stereotypes or prejudice (e.g., Black names and negative words, vs. Black names and positive words).

Implicit Association Test (IAT; Greenwald et al., 1998): Participants are asked to categorize a series of stimuli into four categories (e.g., male, female, strong, weak). An automatic stereotype or prejudice is defined as faster categorizations when the stimuli are in stereotype (prejudice) consistent versus inconsistent groups (e.g., male-strong and female-weak, vs. male-weak and female-strong). A paper-and-pencil version of the IAT has recently been developed for use in large groups (see Lowery et al., 2001). In that version, the participants are given a short amount of time to make as many category judgments as they can, and an automatic stereotype or prejudice is defined as a greater number of category judgments in stereotype (prejudice) consistent versus inconsistent groups.

Lexical Decision Task (LDT; Macrae, Bodenhausen et al., 1994; Macrae et al., 1995): Following an event that ought to activate a stereotype (e.g., exposure to a Black man), participants are shown a series of words and nonwords and they must decide as quickly as possible whether each item is a word or a nonword. An automatic stereotype is defined as faster responses to stereotypic versus nonstereotypic attributes (see also Sequential Priming Task below).

Primed Stroop Task (Kawakami et al., 2000): Participants complete a series of trials that begin with the presentation of a category label, followed by a target trait which is printed in one of four different colors. The participants' task is to name the color that the target word is printed in. An automatic stereotype or prejudice is defined as slower RT (more interference) to targets that are preceded by stereotype (prejudice) consistent category labels versus inconsistent category labels (e.g., skinhead-hostile vs. elderly-hostile).

Sequential Priming Task (Banaji & Hardin, 1996; Fazio et al., 1995; Wittenbrink et al., 1997): Participants complete a series of trials on which a prime stimulus appears briefly (sometimes subliminally) followed by a target stimulus. The participants are asked to make a simple judgment about the target (e.g., male vs. female; good vs. bad; word vs. nonword). An automatic stereotype or prejudice is defined as faster responses to targets that are preceded by stereotype (prejudice) consistent primes versus inconsistent primes (e.g., strong-John vs. weak-John; Black-maggots vs. White-maggots). A flanker task is very similar, except that the prime (flanker) appears simultaneously with and spatially close to the target, and the outcome of interest is response interference rather than response facilitation (see Macrae et al., 1999).

Word Completion Task (Gilbert & Hixon, 1991; Spencer et al., 1998): Simultaneous with or following an event that ought to activate a group stereotype (e.g., an Asian assistant; subliminal presentation of a face), participants are given word fragments to complete (e.g., pol___). An automatic stereotype is defined by the number of stereotypic word completions (e.g., polite) in that condition versus a control condition in which the stereotype should not have been activated.

Table 1. Summary of Research on the Malleability of Automatic Stereotypes and Prejudice. Studies Within Each Section are Listed Alphabetically by Author.

Study	Measure	Primary Results
Motives: Self-image Enhancement		
Sinclair & Kunda (1999, Study 1)	Word Completion	Fewer Black stereotype completions following positive feedback & more completions following negative feedback from a Black manager, compared to responses following feedback from White manager.
Sinclair & Kunda (1999, Study 2)	Word Completion	Equivalent levels of Black stereotype completions following observation only of positive vs. negative feedback given by Black manager.
Sinclair & Kunda (1999, Study 3)	LDT	Weaker Black stereotypes following positive feedback & stronger stereotypes following negative feedback from a Black doctor, compared to responses following feedback from White doctor. Stronger doctor stereotypes following positive feedback & weaker stereotypes following negative feedback from a Black doctor, compared to responses following feedback from White doctor.
Sinclair & Kunda (1999, Study 4)	LDT	Weaker Black stereotypes following positive (vs. no feedback) from Black doctor, but only for high-prejudiced participants. Stronger doctor stereotypes following positive (vs. no feedback) from Black doctor, regardless of prejudice-level.
Spencer et al. (1998, Study 1)	Word Completion	More Asian stereotype completions in the presence of an Asian assistant under high attentional load, following negative vs. positive feedback.
Spencer et al. (1998, Study 3)	Word Completion	More Black stereotype completions following subliminal exposure to a Black face under high attentional load, when prior feedback was negative vs. positive.
Motives: Social Relations		
Lowery et al. (2001, Study 1)	IAT	Less negativity toward Blacks in the presence of Black than White experimenter.
Lowery et al. (2001, Study 2)	IAT	Less negativity toward Blacks in the presence of Black than White experimenter, but not for Asian participants. Experimenter race effect also n.s. for males.

Lowery et al. (2001, Study 3)	IAT	Less negativity toward Blacks when instructed by Black experimenter to be unprejudiced, for both White & Asian Ps.
Lowery et al. (2001, Study 4)	Sequential Priming	Less negativity toward Blacks in the presence of a Black than White experimenter, but not for Asian participants.
Richeson & Ambady (2001)	IAT	Negativity toward Blacks highest with anticipated superior role during interaction with a Black partner; negativity lowest with anticipated subordinate role, & intermediate for anticipated equal-status role.
Sechrist & Stangor (2001)	Sequential Priming	Weaker race stereotypes when participants perceived low vs. high consensus with their beliefs.
<u>Stereotype Strategies: Suppression</u>		
Macrae, Bodenhausen, et al. (1994, Study 3)	LDT	Stronger skinhead stereotypes following suppression vs. no suppression.
Galinsky & Moskowitz (2000, Study 1)	LDT	Stronger elderly stereotypes following suppression vs. no suppression.
Gollwitzer & Schaal (1998)	Stroop	Weaker gender stereotypes following a stereotype suppression implementation intention vs. control intention, but only for specific stimulus person.
Kawakami et al. (2000, Studies 1 & 2)	Stroop	Weaker skinhead stereotypes following extended stereotype negation training vs. control conditions, even after 24 hr. delay.
Kawakami et al. (2000, Study 3)	Sequential Priming	Weaker race stereotypes following extended stereotype negation training vs. stereotype affirmation training.
<u>Stereotype Strategies: Promoting Counterstereotypes</u>		
Blair & Banaji (1996, Study 3)	Sequential Priming	Weaker gender stereotypes with counterstereotype than stereotype expectancy, at both 2000-ms. & 350-ms. SOA.
Blair & Banaji (1996, Study 4)	Sequential Priming	Weaker gender stereotypes with counterstereotype expectancy at 2000-ms. vs. 250-ms. SOA, with priming effect n.s. even in latter condition.
Blair et al. (2001, Studies 1 & 2)	IAT	Weaker gender stereotypes following counterstereotypic mental imagery vs. control conditions.
Blair et al. (2001, Study 3)	IAT	Stronger gender stereotypes following stereotypic vs. neutral mental imagery.

Blair et al. (2001, Study 4)	GNAT	Weaker female stereotypes, following counterstereotypic mental imagery vs. control conditions, but only for female participants.
Blair et al. (2001, Study 5)	False Memory	Weaker gender stereotypes following counterstereotypic vs. neutral mental imagery.
Carpenter & Banaji (2001)	IAT	Weaker gender stereotypes following counterstereotypic vs. neutral mental imagery. But no effect of imagery on gender evaluations.
Dasgupta & Greenwald (2001, Study 1)	IAT	Less negativity toward Blacks following exposure to admired Blacks vs. disliked Blacks or control, even after 24 hr. delay.
Dasgupta & Greenwald (2001, Study 2)	IAT	Less negativity toward elderly after exposure to admired vs. disliked older individuals.
Rudman et al. (2001, Studies 1 & 2)	IAT	Less negativity and weaker stereotypes of Blacks at the end of a semester diversity course, compared to students in control course (quasi-experiment).
<u>Focus of Attention</u>		
Gilbert & Hixon (1991, Studies 1 & 2),	Word Completion	Fewer stereotype completions in the presence of an Asian assistant, under high vs. low attentional load.
Macrae et al. (1997, Studies 1 & 2)	Sequential Priming	Weaker gender stereotypes with attention focused on the presence of a white dot vs. the animateness of the object.
Macrae et al. (1999, Studies 1 & 2)	Flanker Task	Less interference from gender-related flankers when they were placed far vs. near the target.
Mitchell et al. (2001, Study 1)	IAT	Negativity toward Black females & positivity toward White males with stimuli categorized by race; reverse pattern with stimuli categorized by gender.
Mitchell et al. (2001, Study 2)	IAT	More positivity toward Black athletes than White politicians when categorized by occupation; reverse pattern when categorized by race.
Mitchell et al. (2001, Studies 3 & 4)	GNAT	Negativity toward Black females & positivity toward White males with attention focused on race; reverse pattern with focus on gender.
Spencer et al. (1998, Study 2)	Word Completion	Fewer Black stereotype completions following subliminal exposure to a Black face under high vs. low attentional load.
Wittenbrink et al. (2001a)	Sequential Priming	More generalized negativity toward Blacks with evaluative than lexical decisions. Similar negative stereotypes of Blacks with both types of decisions.

Configuration of Stimulus Cues		
Macrae et al. (1995, Study 3)	LDT	Female stereotypes weaker after seeing a Chinese woman use chopsticks & stronger after seeing her put on makeup, compared to control. Reverse pattern of effects for Chinese stereotypes.
Macrae, Hood, et al. (in press, Study 2)	Sequential Priming	Photographed men and women with averted or closed eyes elicited weaker gender stereotypes than those with direct eye gaze.
Wittenbrink et al. (2001b, Study 1)	IAT	Less negativity toward Blacks after seeing a movie clip of Black Americans in a positive vs. negative setting.
Wittenbrink et al. (2001b, Study 2)	Sequential Priming	More negativity to Black vs. White face primes embedded in a street scene; no difference in (positive) evaluations with faces embedded in a church scene.
Characteristics of Individual Category Members		
Livingston & Brewer (2002, Studies 1, 4, & 5)	Sequential Priming	Less negativity toward Blacks with less Negroid facial features, compared to Blacks with more Negroid features.
Macrae, Mitchell, et al. (in press, Study 3)	Sequential Priming	Weaker gender stereotypes with unfamiliar than familiar names.
Mitchell et al. (1999)	IAT	Less negativity toward well-liked vs. disliked Black exemplars.

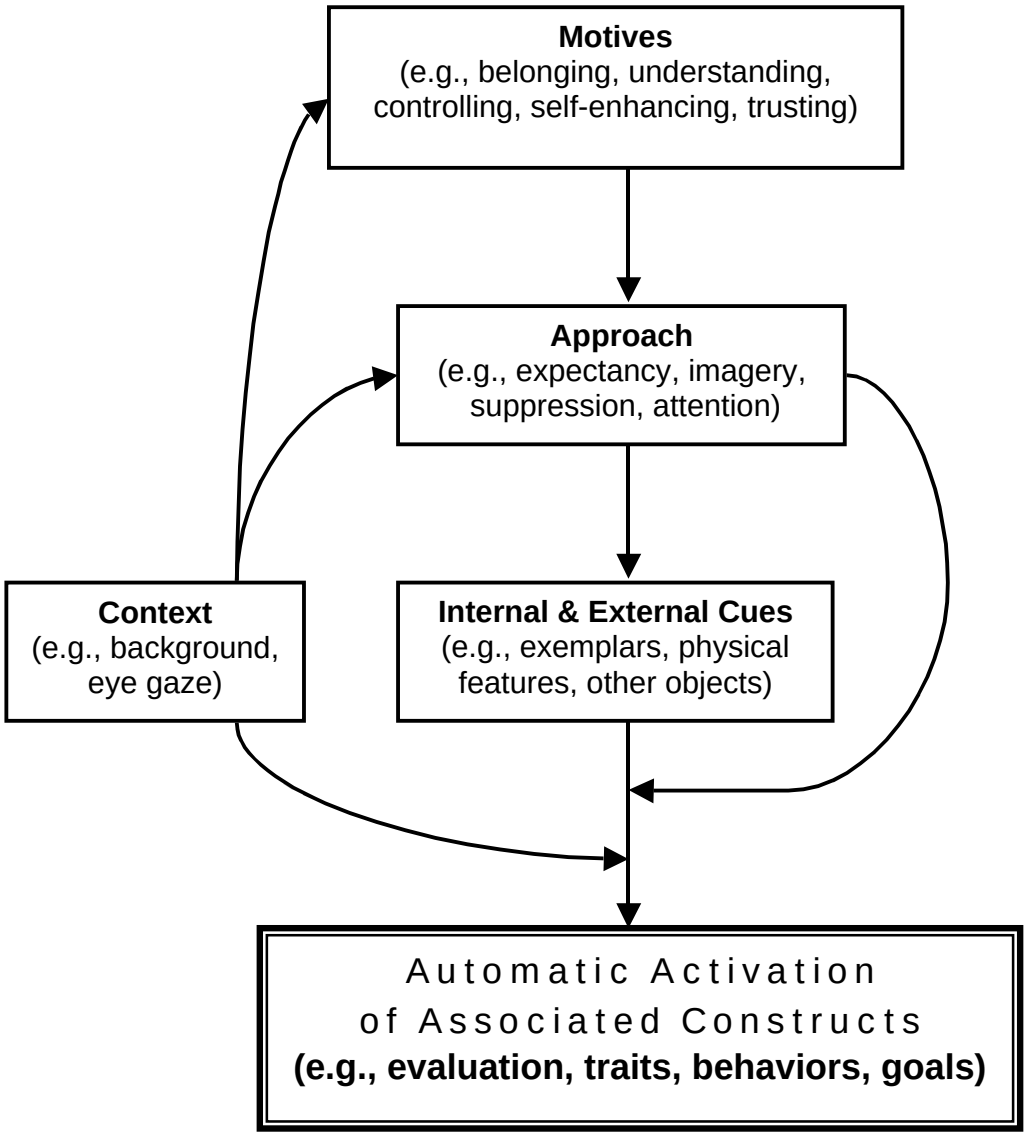


Figure 1. An initial model of early social information processing.