Egalitarian goals trigger stereotype inhibition: A proactive form of stereotype control

Gordon B. Moskowitz a,*, Peizhong Li b

a Lehigh University, Bethlehem, PA, USA
b School of Public Administration, Hohai University, Nanjing, China

Introduction

Ironically, it is often the case that the conscious effort associated with goal pursuit and action control undermines achievement. The very act of trying is one's undoing. For example, thinking hard about losing weight can make weight-loss more difficult (one is focused more on food, especially fattening foods, than when not dieting at all). The current research illustrates how lack of consciousness in a goal pursuit can help us attain what we want. This is illustrated in the domain of controlling weight-loss more difficult when one encounters a member of a stereotyped group. This is a proactive form of control because it prevents a stereotype from ever being retrieved from memory, despite perceivers having categorized the person to a social group. It occurs prior to awareness of either the goal or stereotype activation. It requires no effortful expenditure of mental resources (cf., Fiske, 1989), nor for one to be aware of the bias, nor for one to consciously try to prevent bias from having influence (“correcting” for the stereotype).

Stereotyping others is goal-driven

We begin with two statements about stereotyping that are likely intuitively obvious, and clear to experts and nonexperts alike. First, a stereotype is knowledge that exists in the mind of an individual that is associated with a group of people, knowledge that is learned from, and shared with, others in the culture. Second, stereotyping is a process that unfolds in phases (e.g., Kunda & Spencer, 2003): In an initial phase, the stereotype is pulled from memory due to its association with a category (e.g., old, fat, Black, Jew, woman) that has recently been used to identify a person. In the next phase, this “perceptually ready” stereotype is used to help one understand others and plan behavior toward those others. These two assertions lead to perhaps less obvious third and fourth statements. The third is that, regarding when stereotypes are pulled from memory, it is far more common and silent (we rarely recognize this has happened) than people know (e.g., Devine, 1989). The fourth is that stereotypes are triggered and used for a reason; they serve a goal. As such, they have...
been likened to tools pulled from a “cognitive tool box” when people are encountered (e.g., Gilbert & Hixon, 1991; Macrae & Bodenhausen, 2000).

If stereotype activation is a tool that serves a goal, what goal does it serve? Given that stereotype contributes to discrimination and prejudice (e.g., Allport, 1954), a natural initial answer to this question that received empirical attention was that stereotypes service goal-related motivations, such as the desire to feel superior to others and have positive identity associated with one’s social group (e.g., Tajfel & Turner, 1979); the wish to attain and maintain social, physical, political, and economic power over others (e.g., Jost, 2001; Sherif, 1958; Sidanius & Pratto, 1999) and to protect oneself (and loved ones) from what seems alien, strange, and threatening (e.g., Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950). But stereotyping is found even among people who do not fear others, wish to dominate, or need to boost self-esteem (e.g., Devine, 1989).

Lippmann (1922) introduced the term stereotype in describing a cognitive process used by people because they have a goal of being efficient and economical with their time and mental energy when trying to make sense of (understand) the world around them. Meaning must be attained quickly and without much strain, given both the press of the external world (the speed of life, the multitude of things to process) and the limits of the human processing system. From Allport’s (1954) principle of least effort, up to and beyond Gilbert and Hixon’s (1991) assertion that “people rely strongly on prior information to ease the burden of ongoing perception” (p. 514), stereotyping has been described as in the service of a goal to attain meaning; to know what thoughts and acts are appropriate. We shall refer to this as an epistemic goal.

Heider (1944) described this goal as a central pursuit, a “causal drive,” that guides much of daily life. Because of its chronic and habitual nature, the epistemic goal comes to operate without consciousness. Moskowitz (1993) and Neuberg and Newsom (1993) illustrated that one need not consciously intend to a) categorize or understand another person, b) be efficient, or c) arrive at closure quickly. The epistemic goal is silently pursued, routinely. Stereotyping is an efficient, nonconscious process precisely because the goal it serves is not conscious to the individual.

Proactive versus reactive stereotype control

Given the efficient nature of both the stereotyping process and the goals served by this process, how can stereotyping be controlled? Until recently, control in general was seen as a conscious process, with intended responding equated with effort (Wegner & Bargh, 1998). This is one reason stereotype control has traditionally been depicted as a process whereby a) a stereotype is first triggered, b) the possibility for bias then exists, c) this possibility is detected by the person, and d) at that point a goal to prevent bias, or overturn it if it has already occurred, is pursued. In fact, until recently, stereotype activation was seen as inescapable—a necessary by-product of the categorization process (Allport, 1954; Bargh, 1999; Brewer, 1988; Devine, 1989). Such a view, by definition, precludes the possibility of stereotypes not being activated.

However, many perceivers-based and target-based factors are now known to disrupt stereotype activation. Perceiver-based factors include cognitive load (Gilbert & Hixon, 1991), holding counterstereotypic expectancies (Blair & Banaji, 1996), beliefs (Lepore & Brown, 1997), associative learning (Kawakami, Dovidio, Moll, Hermens, & Russin, 2000), chronic motivations (e.g., Glaser & Knowles, 2008; Moskowitz, Gollwitzer, Wasel, & Schaal, 1999; Plant & Deavin, 1998), and the direction of gaze (Macrae, Hood, Milne, Rowe, & Mason, 2002). Target-based factors also prevent stereotype activation—typically of a target’s face, fame of a target, skin tone, category ambiguity (is the person clearly a woman, Black, etc.?), and whether a target’s other roles/categories are salient and used in categorization. Activation is not inevitable.

Yet, even as research has dispelled the notion of stereotype activation as inevitable, the ability to regulate this cognitive activity through the motivational system has been relatively ignored. Time and again, researchers stop short of claiming that the triggering of stereotypes can be willfully stopped and instead call for the importance of noticing implicit stereotypes so that one can then exert conscious attempts to curtail their influence. Stereotype control is still largely seen as rooted in dual process models (e.g., Brewer, 1988; Fiske & Neuberg, 1990; Posner & Snyder, 1975). These models describe one set of psychological processes inexorably giving rise to specific outputs, but a separate set of processes may be incompatible with those outputs and draw on conscious control to inhibit those outputs from influencing how one ultimately responds. Regardless stereotype control, a first process silently activates the stereotype. A second process, such as the recruitment of explicit goals that are incompatible with the use of stereotypes, prevents the already-activated stereotype from influencing responding.

This is a reactive strategy of stereotype control. Goals allow one to correct-for (e.g., Devine, Monteith, Zuverink, & Elliot, 1991) or decontaminate (e.g., Wilson & Brekke, 1994) cognition from the biasing influence of an already-activated stereotype, overriding it with a more appropriate response (“putting the brakes on prejudice,” e.g., Monteith, Ashburn-Nardo, Voils, & Czopp, 2002). A rich tradition of research has illustrated this approach to be a highly effective strategy (e.g., Devine, 1989; Fiske & Neuberg, 1990; Monteith, 1993; Wegner & Erber, 1992), even impacting a U.S. Supreme Court ruling regarding employer culpability for unintended workplace discrimination. Granting its impact, we argue that this conception of stereotype control is too limiting and can be extended by moving beyond dual process notions.

Control over stereotyping is not only attainable through subsequent intervention of the conscious will to overturn activated stereotypes. Control can be exerted on stereotype activation at the first step of the process. This is a proactive strategy of control, one focused at the level of basic social–cognitive functioning, arguing that goals disrupt the activation of stereotypes. This focus on the role of the control system distinguishes our approach from existing research on the prevention of stereotype activation described above (cognitive load, associative learning, gaze).2

What is often lost when discussing stereotypes is the fact that they are the product of a goal pursuit, and are, as such, willed, or wanted. Not realizing one stereotypes does not mean one does not, or that such unnoticed stereotypes are unintentional. Indeed, explicitly wanting to not stereotype does not preclude having an unconscious goal that relies on stereotyping. Epistemic goals, for example, are compatible with stereotype activation—the goal controls stereotyping, with control here existing in the form of heightening the response tendency. Given this logic, it should also be true that control in the form of weakening the response tendency is also possible if one's goals are incompatible with stereotypes. One has the power to efficiently not stereotype by exercising the same self-regulatory system that at times promotes stereotyping. The cognitive process serves a goal, and whether stereotype activation occurs is thus an issue of what goal the individual is silently pursuing. We examine whether an egalitarian goal that is incompatible with stereotyping will inhibit, not activate, stereotypes. Any person can not stereotype, without even being aware of exerting control, dependent on

---

1 As proof of this efficiency, researchers have pointed to the functionality of stereotype use, such as the increased efficiency produced in one’s thought when one is stereotyping (e.g., Macrae, Bodenhausen, Milne, & Jetten, 1994; Sherman, 2001).

2 Some of that research has focused on the question of shifting the type of category that is used to identify a person (by manipulating gaze, target ambiguity, fame, skin tone, etc.). Activating a stereotype requires first identifying a person as a member of a group, and such research reveals that shifting how a person is categorized affects how (and if) the person is stereotyped. Some of that research has focused on postcategorizing processes where stereotype activation is replaced as one’s dominant reaction to the category due to new learning experiences—by forging new associations to the category or developing counterstereotypic expectancies. None of that research addresses how shifting a given individual’s goals impacts the activation of stereotypes.
what goals are triggered. Stereotyping is a default response only insofar as goals compatible with stereotyping are one’s default goals.

Unconscious goals and stereotype control

The recent proliferation of research on unconscious goals (e.g., Bargh, 1990; Custers & Aarts, 2005a; Förster, Liberman, & Friedman, 2007; Kruglanski et al., 2002; Moskowitz, Li, & Kirk, 2004; Shah, 2005) has made a proactive approach to stereotype control tenable. It was a watershed moment in goals theory when Chartrand and Bargh (1996) found that goals are not always consciously selected. Rather, a goal structure, just like any other cognitive structure, can be primed. Evidence of goal priming is now overwhelming (for reviews, see, Dijksterhuis, Aarts, & Chartrand, 2007; Moskowitz & Gesundheit, 2009). Once a goal is primed, to what extent are the steps taken that move one toward goal completion also occurring outside of consciousness?

Compensatory responding

A goal specifies an as-yet-unattained end state. The steps or responses taken that move the person toward that end are compensatory—they compensate for the failure of not reaching the goal (e.g., Wicklund & Gollwitzer, 1982). Traditionally, the study of goals focused on explicit steps taken to help one produce outcomes specified by explicit goal states. However, compensatory responses are now known to include not only overt actions that are in the service of the goal but also cognitive operations and automatic processes.

Examples of such compensatory cognition include the following: A goal (such as to be egalitarian) directs selective attention to goal-relevant stimuli in the environment that one is not consciously able to detect (Moskowitz, 2002; Moskowitz, Li, Ignarri, & Stone, in press). A goal (to choose between various products or to solve a problem) leads one to implicitly ruminate on material related to the goal and strengthen memory traces for information, even when one has consciously stopped thinking about the task (Dijksterhuis, Bos, Nordgren, & Van Baaren, 2006; Koole, Smeets, van Knippenberg, & Dijksterhuis, 1999). A goal (to form attributions) leads one to implicitly follow rules of inference (Jones & Davis, 1965) and infer traits (Moskowitz, 1993). A goal (to suppress thoughts) leads one to implicitly monitor cognitive processing to detect unwanted thoughts and prevent them from entering consciousness (Postman, Bruner, & McGinnies, 1948; Wegner, 1994). A goal (to do business in a specific location) facilitates processing of goal-relevant information (features associated with the desired location, e.g., Aarts, Dijksterhuis, & Midden, 1999). And a goal (to memorize) impacts how information is clustered in memory (e.g., Chartrand & Bargh, 1996).

Goal shielding

Kruglanski et al. (2002) argue that goals, as representations, exist among a system, or network, of related representations. Movement toward one goal impacts standing on another (e.g., Fishbach & Trope, 2008; Shah & Kruglanski, 2002). Goals may be compatible and facilitate each other. Goals may also compete and result in the inhibition of a competing goal to “shield” a focal goal (e.g., Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trötschel, 2001; Fishbach, Friedman, & Kruglanski, 2003; Shah, Friedman, & Kruglanski, 2002). Goal shielding describes the general set of processes by which a given goal is promoted through cognitive operations that inhibit distractions to the goal and facilitate the detection and processing of goal-relevant stimuli (as well as other goals) that are compatible with the focal goal.

For example, Shah (2003) illustrated that priming a goal that was related to a focal goal improved task performance on the focal goal. Shah (2003) further illustrated that incompatible goals produce an inhibitory effect on the accessibility, commitment, and pursuit of one of the goals. When a goal incompatible with verbal fluency was primed prior to a task assessing verbal fluency, the goal system shielded the primed goal by inhibiting the goal of verbal fluency. Aarts, Custers, and Holland (2007) also illustrated inter-goal inhibition. Participants were primed with the goal of socializing while they had been pursuing an incompatible goal—studying. Goal shielding emerged in that people with studying goals inhibited the goal to socialize. Goal shielding research shows that goals reside in a hierarchically-organized network with inhibitory and facilitative associative links among goals that are, respectively, incompatible or compatible.

Stereo
type inhibition

Research on goal priming, goal shielding, and compensatory cognition each inform stereotype control. Goals incompatible with stereotyping can be primed and held by the person outside of his/her awareness. Such goals, despite not being consciously detected, can guide how one responds, activating and inhibiting associated goals, shaping both behavior and cognition. When lower-order goals that implicate the activation of stereotypes are compatible with higher-order goals, then goal shielding should facilitate stereotype activation. When lower-order goals that implicate the activation of stereotypes are incompatible with higher-order goals that denounce or reject stereotyping, then stereotypes should be inhibited.

Consider an example of such goal shielding relating to a White perceiver encountering an African American man. The person is typically perceived in the context of a high-order goal to categorize in terms of race, gender, or age. A category (Black man) once activated will lead to heightened accessibility of category-relevant information, with the retrieval of that information also under the direction of goals. High-order goals, such as to understand the person’s behavior, trigger lower-order goals, such as to understand behavior using the least effort possible, and to retrieve stereotypes to facilitate fast, effortless, prediction. These goals determine which knowledge, from among many possible types of category-relevant knowledge, is retrieved and activated. It facilitates the retrieval and use of social stereotypes associated with the group.

However, the processing of these same stereotypes can be inhibited if goals incompatible with stereotype use, such as egalitarian goals, are in place. It has already been shown that chronic egalitarian goals inhibit stereotypes (Moskowitz et al., 1999). Inhibition due to chronic goals is accompanied by other forms of goal shielding, such as strengthening the association between the goal and group members (Moskowitz, Salomon, & Taylor, 2000) and the scanning of the environment for goal relevant items (Moskowitz, 2002; Moskowitz et al., in press). Thus, if a category (e.g., Black man) activates an egalitarian goal, the nature of the spreading activation from the category is altered relative to when the category is not associated with such a goal. Egalitarian goal activation is often incompatible with the goal of retrieving stereotypes, and should, under those conditions, activate lower-order goals such as the goal to inhibit stereotypes.

Experiment 1

The current experiment examines if egalitarian goals that are not chronically held lead to control of stereotype activation by triggering associated goal shielding operations that include the inhibition of stereotypes. Egalitarian goals are triggered in this experiment by asking participants to reflect on a failure experience. Many models of goal selection (e.g., Carver & Scheier, 1981; Lewin, 1936; Powers, 1973; Wicklund & Gollwitzer, 1982) reveal that a goal is triggered when one contemplates failure in the goal domain; by a person detecting a discrepancy between their actual responses and a desired response. This
discrepancy is said to produce a psychological tension that impels the organism to reduce the tension and approach the standard.\(^4\) Such a discrepancy is sometimes introduced via negative feedback (e.g., Koole et al., 1999; Spencer, Fein, Wolfe, Fong, & Dunn, 1998). In the current research, as in other research (e.g., Monteith, 1993; Moskowitz, 2002), the discrepancy is introduced by having participants reflect on how their past behavior is discrepant with a desired goal (in this case, the goal of being egalitarian to Black men).

In past research where goals are triggered by discrepancy detection, the discrepancy is typically salient to the individual, and he/she is then placed in a situation that clearly affords an opportunity to behave in a manner that addresses the failure (e.g., Gollwitzer, Wicklund, & Hilton, 1982; Wicklund & Gollwitzer, 1981). For example, Monteith (1993) had participants write how they should versus would act toward a minority (triggering a discrepancy). They then respond to a task explicitly about that minority group. In such situations, the goal is the focus of conscious attention as one is responding. The compensatory responses are driven by awareness of the connection between the discrepancy-arousing and -reducing tasks.

But, a response can be “compensatory” without the individual being aware of its association to a goal (or of the goal itself). Such nonconscious goal pursuit is examined here. It is evidenced in two ways. First, by having goal activation temporally distanced from the response (so that when the response is made the goal is accessible, but the person is not aware of this). Second, by making the response one that is not explicitly linked to the goal in the mind of the individual. This is achieved by using an implicit measure of stereotype accessibility as the response in question. Participants responded to words that were either relevant or irrelevant to the stereotype of Black men following faces of either Black or White men. A robust finding in the psychological literature is that responses are facilitated to concepts that have been activated and slowed to concepts that have been inhibited. Using this response time logic, Experiment 1 illustrates that stereotypes are activated when White participants encounter a Black man unless an egalitarian goal has been triggered. Egalitarian goals instead initiate stereotype inhibition.

Method

Research participants

Participants were 42 White students at Lehigh University who participated to partially fulfill a requirement in their introductory psychology course. All were native English speakers.

Materials

Goal status manipulation. A paper-and-pencil task that was described as a goals survey was given to participants. The survey contained 26 items (e.g., equality, freedom, intelligence, respect for tradition, humility, egalitarian). Participants were asked to rate the importance of each goal on a Likert scale, with values ranging from 1 (not at all important) to 9 (extremely important). When finished, they turned the page and read the instructions for the next task.

These instructions informed participants each person in the experiment was focusing on one goal from the survey. Half were told they were to focus on egalitarian goals. Their task was to write a description of an instance in which they failed to live up to the ideal specified by an egalitarian goal, and to do so by relaying an event relating to African American men. The instructions stated the experiment was not interested “in fleeting moments of negative thoughts towards others, but for you to reflect on instances in your life when you truly performed in a way that led you to violate the egalitarian ideal.” An egalitarian goal was defined for them as:

Egalitarian goals are part of the “American Creed” and involve acting fair and open-mindedly, being tolerant of others despite the fact they differ from you, and treating people equally regardless of their ethnicity, religion, gender, race, physical appearance.

The remaining half of the participants were each asked to similarly describe a failure, but in the domain of “respect for traditions.” Respect for tradition was defined as: “acting in a manner that recognizes the value of established practices and institutions, honoring these traditions, and continuing to practice and observe the rituals associated with the traditions.”

Since these tasks were of a personal nature, participants were reminded responses would be anonymous and that they should be truthful, even if admitting things they do not like.

Computer. The computer was a Dell Dimension 8200 with a Pentium 4 processor and 256 MB RAM attached to a 17-inch Dell E770 Color Display monitor.

Pictures used as primes. Photographs were drawn from a pool of 180 black-and-white yearbook pictures of White and Black men. On the critical trials half of the photos depict White, while the other half Black, male individuals. Filler trials contained White male faces.

Target words. Eight words relevant to the stereotype of African Americans and eight neutral words matched for length, frequency, and valence were used on critical trials. Attributes were selected as relevant or irrelevant to this stereotype based on the research of Moskowitz et al. (2000), Devine (1989), and Devine and Elliot (1995). Stereotype-relevant attributes (crime, stupid, lazy, violent, prison, poor, hostile, and welfare) and control attributes (annoying, nervous, indifferent, cruel, selfish, death, and nosy) were all negative in valence.

Reaction time task. The task contained 24 practice trials and 96 trials (divided into two blocks of 48 trials). Each trial contained two phases—a priming phase, where pictures (primes) were presented along with irrelevant stimuli, and a response phase, where participants performed a lexical decision task. The first phase consisted of a male face presented simultaneously with a pair of letter-strings. These pairs of letter-strings were relevant to the cover story. Participants were told the task was to determine if the strings in the pair were identical by pressing buttons marked “same” versus “different” on the keyboard. One-quarter of the trials contained a pair of identical words, one-quarter contained identical letter strings that were not words, one-quarter contained two separate words, and one-quarter contained two separate letter strings. Faces were ostensibly presented as a distracter to the “focal” task. Participants were instructed to attempt to ignore them and to perform the string identification task as quickly as possible. Faces and string-pairs remained on the screen until a response was made or until 1000 ms elapsed (whichever comes first). From among these 96 trials, there were 16 designated as critical trials. On critical trials, the type of face was manipulated—eight contained Black and eight contained White faces. The remaining 80 trials were filler trials with 90% White and 10% Black faces. Phase 1 served as a way to subtly present (by having them be ignored) primes. Responses to the letter-strings pairs were not of interest. The response of interest was made during the next phase of the trial.

Phase 2 of a trial commenced 10 ms after a response was made in phase 1. It contained a single string of letters, and participants were

\(^4\) In various literatures, this tension has been referred to as a self-discrepancy (e.g., Higgins, 1989), a feeling of incompleteness (e.g., Wicklund, & Gollwitzer, 1982), a failure (e.g., Carver, & Scheier, 1981), and a lack of affirmation (e.g., Steele, 1988). Thus, inherent to having a goal is a state of negative feedback that informs one, via recognition of one’s standing regarding the pursuit of a goal, need not require information (e.g., Steele, 1988).
instructed to respond to each string by making a lexical decision—is this a word in the English language? To respond, participants were asked to press either a button marked “yes” or “no” on the keyboard. On half of the trials the letter-string contained a word, the remaining half had nonwords. Critical trials contained eight stereotype-relevant and eight stereotype-irrelevant words. On these critical trials, half of the words followed Black faces from phase 1 of the trial and half followed White faces.

Procedure and design

Upon arrival at the laboratory, participants were taken to a private room to work on a first task. They were informed that while they had agreed to participate in a perception experiment, they would first have an opportunity to perform an unrelated task to help a graduate student preparing a project examining how people remember positive versus negative information about the self. When finished with this task, they were to seal the forms in an envelope and get the experimenter, who would take them to a separate room for the perception experiment. They were then given the goal status manipulation and an empty envelope, and the experimenter exited.

Upon finishing the task, participants were next brought to a cubicle with a desk and computer. Task instructions were presented on the monitor. The task was described as having multiple trials, with each trial consisting of two phases. Phase 1 was said to contain three items presented simultaneously: A face, which they are to ignore, and two letter strings presented one on top of the other. The task was to decide if the letter strings were identical. Phase 2 contained a string of letters, and the task was to decide if the letters comprised an English-language word—a lexical decision task (LDT). For each phase, they were instructed to respond as accurately and quickly as possible. When finished with all the trials they were debriefed, thanked, and dismissed.

The experiment had a Goal (egalitarian, tradition) by Prime Type (Black, White) by Word Type (stereotype-relevant, stereotype-irrelevant) mixed-factorial design. Goal was a between-participant variable, Prime Type and Word Type were within-participant variables.

Results and discussion

If stereotypes are activated by the faces/primes, people should be facilitated in their ability to respond to words that are relevant to the stereotype. Thus, seeing a Black face on phase 1 of the trial should make one quicker to respond to such words on the lexical decision task presented at phase 2. Stereotype control is evidenced by this effect disappearing so that stereotypic words no longer are responded to more quickly, and instead, due to spreading inhibition, are responded to more slowly following faces of Black men (Moskowitz et al., 1999).

Two independent coders rated each of the statements produced by participants on the “goal status manipulation” task. Each statement was rated on a 7-point scale ranging from responses that did not attempt to describe what was asked (a score of 1) to detailed answers with clear examples of a failure (a score of 7). People with scores less than 3 (eight participants) were excluded from analyses. Interrater reliability was high (r = .70, p < .001). In addition, ratings on the Likert scales to the items “egalitarian,” “equality,” “fairness,” and “tolerance” were examined to determine if participants valued being egalitarian. The goal shielding logic requires the goal is not one participant’s devalue or dislike. Participants were considered to value the goal if they averaged a score of six or higher on these items. Across all experiments reported in this paper, no participants were excluded based on this assessment (all valued egalitarianism). For the computerized portion of the experiment, reaction times more than 3 standard deviations from the mean (5.2% of responses) were considered outliers and eliminated from the analyses. No participants expressed suspicion the words to which they responded were related to the faces.

To most clearly illustrate processes of inhibition and activation, difference scores were created for each type of word. Reaction time to a given word following a White face was subtracted from the reaction time to the same word following a Black face. Inhibition would thus be represented by positive difference scores (slower to words—larger response latencies—that follow Black faces), while activation/facilitation would be represented by negative difference scores (faster to words—smaller response latencies—that follow Black faces). Difference scores for stereotype-relevant and stereotype-irrelevant words were submitted to a Word type × Goal (egalitarian, respect for traditions) mixed-factorial analysis of variance (ANOVA). As predicted, the only effect to emerge was a reliable interaction, F(1,32) = 15.67, p < .001 (which also appears as a three-way interaction if difference scores are not computed, and reaction times are submitted to a Goal Status × Prime Type × Word Type mixed-factorial ANOVA; see below).

To examine the specific prediction regarding activation and inhibition, comparisons of the difference scores were made between the goal conditions for each word type. No reliable effects emerged when examining difference scores computed for control words—the difference as a function of prime type in response to these words is not significantly different between people who had egalitarian goals triggered and those who had respect for tradition goals triggered. However, for stereotype-relevant words, there is an influence of prime type on responding, and that influence is in opposite directions for people in the egalitarian versus tradition conditions. For people in the egalitarian condition, reactions times following Black faces are slower than reaction times following White faces; the reverse pattern is seen in the tradition condition (see Fig. 1). Participants with egalitarian goals have a mean difference score reliably larger (M = 86.12) than the mean difference score (M = −42.88) of participants with tradition goals, t(32) = 3.56, p < .05. These effects reflect inhibition and facilitation respectively. To further illustrate this, the difference score for stereotype-relevant words were compared against the difference scores for control words. Egalitarian goal participants had reliably slower difference scores to stereotype-relevant words compared to the response times to control words, t(16) = 4.07, p < .01. This indicates a significant slow-down in responding to stereotype-relevant words when they follow Black faces that does not occur for control words. Participants with respect for tradition goals show facilitation to stereotype-relevant words when they follow Black faces such that the difference score is significantly smaller for stereotype-relevant words compared to control words, t(16) = −2.20, p < .05.

![Fig. 1. Goal effects on response times (RTs): Stereotype-relevant words following Black faces minus stereotype-relevant words following White faces in Experiment 1.](image-url)
These analyses can also be approached by examining raw scores (opposed to difference scores). Consistent with the hypothesis, a Prime Type × Goal ANOVA revealed no reliable main effects ($p > .4$) or interaction ($p < .1$) when examining reaction times to stereotype-irrelevant words. Importantly, when examining stereotype-relevant words, the expected interaction between Goal and Prime Type was reliable, $F(1,32) = 12.67$, $p < .01$. Participants who had tradition goals undermined had response times to stereotype-relevant words that were reliably faster when they followed African American faces ($M = 646$ ms) than when they followed White faces ($M = 689$ ms), $t(16) = -2.08$, $p < .05$. Thus, as predicted, they exhibited stereotype activation. However, participants whose egalitarian goals had been undermined exhibited inhibition. Stereotype-relevant words following African American faces ($M = 746$ ms) yielded response times reliably slower than those following White faces ($M = 660$ ms), $t(16) = 2.88$, $p < .05$.

Stereotype activation and inhibition were exhibited, with the type of processing initiated dependent on the type of goal the individual had accessible at the time of the reaction time task. One potential argument against the inhibition exhibited here is that the LDT and the prime were separated in time by a full second, which is enough time for conscious control to intervene. However, during targeted questioning in the debriefing, no participants expressed any conscious intent to inhibit stereotypes on the task, nor saw any of the tasks performed during the computerized portion of the experiment as related to the egalitarian goals they had undermined earlier in the session. The reaction time task was not consciously seen as a way to address an egalitarian goal or as having anything to do with stereotyping.

A second potential argument might assert that the stereotype inhibition exhibited here is not a specific reaction to a goal. Rather than goal shielding, in which a specific compensatory response is associated with the category “African American,” participants might be instead displaying a generalized behavioral inhibition. African American primes might have triggered the behavioral inhibition system (BIS; e.g., Monteith et al., 2002). This could occur if the undesired behavior written about in the first task led participants at that time to adopt a strategy of being cautious and to respond more slowly whenever an African American is encountered. This would cause them to respond slower to items following Black faces on the computer task. This argument does not explain the current findings adequately because reaction times were not generally slowed following faces of African Americans. Reaction times to irrelevant stimuli (in both phase 2 and 1) did not differ as a function of the primes. The inhibition displayed is specific to stereotype-relevant words, suggesting a goal-specific, compensatory response.

Experiments 2A and 2B

Experiment 2 sought to replicate the goal shielding effect of Experiment 1 while directly addressing the two concerns just discussed. Rather than separating the prime and LDT tasks by a full second and describing them as two, separate tasks in the trial, the two events were described as part of one phase of a trial with a stimulus onset asynchrony (SOA) of only 100 ms.

To rule out that a generalized behavioral inhibition was being initiated as an explicit response to the failure, experiment 2 used a method of assessing inhibition in which faster (rather than slower) responses evidenced inhibition. The task was to indicate whether two words presented simultaneously were written in the same color. If the words are related in semantic content to a stereotype, and if that stereotype is accessible, the color-naming response will be slowed. If the stereotype has been inhibited, then there is reduced interference from the semantic content and the color-naming response will be facilitated. This task is modeled after the Stroop (1935) task, and slower response times indicate increased activation rather than inhibition. Goal shielding in the form of stereotype inhibition would be exhibited by the exact opposite response than expected if inhibition was due to the operation of the BIS—following African American faces responses should be faster, but only responses to specific target words (stereotypic words).

Experiment 2A

The purposes of this experiment are to introduce a new reaction time paradigm in which slower responses indicate increased accessibility while faster responses indicate inhibition and to examine whether this paradigm can be used to adequately measure stereotype activation.

Method

Research participants. Participants were 18 White students at Lehigh University who participated to partially fulfill a requirement in their introductory psychology course. All were native English speakers.

Materials

Computer. Identical to that used in Experiment 1.

Pictures used as primes. Identical to those used in Experiment 1.

Target words. The words were identical to those used in the LDT of Experiment 1. Unlike the LDT, on a given trial, words in the current task were presented in pairs. Because there are eight words of each type, four sets of randomly paired words were produced for each word type.

Reaction time task. The experiment included 18 practice trials and 96 trials. Each trial started with a photograph displayed for 100 ms. The image (a man’s face) appeared randomly in one of the four quadrants of the screen (upper left, upper right, lower left, and lower right). The participants were informed by the instructions that the picture would serve as a gaze-fixation cue for the location of a word pair. On each trial, the image was followed immediately by a pair of words in the same position. Reactions to the word pair were the focal task. Participants were asked to indicate if the two words were written in the same or different color as quickly and accurately as possible. Words remained on the screen until a response was made by pressing keys marked “different” or “same.” The position of the photograph and word pair was selected randomly on each trial. After the response was recorded, feedback information was displayed on the screen, informing the participant of his/her accuracy and reaction time.

There were 16 critical trials. Eight had pairs of stereotypic words (four following Black, four following White, faces). Eight had pairs of control words (four following Black, four following White, faces). Each word pair appeared in two trials, once following a photo of a white male, once following a photo of a black male. Critical trials were always pairs of different words in different colored font. The remaining trials had pairs of “filler words.” On filler trials, half of the trials had word pairs printed in the same color, half had word pairs in different colors.

Procedure and design. The experimenter took each participant to a small room containing a desk and a computer where they were to work on a perception task in private. Each participant was informed the task had multiple trials, with each trial consisting of a pair of words written in color. The task was to determine if the colors of the words were the same or different (a color comparison task). After all the trials had been completed, each participant was debriefed, thanked and dismissed.

The experiment had a Prime Type (Black, White) by Word Type (stereotype-relevant, control) factorial design. Prime Type and Word Type were within-participant variables.

Results and discussion

If stereotypes are activated by the faces, then people should have attention directed to words that are consistent with the stereotype.
With attention displaced from the color-naming task toward reading the words, responses to the color-naming task would be slowed. Thus, stereotype activation is evidenced by slower reaction times to stereotypic words following Black faces, but no differences between stereotypic and control words following White faces.

To most clearly illustrate activation, difference scores were created for each type of word. Reaction time to a given word-pair following a Black face was subtracted from the reaction time to the same word-pair following a White face. Stereotype activation would be represented by negative difference scores (slower to words following a Black face). Difference scores for stereotype-relevant words were reliably smaller ($M = −81.59$) than difference scores for stereotype-irrelevant words ($M = 15.34$). No reliable main effects ($t(16) = 2.74, p>.05$). No such differences as a function of prime emerged on responses to pairs of control words ($p>.23$). Reaction times were slowed to the color-naming task only when stereotypic word content followed a Black face, indicating the stereotype had been activated.

Experiment 2B

After having established the utility of the reaction time paradigm for assessing stereotype activation, the manipulation of egalitarian goals was introduced in Experiment 2B.

Method

Research participants. Participants were 38 White students at Lehigh University who participated to partially fulfill a requirement in their introductory psychology course. All were native English speakers.

Materials

Goal status manipulation. Identical to that used in Experiment 1.

Computer. Identical to that used in Experiment 1.

Reaction time task. Identical to that used in Experiment 2A.

Pictures used as primes. Identical to that used in Experiment 1.

Target words. Identical to that used in Experiment 2A.

Procedure and design. Each participant was taken to a room to work on the first task in private. They were informed that while they had agreed to participate in a perception experiment, they would first have the opportunity to perform an unrelated task to help a graduate student preparing a project examining memory for positive versus negative information about the self. When finished, they were to seal the forms in an envelope and get the experimenter. They next moved to another private room for their so-called perception experiment, where they completed a task identical to that in Experiment 2A. When finished with this task they were debriefed, thanked, and dismissed.

The experiment had a Goal (egalitarian, tradition) by Prime Type (Black, White) by Word Type (stereotype-relevant, stereotype-irrelevant) mixed-factorial design. Goal was a between-participant variable, Prime Type and Word Type were within-participant variables.

Results and discussion

Stereotype activation is evidenced by slower reaction times to stereotypic word-pairs following Black faces, but no differences between stereotypic and control word-pairs following White faces. Stereotype control is evidenced by this effect disappearing so that stereotypic words no longer displace focused attention, even when Black faces have been presented as primes.

Two independent coders rated each of the statements produced by participants on the first task following the same rating scale used in Experiment 1. Based on these ratings six participants (three in each condition) were eliminated from the analyses. Interrater reliability was high ($r = .72, p < .001$). In addition, reaction times more than three standard deviations from the mean (4.8% of responses) were considered outliers and eliminated from analyses. No participants expressed suspicions that word-pairs were related to the faces.

To most clearly illustrate activation and inhibition, difference scores were created for each type of word following the same procedure used in Experiment 2A. Difference scores for stereotype-relevant and stereotype-irrelevant words were submitted to a Word type × Goal (egalitarian, respect for traditions) mixed-factorial analysis of variance (ANOVA). As predicted, the only reliable effect was the interaction, $F(1,32) = 12.67, p<.01$.

To examine the specific prediction regarding activation and inhibition, comparisons of the difference scores were made between the goal conditions for each word type. No reliable effects emerged when examining difference scores computed for control words. However, for stereotype-relevant words, there was an influence due to one’s goals. For people in the egalitarian condition, reaction times following Black faces are faster than when following White faces; the reverse pattern is seen in the tradition condition (see Fig. 2). Participants with egalitarian goals have a mean difference score reliably larger ($M = 86.73$) than the mean difference score of participants with tradition goals ($M = −103.94$), $t(32) = 3.496, p<.01$. These effects reflect inhibition and facilitation respectively. To further illustrate this, the difference score for stereotype-relevant words was compared against difference scores for control words. Egalitarian goal participants have reliably slower difference scores to stereotype-relevant words than control words, $t(16) = 3.46, p<.01$. Participants with respect for tradition goals show facilitation to stereotype-relevant words that follow Black faces such that the difference score is significantly smaller for stereotype-relevant compared to control words, $t(16) = −2.06, p<.05$.

These analyses can also be approached by examining raw scores. Consistent with the hypothesis, a Prime Type by Goal ANOVA revealed no reliable main effects ($p>.52$) or interaction ($p>.37$) when examining reaction times to stereotype-irrelevant words. Importantly, when examining reaction times to stereotype-relevant words, the expected interaction between Goal and Prime Type was reliable, $F(1,32) = 12.21, p<.01$. Participants who had tradition goals undermined had response times to stereotype-relevant words that were slower when following African American faces ($M = 585$ ms) versus White faces ($M = 482$ ms).
Evidence for implicit control of specific stereotypes as opposed to broad inhibition

Experiments 1 and 2 provide converging evidence, using two separate reaction time procedures, that African American faces do not lead to a generalized inhibition due to the BIS, they cause (among people with egalitarian goals) specific compensatory responses in the form of inhibiting specific semantic content—stereotypes and stereotypes alone. The evidence against the BIS triggering a generalized response inhibition is the fact that such a process would be indicated by a general slow down in responding to all stimuli that followed the face of an African American. Yet, the inhibition is specific to stereotype-relevant items. Further, in Experiment 2, inhibition is evidenced by faster responding to the color-naming task, not a response slow down. Thus, the exact opposite response pattern than that predicted by a BIS effect was found—following African American (and not White) faces responses to stereotypic words (and not control words) were faster. This occurred only if egalitarian goals were induced.

The unconscious nature of this control is evidenced by the fact that participants are unaware the task has to do with stereotyping or its control and undertake no explicit strategy to exert control. In addition, the control occurs quickly, at a SOA too fast for explicit control. Fazio, Jackson, Dunton, and Williams (1995) described tasks in which responses to a stimulus are initiated within approximately a half-second of perceiving the stimulus as useful for measuring implicit processing because conscious control takes longer than this to be engaged.

The nature of inhibition: Concept incompatibility or goal shielding?

Experiments 1 and 2 illustrated the implicit inhibition of a stereotype. Is this due to an egalitarian goal? The goal conditions differed in that while each introduced a goal through a failure experience, only one introduced the concept of egalitarianism. Perhaps the concept “egalitarian” is antithetical with the concept of stereotyping and what data reveal is the inhibition that typically accompanies interconstruct incompatibility. Glucksberg, Newsome, and Goldvarg (2001) illustrated that, when a metaphor is used to refer to a person as a “pool shark,” the concept “shark” in its typical meaning as an aquatic creature is actively inhibited. The same type of inhibition of incompatible meanings between concepts is found in idiom comprehension. A concept that is used in one context with one meaning inhibits the incompatible meanings.

Interconcept inhibition is also seen in social psychology. Kunda and Thagard (1996) parallel–constraint–satisfaction theory postis inhibition of competing thoughts during concept activation. Kunda, Sinclair, and Griffin (1997, p. 721) offer an example: “Aggressive may activate both punch and argue. The context in which aggressive is activated may serve to narrow its meaning. Thus, courtroom may activate argue while deactivating punch.” Dijksterhuis and van Knippenberg (1996), as well as Macrae, Bodenhausen, and Milne (1995), illustrated such inhibition of concepts incompatible with a stereotype. Macrae et al. demonstrated that the same person can trigger the activation and deactivation of different traits dependent on what stereotype is primed. A Chinese woman viewed in a context that promotes categorizing her as a woman leads to the priming of gender stereotypes, not stereotypes of Chinese people. The same woman viewed in a context that promotes categorizing her as Chinese produces the opposite result—the stereotype of Chinese people is primed, and gender stereotypes are no longer accessible.

Perhaps in Experiments 1 and 2 participants are illustrating this functioning of the lexical system: the primed semantic concept “egalitarian” is inhibiting incompatible semantic items (the stereotype) rather than the stereotype being inhibited due to goal shielding. The next two experiments illustrate that priming the concept “egalitarian” is not sufficient to initiate inhibition. Indeed, under some conditions, priming “egalitarian” promotes stereotype activation. The next two experiments illustrate that the inhibition versus triggering of stereotypes depends upon whether the goal to be egalitarian is triggered (as opposed to mere concept activation).

Experiment 3

The hypothesis that inhibition results from priming the semantic concept “egalitarian” is pitted against a goal shielding hypothesis. Goal accessibility is differentiated from concept accessibility in a number of ways (see Förster et al., 2007; Martin & Tesser, 2009). One feature that defines accessible goals is they possess a tension state. And once a tension state is addressed, goal operations are eliminated or halted (e.g., Bargh et al., 2001; Koole et al., 1999; Liberman & Förster, 2000; Moskowitz, 2002; Spencer et al., 1998). Martin and Tesser (2009) refer to this as the “persistence until” feature of a goal—pursuit persists until the goal is successfully moved toward, at which point goal activity ceases. Successful movement toward the goal produces a state of affirmation and with affirmation goal-relevant responding shuts down (e.g., Carver & Scheier, 1981, 1998). This can be used to distinguish goal accessibility and concept accessibility, which does not have such tension states. Thus, unlike concept accessibility which fades with time (Higgins, Bargh, & Lombardi, 1985), goal accessibility persists over time if the goal is not addressed. And unlike concept accessibility which increases as one thinks about a concept, goal accessibility can decrease as one thinks about a concept if that thought resolves the tension state.

Liberman and Förster (2000) provided an illustration of this fact in an experiment in which a goal-completion procedure was shown to shut down the monitoring system thus eliminating the tension and its consequent effects. Participants with an implicitly triggered goal who were allowed to satisfy the goal showed no evidence of goal-relevant responding. When there was no opportunity to satisfy the implicit goal, participants persisted in their goal-relevant responding (see also Bargh et al., 2001). Förster, Liberman, and Higgins (2005) similarly found that there was a stronger tension state when a goal had not been fulfilled, as illustrated by participants being faster at recognizing goal-relevant words in a lexical decision task. However, after having completed the goal, these same words were responded to more slowly.

Goal shielding in stereotype control can be evidenced by utilizing this persistence until feature of goals. If the inhibition observed in our experiments is due to incompatibility between the concept “egalitarian” and semantic content associated with a stereotype, then increased accessibility of the concept “egalitarian” will trigger greater inhibition. If the inhibition observed in our experiments is due to goal shielding, a different prediction can be made. The persistence until logic argues that decreased activation of the goal “egalitarian” can be produced by an increased focus on the concept “egalitarian” if that focus is on resolving the tension associated with the goal. Thinking about egalitarianism while contemplating one’s successes at it would, ironically, heighten accessibility of the concept “egalitarianism” while simultaneously lowering the accessibility of the goal to be egalitarian. This is ironic because thinking about success at the goal would eliminate the stereotype inhibition that accompanies the goal, thus restoring stereotype activation to the very people contemplating success at being egalitarian. In contrast, thinking about failure at being egalitarian primes the goal and triggers stereotype inhibition.

The persistence until logic in Experiment 3 followed the procedure of Moskowitz (2002) to manipulate egalitarian goals versus the concept “egalitarian.” Activation of the concept “egalitarian” in half of the participants was achieved by having them contemplate a success at being egalitarian. These participants were contrasted against participants who had egalitarian goals activated by contemplating a failure at being
egalitarian. We predict that thoughts of failure should lead to goal accessibility, which persists until the goal is met, thus increasing stereotype inhibition. Thoughts of success should trigger the same concept, yet fail to trigger a goal (or tension state), and thus fail to initiate inhibition as a compensatory response. Activation and inhibition were assessed using a reaction time procedure similar to Moskowitz et al. (1999).

Method

Research participants

Thirty-eight White undergraduates participated to fulfill a requirement in their introductory psychology course. All were native English speakers.

Materials

Goal status manipulation. This was identical to Experiment 1 with one exception. Rather than half the participants describing lack of respect for traditions, they affirmed their sense-of-self as egalitarian. They described success at being egalitarian toward African Americans. The instructions asked not to focus on “fleeting moments of positive thoughts, but to reflect on instances in life when you truly performed in a way that led you to uphold the egalitarian ideal.”

Target attributes. The attributes were identical to those used in Experiment 1, with minor additions. Ten words of each type were included. This meant the addition of two stereotype-relevant words (showy, dangerous) and two stereotype-irrelevant words ( tiresome, cheap).

Pictures used as primes. Identical to those used in Experiment 1.

Apparatus. The experiment was presented on a Macintosh Powerbook 180 attached to an “Applecolor High-Resolution RGB” monitor. Participants indicated responses for the recognition measure and the lexical decision task on the Powerbook’s keyboard. The “J” key was labeled “yes,” the “F” key was labeled “no” for all participants. The program was written in MacProbe.

Reaction time task. The task was described as a memory experiment in which two images would be presented (one in a red, one in a blue, border). The image in the red border was to be memorized and compared against another image that would shortly appear to determine if they were a match. The image in the blue border was to be ignored. In addition, a lexical decision task occurred in the short interval between image presentation and the memory test.

The task began with 20 practice trials followed by 140 trials with an identical format. Each trial began with a fixation cross lasting 500 ms as a sign that the trial was about to start, followed by a blank screen for 500 ms. Two pictures were then presented simultaneously for 200 ms. One picture was in a red border, the other in a blue one, with location on the screen of each border-type (right or left of where the fixation cross appeared) randomly determined on each trial. After a 15-ms interval following the disappearance of the images from the screen, a letter string appeared. Participants were to decide if it was a word by pressing keys marked “yes” and “no” on the keypad. A 500-ms interval after the lexical decision was followed by an image. The task was to indicate whether the image was the same as the red-bordered image from that same trial by pressing keys marked “yes” and “no.” After a 3-s pause, the next trial commenced.

There were 40 critical trials, 20 where African American and 20 where White faces were presented in red borders. On these trials the letter strings were words (attributes) randomly paired with the preceding, red-bordered images (given the constraint that 10 stereotype-relevant and 10 irrelevant words had to be paired with each type of face/prime). The program randomized the order of the trials and the pairing of specific faces with specific words within each trial.

Procedure and design

Each participant was taken to a room to work on the first task in private following the procedure of Experiment 1. Next, they moved to another private room for their so-called perception experiment. Upon completion, they were debriefed, thanked, and dismissed.

The experiment had an Egalitarian Goal (affirmed, undermined) by Prime Type (Black, White) by Word Type (stereotype-relevant, stereotype-irrelevant) mixed-factorial design. Goal was a between-participant variable, Prime and Word Type were within-participant variables.

Results and discussion

The hypothesis was that African American faces would activate stereotypes of African Americans, unless one had the goal to be egalitarian, in which case the same faces would trigger inhibition. Participants affirmed in their status as egalitarian individuals were not predicted to have the goal to be egalitarian triggered and should have faster responses to stereotype-relevant words following African American versus White faces (activation). People whose status as egalitarians were undermined were predicted to have egalitarian goals triggered and should have slower responses to stereotype-relevant words following African American faces (inhibition).

Reaction times from the lexical decision task were submitted to a Goal (affirmed, undermined) × Prime (African American, White face) × Attribute (stereotype-relevant, irrelevant) mixed-factorial ANOVA. Reaction times more than 3 standard deviations from the mean (5.5% of responses) were treated as outliers and eliminated from analyses. No participants expressed suspicion that responses to the words were related to the images that preceded them.

As predicted, a reliable three-way interaction was found, F(1,36) = 5.94, p<.03. Separate Goal × Prime ANOVAs were conducted for each word type. If an egalitarian goal was able to disrupt stereotype activation, we would observe a Goal × Prime interaction in examining reaction times to stereotype-relevant words. However, no such interaction would emerge when examining responses to stereotype-irrelevant words. Consistent with the hypothesis, no reliable main effects or interactions emerged in examining reaction times to stereotype-irrelevant words.

Importantly, when examining reaction times to stereotype-relevant words, the interaction was reliable, F(1,36) = 4.90, p<.04. As predicted, participants who had their egalitarian goals affirmed exhibited stereotype activation—facilitated responses to stereotype-relevant words following stereotype-relevant, but not stereotype-irrelevant, primes. When stereotype relevant words followed African American faces (M=664 ms), response times were reliably faster than when they followed White faces (M=744 ms), t(19) = 3.53, p<.01. However, participants who had egalitarian goals undermined failed to show this facilitation effect. For these participants, when stereotype relevant words followed African American faces (M=738 ms) response times did not reliably differ from those following White faces (M=758 ms), t(19) = 1.35, p = .19.

In addition, participants did not reliably differ in responses to stereotype-relevant words following White faces as a function of whether egalitarian goals were affirmed or undermined (F = .18, p = .68), but responded reliably faster to stereotype-relevant words following African American faces when egalitarian goals had been affirmed versus undermined (F = 7.8, p < .01).

In summary, people who contemplated successfully being egalitarian did not have egalitarian goals regulating their responding—a fairly typical, stereotype activation effect emerged when a Black face was seen. In contrast, people who had contemplated failure at egalitarianism had a goal activated and displayed the evidence of goal shielding. These individuals controlled stereotype activation, regulating goal pursuit in the preconscious.

In this regard, Experiment 3 replicates a point previously demonstrated by Aarts and Dijksterhuis (2003) and Moskowitz (2002).
Exposure to a goal concept does not necessitate there be downstream consequences on behavior and cognition. Simply thinking about the goal of being egalitarian may trigger the semantic meaning associated with it, but goal-relevant responses are only initiated when there is an energy to pursue the goal triggered as well. Thoughts of success at being egalitarian did not supply this energy required for compensatory cognition to occur. It also did not trigger inhibition of stereotypes as the parallel–constraint–satisfaction theory could predict (if these concepts were semantically linked as oppositional constructs). Instead, stereotype activation emerged when African American faces were detected by the perceptual system.6

The procedure used in this experiment, unlike the previous two experiments, did not find evidence for active inhibition. Stereotype activation was controlled when egalitarian goals were accessible, but reaction times were not slowed to the stereotype-relevant words following African American faces relative to White faces. There are several possible reasons for this. The competing goal of memorizing faces was active as participants responded to the LDT. The goal of memorizing faces could have interfered with the inhibitory processes associated with the goal of being egalitarian. Or, it could be that these goals are not incompatible, but that the critical LDT, being an intermediary task performed while awaiting an upcoming task, had variance in responding introduced that masked the ability to detect any inhibitory processes that occurred. Finally, it could be that negative variance in responding introduced that masked the ability to detect any inhibitory processes that occurred. In this procedure, the target faces were the stimuli being attended to in the memory task. Moskowitz et al. (1999) had the reaction time task. In this procedure, the target faces were the stimuli associated with the goal of being egalitarian. Or, it could be that these goals are not incompatible, but that the critical LDT, being an intermediary task performed while awaiting an upcoming task, had variance in responding introduced that masked the ability to detect any inhibitory processes that occurred. Finally, it could be that negative variance in responding introduced that masked the ability to detect any inhibitory processes that occurred. In this procedure, the target faces were the stimuli being attended to in the memory task. Moskowitz et al. (1999) had the target faces in the blue borders, not being attended to (but ignored). Perhaps inhibition would be found if such a procedure had been used. Since Experiment 1 relied on a procedure where participants were to ignore faces, and that procedure proved reliable for showing inhibition, Experiment 4 returned to that procedure.

**Experiment 4**

Experiment 4 uses the persistence until logic once again, this time first introducing the goal to be egalitarian to all participants. In this way, the goal is in place, and it will subsequently either persist or be shut down as a function of the task that follows. Failure to adequately address the goal will lead the goal to persist and stay accessible, resulting in stereotype inhibition. In contrast, goal attainment would eliminate the tension and shut down goal shielding, thus causing stereotype inhibition to cease (and stereotype activation to occur). This allows for a stronger test of the goal shielding hypothesis than the manipulation used in Experiment 3 because the competing hypothetical process—concept activation—makes a diametrically opposing prediction. It predicts that once the concept “egalitarian” is primed, this activation fades if one subsequently does not think about the concept. And with it fades inhibition of the opposing concepts (stereotypes). Conversely, accessibility of the concept ‘egalitarian’ should increase, as would inhibition, if one subsequently thinks about the concept. Goal shielding logic dictates the exact opposite effects of thinking, versus not thinking, about success at egalitarianism, as described above.

To control for the fact that thinking about success at being egalitarian results in a self-affirmation as well as results in goal attainment, participants who do not attain the goal also contemplate a self-affirmation; however, it is in a domain irrelevant to egalitarianism. Thus, each group of participants is given an egalitarian goal, and each then thinks about an affirmation. What is manipulated to induce goal completion versus incompletion is whether the affirmation is specific to egalitarianism or irrelevant to egalitarianism. Such a procedure will also allow for a test of whether affirmation processes in general will alleviate the tension state associated with a goal or whether the goal persists until the specific goal-related discrepancy has been addressed.

**Method**

**Research participants**

Participants were 36 White students at Lehigh University who participated to partially fulfill a requirement in their introductory psychology course. All were native English speakers.

**Materials**

**Goal status manipulation.** The packet of materials began identically to that used in Experiment 1—a set of values to be rated with Likert scales was followed by the task of describing a failure at pursuing one of the items from the set. All the participants, in a fashion identical to that followed by half the participants in Experiment 1, describe a failure at being egalitarian. Following this task, all participants experience the positive affect associated with affirming the sense of self but are divided into two groups. Half of the individuals next write about success at being egalitarian in a fashion identical to Experiment 3. This group has had their egalitarian goal affirmed. The remaining half write about success at respecting traditions (as described in Experiment 1). This group has an egalitarian goal that is still incomplete.

**Computer.** Identical to that used in Experiment 1.

**Reaction time task.** Identical to that used in Experiment 1.

**Pictures used as primes.** Identical to that used in Experiment 1.

**Target words.** Identical to that used in Experiment 1.

**Affect measure.** Participants indicated how they felt after completing the tasks on the Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988). Feelings (e.g., upset, distressed, proud) were rated on a 5-point scale (1 = very slightly or not at all; 5 = extremely).

**Procedure and design**

Each participant was taken to a room and told they would work on the first task in private following the procedure used in Experiment 1. Next, they moved to another private room for their so-called perception experiment. Before the task began, they completed the affect measure and then the computerized task. When finished, they were debriefed, thanked, and dismissed.

The experiment had a Goal (incomplete, affirmed) by Prime Type (Black, White) by Word Type (stereotype-relevant, stereotype-irrelevant) mixed-factorial design. Goal was a between-participant variable, Prime Type and Word Type were within-participant variables.

**Results and discussion**

If stereotype inhibition is due to incompatibility between the concept “egalitarian” and concepts associated with African American stereotypes, then continued activation of the concept “egalitarian” will maintain inhibition. Conversely, decreased accessibility of the concept “egalitarian” (such as when thoughts are focused on concepts unrelated to egalitarianism) will decrease inhibition and stereotype activation will return. Goal shielding predicts an opposite pattern. Thoughts about success at being egalitarian will eliminate goal accessibility and goal shielding (and inhibition). However, if thoughts are unrelated to egalitarianism, the egalitarian goal will not fade. And, as part of the goal shielding process, stereotype inhibition persists. Two independent coders rated statements made by participants on the first task. People failing to adequately describe a failure or success

---

6 This was not a “licensing effect” as described by Monin and Miller (2001) since participants did not know the reaction time task afforded them an opportunity to stereotype. They were also not afforded the time to make a conscious response.
Consistent with the hypothesis, a reliable three-way interaction goals, difference score (a mean difference score (activation) pattern is seen among people whose egalitarian goals had been inhibition displayed in incomplete (and still accessible). In the incomplete condition, there is no reliable effects emerged when examining differences scores computed for control words. The difference as a function of each word type. However, for stereotype-difference scores (faster to words—smaller response latencies—that follow Black faces), Difference scores for stereotype-relevant and -irrelevant words were submitted to a Word type × Goal (incomplete, affirmed) mixed-factorial ANOVA. As predicted, a reliable interaction emerged, F(1,31) = 11.50, p < .01. To examine the predictions regarding activation and inhibition, comparisons of difference scores were made between goal conditions for each word type. No reliable effects emerged when examining difference scores computed for control words. The difference as a function of prime type in response to these words is not significantly different between people who had egalitarian goals affirmed and those for whom the goal was incomplete (still accessible). However, for stereotype-relevant words, there was an influence of prime type on responding and that influence is in opposite directions for people whose goals were affirmed (and thus shut down) versus those whose egalitarian goals were incomplete (and still accessible). In the incomplete condition, there is inhibition displayed—reactions times following Black faces are slower than reaction times following White faces. The reverse (stereotype activation) pattern is seen among people whose egalitarian goals had been affirmed (see Fig. 3). Participants with an incomplete egalitarian goal have a mean difference score (M = 55.02) reliably larger than the mean difference score (M = −39.56) of participants with affirmed egalitarian goals, F(1,31) = 34.2, p < .01. These analyses can also be approached by examining raw scores. Consistent with the hypothesis, a reliable three-way interaction emerged, F(1,31) = 11.50, p < .01. Goal by Prime Type ANOVAs were conducted for each word type. As predicted, the interaction examining reaction times to stereotype-irrelevant words is not reliable (p > .79). Importantly, for reaction times to stereotype-relevant words, the expected interaction between Goal and Prime Type was reliable, F(1,31) = 34.82, p < .01. Participants whose egalitarian goals were affirmed had response times to stereotype-relevant words that were reliably faster when they followed African American (M = 488 ms) versus White faces (M = 528 ms), t(15) = −3.96, p < .01. This facilitated reaction time to words following Black faces does not occur for control words (p > .8). In addition, response times to stereotype-relevant words following Black faces (488 ms) were reliably faster than reaction times to control words (515 ms) following Black faces, t(16) = −4.58, p < .01. Thus, as predicted, these participants exhibited stereotype activation.

However, people whose egalitarian goals had not been affirmed (incomplete participants) exhibited inhibition. Stereotype-relevant words following African American faces (M = 562 ms) yielded response times reliably slower than those following White faces (M = 507 ms), t(16) = 4.44, p < .01. In addition, response times to stereotype-relevant words following Black faces (562 ms) were reliably slower than reaction times to control words (526 ms) following Black faces, t(16) = −4.58, p < .01. Comparing across goal conditions, the only comparison where reaction times differ is to stereotype-relevant words following Black faces. There are no differences for stereotype-relevant words following White faces, or stereotype-irrelevant words following either face type (p’s > .34). Thus, as predicted, these participants inhibited stereotypes.

These results illustrate that affirmed participants have stereotypes activated upon detecting a Black face, whereas incomplete participants have stereotypes inhibited. This occurs despite the fact that affirmed participants have thought more about the concept “egalitarian” and have stronger accessibility for the concept. The incomplete participants inhibit despite not having consciously thought about egalitarianism since the initial stage of the experiment. The pattern of findings suggests the manipulations trigger a goal that is then either shut down or accessible. As such, goal shielding is either shut down or remains active. This occurs without awareness that the reaction time task is relevant to stereotyping, without being conscious a goal is accessible during the task.

These findings also replicate an important point about the nature of goal affirmation recently reported by Moskowitz et al. (in press). What type of success “shuts down” compensatory responding? Self-compensation theory (Wicklund & Gollwitzer, 1982) suggests the affirmation must occur in the specific domain in which a discrepancy was detected. However, self-affirmation theory (Steele, 1988) says any restoration of the self-system, an increase in self-esteem, should resolve the tension, portrayed instead as a threat to the global self-system. Our findings reveal that the tension associated with being non-egalitarian to Black men is not satisfied by global affirmation, but only by affirming the self-as-egalitarian. This is not to say global self-affirmation does not resolve tension states. However, we would argue this would only occur if one is pursuing the goal of affirming the global sense of self. Research illustrating the successful resolution of tension states resulting from processes of global affirmation (e.g., Koole et al., 1999; Spencer et al., 1998; Steele, 1988) typically create tension by challenging self-esteem at a somewhat broad level (providing feedback suggesting one is not intelligent enough to lead a successful life, or that one is a bad person, or not well received by others). Devastations to the global self-system such as these would be resolved (goal completion) by restorations of that

Fig. 3. Goal effects on response times (RTs): Stereotype-relevant words following Black faces minus stereotype-relevant words following White faces in Experiment 3.
global self-system. However, more specific goals should not be resolved by global affirmations.

This finding is consistent with the research of Galinsky, Stone, and Cooper (2000) who found that the tension associated with cognitive dissonance was only seemingly reduced by a global self-affirmation. That tension state (and dissonance) was easily reintroduced if one's global affirmation strategy was subsequently challenged. Participants then adopted a strategy of goal completion—addressing the specific threat directly. Perhaps the ease of regenerating the dissonance is a sign that the tension had never really dissipated at the implicit level and could have been detected by implicit measures such as the reaction time tasks we employ. Our results reveal that below conscious awareness, such tensions are not resolved, as Galinsky et al. (2000) suggest.

General discussion

Despite the fact stereotyping often occurs outside of awareness, so too do processes of control. Stereotyping may not require one to think (consciously), but one also need not think (consciously) to control it. Indeed, the logic of the ubiquity of stereotype activation is turned on itself. Because stereotype activation is an unconscious response that is in the service of one type of goal (one where stereotyping is required), it is controlled by exercising the same self-regulatory system that makes stereotype activation efficient. These experiments illustrate such a proactive form of stereotype control—preventing a stereotype from ever being activated upon categorizing a person as a member of a stereotyped group. One can control thought, even low-level processes of activation/inhibition (even if one is not consciously aware of the goal).

Shifting from types of people to all people

Since the so-called “New Look” (e.g., Bruner, 1957), it has been argued that goals direct what is perceived, even when the items being perceived are subliminally presented. However, the “New Look” and related work focused on chronic states, not goals that are adopted in the moment (even unknowingly primed in the moment). Prior research extended that approach to stereotyping, showing that chronic goals determine if stereotypes are activated or inhibited (Moskowitz et al., 1999). Our experiments move the question away from such individual differences to ask whether any person can adopt a goal to inhibit stereotypes without awareness or conscious intent to inhibit the stereotype at the time such inhibition occurs.

We apply goal shielding logic to control of stereotype activation. People are associated with many goals, some compatible and some incompatible to each other. Some goals promote stereotyping, some oppose it. An African American man provides one with the opportunity to be fair and egalitarian toward a member of a stereotyped group. He also represents an opportunity to pursue an alternative goal of quickly and efficiently categorizing him, a goal that may involve lower-order goals of recruiting stereotypes to ease the process. Each of these goals—in-conflict can be triggered, each triggering stereotype-relevant processing outside of awareness.

Our research finds that when egalitarian goals are accessible, the presence of an African American man triggers that goal and its associated operations. This will include inhibition of incompatible goals and stereotypes (as well as heightened accessibility of the egalitarian goal, e.g., Moskowitz, 2002, and readiness to detect goal-relevant people in the environment, e.g., Moskowitz et al., in press). An individual can control stereotyping without knowing a stereotype or a goal exists. Consciousness is not required. One’s wants, even implicit wants, can direct thoughts. Consistent with research outside the domain of stereotyping (e.g., Bargh & Huang, 2009; Dijksterhuis et al., 2006; Sassenberg & Moskowitz, 2005), it is possible that at times the implicit pursuit of one’s goals is even more efficient and successful than the conscious pursuit.

Not all egalitarian goals are equal

The current experiments required participants to define egalitarian as antithetical to stereotyping by forcing them to write essays in which they reported having acted in a stereotypic fashion. However, not all people will, outside of laboratory manipulations that control the definition of the goal, associate egalitarianism as necessarily being antithetical with stereotypes. Two general forms of egalitarian goals are represented in the culture. One stresses egalitarianism being defined by explicitly paying no mind to culture, race, religion, ethnicity, etc. Color-blind is the term used to describe strivings marked by evaluating people based on merit with avoidance of categories. A second stresses egalitarianism being defined by diversity. Multiculturalism is the term used to describe egalitarian strivings marked by cultural and ethnic diversity being the desired end (Richeson & Nussbaum, 2004). These two forms of egalitarian goal pursuits are often at odds in terms of the cognitive and behavioral means they suggest be implemented en route to fairness (Vorauer, Gagnon, & Sasaki, 2009). Multiculturalism requires one take note of group membership while a color-blind approach requires ignoring it. Multicultural and color-blind goals thus impact stereotyping differently because of these differing ends that are specified.

Multicultural goals lead perceivers to attend to group membership, categorize others into groups, and draw on stored representations about the group—stereotypes. This goal, however, also involves recognizing and valuing positive group differences. It requires open-mindedness and for one to consider a wide range of qualities—positive as well as counter-stereotypes—that stand in contrast to the traditional negative stereotypes. Multicultural goals, rather than being antithetical with stereotyping, control stereotype activation by facilitating it. Wolsko, Park, Judd, and Wittenbrink (2000) found such effects. In a task where participants estimate what percentage of Whites and Blacks possess certain attributes, participants in the multicultural group stereotyped (both positive and negative) Blacks more than Whites, while there was no difference between the color-blind and control groups. They also found the multicultural group had a more accurate understanding of the differences that exist between Blacks and Whites.

Moskowitz and Ignarri (2009) suggest that because color-blind goals require one to ignore categories, this approach narrows the mind-set with which a perceiver approaches interpersonal perception, and should thus be able to control stereotype activation in two distinct ways. First, by preventing a category from ever being activated (and hence no activation of the stereotype linked to the category). Second, by the category being activated, and that category triggering fairness goals that inhibit, rather than activate, stereotypes through the goal shielding processes we have here illustrated. These two forms of control are reviewed next.

Control despite categorization versus control by blocking categorization

Stereotype activation involves the intermediary step of categorizing the person as “a person” and as a member of a group to which the stereotype is associated. Categorization itself is a multitstep process in which preattentive processes locate features, and those features then get matched against existing categories. Livingston and Brewer (2002) argued that not even the categorization of people into relevant “person groups” must occur automatically but requires one has a goal. Further, people are complex and multifaceted targets of perception, and as such, the way one categorizes (which category is selected) is goal-based. For example, giving a perceiver the goal of reacting to a person in terms of race inhibits categories such as gender and occupation (e.g., Dijksterhuis & van Knippenberg, 1996; Macrae et al., 1995). In contrast, goals that promote categorizing to occupational groups may promote lawyer, professor, and fireman stereotypes but inhibit racial stereotypes also associated with, for example, Black firemen (Taylor, 1981). Goals can promote category use, but in so doing inhibit competing categories.
The main point of our experiments is not "goals specify which categories, from among many possible categories, are triggered." Nor is it that "stereotypes of Blacks can be blocked if one categorizes the person instead according to gender or occupation." It is that given a specific category is activated, people may be perceived in a fashion that does not imply a stereotype associated with that category if goals inhibit such processing. Stereotype control is accomplished even when categorization has occurred, despite categorization; goals direct processing such that stereotypes are inhibited and other knowledge associated with the category is retrieved instead.

Conclusion

Stereotype control is not only something we consciously perform to overt, or react to, unwanted thoughts. It is something we proactively engage, outside of conscious awareness, to help produce desired cognition in the first place, even inhibiting unwanted thoughts before they occur. This distinction between reactive and proactive control is perhaps more recognizable outside the domain of psychology. For example, British Petroleum recently took a reactive strategy to its drilling operations in the Gulf of Mexico by ignoring several preventative measures that could have been undertaken in advance of an accident. Instead, they waited for an oil spill and then applied attempts at control, perhaps only after substantial damage had already accrued. A proactive strategy would have anticipated problems with drilling and eliminated the likelihood of a spill ever happening. Proactive strategies are difficult for people to contemplate because they ultimately result in a negative event; success at the goal is nothing happening (such as a global computer meltdown not happening due to Y2K preparations, a deep depression being avoided because of economic bailouts, terrorist acts failing to materialize from proactively passing and enforcing antiterrorism laws, and oil spills at rigs that never happen).

Returning to the psychology of stereotyping, both approaches are viable, each with limitations. Reactive strategies require being aware of stereotypes, yet stereotypes are hard to detect even when one is not motivated to deny them exist. They also require knowing how to eliminate bias and having the desire and ability to implement that strategy, with each of these components often lacking (e.g., Wilson & Brekke, 1994). Proactive strategies require commitment to a goal, which may not always be present. They also are at the mercy of the environment and the multitude of competing goals that might get triggered and rotate the egocentric goal out of one's current concerns (e.g., Shah, Hall, & Leander, 2009), however, the possibility for control is even more pervasive and efficient than researchers have often assumed. And in the case of stereotyping, proactive control produces a negative event (no stereotype activation) that is easy to contemplate because it can be seen. It does not result in nothing happening, but something measurable in the form of stereotype inhibition and other forms of goal shielding.

Just as with stereotype activation, stereotype control is functional. It has a purpose, or it serves a goal. Not stereotyping can be just as motivating and cognition: Foundations of social behavior, Vol. 2. (pp. 93–130). New York: The Guilford Press.


