

# Stereotype Threat and Performance: How Self-Stereotypes Influence Processing by Inducing Regulatory Foci

Beate Seibt  
Universität Würzburg

Jens Förster  
International University Bremen

The authors hypothesized that activated self-stereotypes can influence the strategies of task solution by inducing regulatory foci. More specifically, positive self-stereotypes should induce a promotion focus state of eagerness, whereas negative stereotypes should induce a prevention focus state of vigilance. Study 1 showed that a negative ascribed stereotype with regard to task performance leads to better recall for avoidance-related statements whereas a positive stereotype leads to better recall for approach-related statements. In Studies 2 and 3, both an experimental manipulation of group performance expectation and the preexisting stereotype of better verbal skills in women than in men led to faster and less accurate performance in the positive as compared with the negative stereotype group. Studies 4 and 5 showed that positive in-group stereotypes led to more creative performance whereas negative stereotypes led to better analytical performance. These results point to a possible mechanism for stereotype-threat effects.

“Blonds are dumb”; “women can’t do math”; “White men can’t jump.” If you bring members of a stereotyped group into an achievement situation where the stereotype could be applied, they often dramatically underperform if they believe their ability in that particular domain is measured. This was first shown by Steele and Aronson (1995) in their seminal work on the effect of stereotype activation on intellectual test performance of Blacks. In a series of studies, Blacks performed considerably poorer on a standardized test when it was presented as diagnostic of their abilities than when the same test was framed as a simple problem-solving task. No such difference emerged for Whites. This basic finding, named *stereotype threat*, has since been replicated several times with different groups and various stereotypes (e.g., Aronson et al., 1999; Croizet & Claire, 1998; Inzlicht & Ben-Zeev, 2000; Levy, 1996; Shih, Pittinsky, & Ambady, 1999; Spencer, Steele, & Quinn, 1999; Stangor, Carr, & Kiang, 1998).

What is the psychological mechanism behind this pervasive effect? According to the stereotype-threat theory formulated by

Steele (1997), performance suffers when negative stereotypes are activated in a performance situation and when the performance domain is self-relevant for the individual. The theory assumes that this underperformance stems from the keenly felt threat triggered by the possibility of conforming to the stereotype or of being treated and judged in terms of it. The experienced threat then undermines performance by setting up a frame for interpreting any performance frustration in the light of the stereotype. However, the search for mediators of stereotype-threat effects has yielded some inconsistent results (cf. Steele, Spencer, & Aronson, 2002).

Given the various aspects of stereotypes, it seems likely that several mechanisms contribute to the described performance deficits. In this article, we propose a model according to which the activation of negative as well as positive self-stereotypes affects performance strategies by inducing regulatory foci (Higgins, 1997). We argue that negative stereotypes induce a prevention focus and thus instigate an avoidance strategy, in particular avoiding errors of commission, and that positive stereotypes induce a promotion focus, characterized by an approach strategy, in particular approaching hits. This means that the activation of negative stereotypes should foster accuracy whereas the activation of positive stereotypes should foster speed. In addition, regulatory foci entail certain cognitive styles, which should, according to the proposed model, influence task performance under activated stereotypes. In particular, a prevention focus leads to a risk-averse, analytic, and perseverant processing style whereas a promotion focus leads to an explorative and more creative processing style (Friedman & Förster, 2001). Consequently, when a negative stereotype is activated, performance in creative tasks should become poorer and performance in analytic tasks better than when no stereotype is activated. Conversely, when a positive stereotype is activated, performance in creative tasks should become better and performance in analytic tasks poorer than when no stereotype is activated. These predictions are developed in greater detail below.

---

Beate Seibt, Lehrstuhl für Psychologie II, Universität Würzburg, Würzburg, Germany; Jens Förster, School of Humanities and Social Sciences, International University Bremen, Bremen, Germany.

This research was supported by Deutsche Forschungsgemeinschaft Grant FO 244/9-1 to Jens Förster. The model was introduced during the habilitation hearing of Jens Förster in 2000 at Universität Würzburg. We thank Jeannine Dreyßig, Monika Ott, Tatjana Taraszow, Sabine Thoma, Julian Wucherpfennig, Maren Breuer, Katarzyna Kubacka, Florian Albert, Nina Burger, Sebastian Karban, Kirils Jegorovs, Janina Marguc, Alan Langus, and Joanna Styczynska, who served as experimenters. Fritz Strack, Thomas Schubert, Roland Neumann, and Michael Häfner are acknowledged for invaluable suggestions and discussions. Special thanks go to Markus Denzler for invaluable discussions. We also thank Dave Kenny for his methodological advice.

Correspondence concerning this article should be addressed to Beate Seibt, Universität Würzburg, Lehrstuhl für Psychologie II, Röntgenring 10, 97070 Würzburg, Germany. E-mail: seibt@psychologie.uni-wuerzburg.de

### Stereotype Activation and Self-Regulation

Regulatory focus theory (Higgins 1997, 1999) distinguishes between two kinds of goal pursuit that vary as a function of regulatory focus: concerns with attainment of aspirations and accomplishments (*promotion focus*) and concerns with responsibilities and safety (*prevention focus*). These distinct regulatory concerns can be emphasized either chronically or momentarily. To illustrate, teacher–student interactions can chronically emphasize goal pursuit with either promotion focus concerns or prevention focus concerns. The teacher’s messages can be communicated in reference to a state of the student that does or does not meet promotion concerns, either “this is what I ideally would like you to do” or “this is not what I ideally would like you to do,” respectively. In this case, the regulatory focus is one of promotion, a concern with advancement and accomplishment, hopes, and aspirations (ideals). The students focus on the presence or absence of positive outcomes. Strategically, within a promotion focus, individuals are eager to approach matches to the desired end state (i.e., pursue all means of advancement). Students in teacher–student interactions that involve a prevention focus are rewarded for avoiding potential dangers or misbehaviors and punished for being irresponsible or careless. Here, positive and negative emotions are experienced as the absence or presence of negative outcomes, respectively. The teacher’s messages are communicated in reference to a state of the student that does or does not meet some prevention concerns, either “this is what I believe you ought to do” or “this is not what I believe you ought to do,” respectively. The regulatory focus is one of prevention, a concern with protection and safety, duties, and responsibilities (oughts). Strategically, within a prevention focus, individuals are vigilant to avoid negative outcomes.

As Higgins (1997) detailed in his model, any communication conveying gain or nongain information should induce a promotion focus whereas communications about nonloss or loss should induce a prevention focus. For example, feedback messages such as “you did not do badly” introduce badness as the reference point and establish that the goal of not being bad has been met. A goal not to do badly is a minimal goal because it specifies the minimal expected performance a person tries to exceed. Not meeting a minimal goal is thus a negative event, and meeting it is a non-negative event. According to Brendl and Higgins (1996), when feedback introduces a negative reference point (e.g., loss, failure, or badness), an individual will tend to set a minimal goal in the next task. A focus on minimal goals is a situationally induced prevention focus and strategically involves avoidance means.

When a negative self-stereotype (such as “your in-group is worse in doing *X* than the out-group”) is activated, it should establish a negative reference point and thereby lead to adopting a minimal goal. To illustrate, when a negative stereotype is activated, confirming the stereotype would be a loss whereas disconfirming it would be a nonloss. Such a situation induces a prevention focus, as reviewed above. It orients the individual toward avoiding mismatches to the desired end state. Activating negative stereotypes should thus induce a prevention focus state of vigilance.

A regulatory focus perspective also allows conclusions for the activation of positive stereotypes. Just as the feedback “you didn’t do badly” introduces a negative reference point, the feedback “you

did well” introduces a positive reference point and establishes that the goal of doing well has been met. A goal of doing well is a maximal goal in that it separates the positive event of doing well from the nonpositive event of not doing well (Brendl & Higgins, 1996). Positive reference points orient individuals toward gains and nongains and should thus induce a promotion focus state of eagerness with a strategic approach tendency. Positive stereotypes, just like feedback, introduce a positive reference point. For example, the stereotype “women are good at verbal tasks” specifies a positive outcome (or a gain) to be approached. It should thus orient women toward approaching matches to this desired end state. This approach orientation is the defining feature of a promotion focus. Hence, positive stereotypes should induce a promotion focus. Let us consider the implications of this theoretical notion for memory, speed–accuracy, and thinking styles.

### Recall of Approach Versus Avoidance Information

Although the main focus of this article is on the strategic consequences of regulatory foci for task performance, we first sought to get more direct evidence for the idea that the activation of stereotypes does indeed induce different foci. The best indicator for a prevention focus is the inclination to use avoidance strategies. Likewise, the best indicator for a promotion focus is the inclination to use approach strategies. These inclinations should lead to higher sensitivity to and better recall for episodes exemplifying the corresponding strategies. Accordingly, a prevention focus has been shown to enhance incidental recall for avoidance strategies and a promotion focus for approach strategies (Higgins, Roney, Crowe, & Hymes, 1994; Higgins & Tykocinski, 1992). Thus, selective memory performance for approach–avoidance information is one possibility to examine whether differentially valenced stereotypes induce regulatory foci. Therefore, Study 1 tested the prediction that induced positive stereotypes lead to better recall of information exemplifying approach strategies whereas induced negative stereotypes lead to better recall of information exemplifying avoidance strategies, using the paradigm developed by Higgins et al. (1994).

### Speed and Accuracy of Performance

In general terms, both foci involve different motivational orientations: Whereas individuals in a promotion focus, with their inclination to approach matches to their goal, are eager to attain advancements and gains, individuals in a prevention focus, with their inclination to avoid mismatches, are vigilant to ensure safety and nonlosses. In signal detection terms (e.g., Green & Swets, 1966; Tanner & Swets, 1954), individuals in a state of eagerness resulting from a promotion focus are motivated to ensure hits and ensure against errors of omission (i.e., a lack of accomplishment). In contrast, individuals in a state of vigilance resulting from a prevention focus are motivated to ensure correct rejections and ensure against errors of commission (i.e., making a mistake). Accordingly, Crowe and Higgins (1997) found in a recognition task that individuals in a promotion focus had a risky bias to say “yes,” whereas individuals in a prevention focus had a conservative bias to say “no” (see also Friedman & Förster, 2001).

In sum, promotion focus emphasis on strategic eagerness should lead to a riskier processing style that is concerned with getting hits,

whereas prevention focus emphasis on strategic vigilance should lead to a more careful processing style concerned with avoiding mistakes. Thus, in tasks where speed is in the service of hits, one would expect participants in a promotion focus to be faster and at the same time less accurate and participants in a prevention focus to be more accurate yet slower. This difference is due to strategic concerns with gains and nongains (promotion) versus losses and nonlosses (prevention) rather than due to a built-in trade-off. These predictions were confirmed in a recent research program (Förster, Higgins, & Taylor Bianco, 2003). In three studies, participants with a chronic or situationally induced promotion focus showed faster performance and less accuracy in simple drawing tasks compared with participants with a prevention focus. In a fourth study, the results were replicated with a situationally induced regulatory focus using a more complex proofreading task. The authors found that promotion focus led to faster proofreading whereas prevention focus led to higher accuracy in finding the more difficult errors. Note that the effects of speed and accuracy in the studies were independent from one another, so that speed did not necessarily lead to deficits in accuracy or vice versa. This finding suggests that when individuals have strong promotion focus concerns, they naturally are eager for hits, and when they have strong prevention focus concerns, they naturally are vigilant against mistakes.

Consequently, in tasks where speed is in the service of hits, participants under negative stereotypes should be slower than participants in a control group, whereas participants under positive stereotypes should be faster. For accuracy, the reverse pattern is to be expected: Here, negative stereotypes should lead to more accuracy and positive stereotypes to less. Study 2 tested these hypotheses with an existing stereotype, and Study 3 tested them with an induced stereotype.

It was further hypothesized that expectancy and general motivation would not mediate the effects of stereotype activation on speed and accuracy of performance. The regulatory model predicts that activated stereotypes, through inducing a regulatory focus, influence specific kinds of motivation and strategic tendencies and not general motivation. Of course, activating stereotypes could change participants' motivation or expectancies and thereby performance. For example, when judging one's performance expectation, the self could be assimilated to the group standard or, under certain conditions, it could be contrasted away from it (Mussweiler, 2003). Empirically, some stereotype-threat studies have found an assimilative effect of stereotype activation on performance expectancies (Cadinu, Maass, Frigerio, Impagliazzo, & Latinotti, 2003; Stangor et al., 1998), whereas others have not (Spencer et al., 1999; Stone, Lynch, Sjomeling, & Darley, 1999). According to Higgins (1997), regulatory focus and expectation both influence self-regulation independent from each other. Therefore, if activated stereotypes influence performance expectation or motivation, these should not affect speed and accuracy the same way regulatory focus does. In particular, if lowered expectancy or lowered motivation results from activated negative stereotypes, both speed and accuracy should drop.

### Creativity and Analytic Thinking

Recently, Friedman and Förster (2001) predicted that the risky explorative processing style elicited by promotion cues relative to

the risk-averse, perseverant processing style elicited by prevention cues would facilitate creative thought. The process by which promotion focus facilitates creative thought is assumed to involve cognitive tuning (Clore, Schwarz, & Conway, 1994). Specifically, activation of a promotion focus, a focus on nurturance, may be seen as signaling that the environment is prospectively benign, thereby leading to adoption of a riskier, more explorative processing style and bolstering creativity. Participants in their experiments were primed with either an approach–promotion maze or an avoidance–prevention maze. In both conditions, Friedman and Förster (2001) depicted a cartoon mouse trapped inside the maze and asked participants to find the way for the mouse by drawing the path. In the promotion cue condition, a piece of Swiss cheese was depicted as lying outside the maze. It was predicted that approaching food would activate a promotion focus (see Förster, Higgins, & Idson, 1998). The other half of their participants had to solve the very same maze with one exception: Instead of the cheese, an owl was hovering above the maze, presumably ready to capture the mouse unless it could escape the maze. Completion of this task was posited to activate a prevention focus. The authors could show that after having completed the promotion maze participants performed better in creative generation tasks (such as creative generation of uses for a brick) than participants who performed the prevention maze. Moreover, the promotion maze bolstered memory search for novel responses compared with the prevention maze, whereas the prevention maze led to an attentional perseverance on initial responses (or their associates). Finally, participants with a chronic promotion focus were better at solving the Gestalt Completion Task (Ekstrom, French, Harman & Dermen, 1976) than participants with a chronic prevention focus. This task is supposed to measure a classic component of creative thinking, namely, breaking context-induced mental sets (Schooler & Melcher, 1995). Moreover, Friedman and Förster (2000) showed that an approach motor action (arm flexion) led to more inclusive categorization than an avoidance motor action (arm extension), as reflected in accepting an atypical member (e.g., a camel) for a category (e.g., vehicles). To summarize, it seems that a promotion focus enhances the diverse elements of creativity as known from the literature, whereas a prevention focus impedes them.

However, given that performance outcomes depend on the fit between the strategic inclinations and processing styles of an individual's regulatory focus and the requirements of the task, performance on some tasks should profit from prevention focus vigilance. Such tasks are analytic tasks, because analytic tasks can be solved with conventional algorithms that worked in the past. These algorithms are reliable and approved and therefore represent conservative means to solve a problem. More heuristic thinking in creative tasks leads to solutions that are innovative and might be wrong. Prevention focus is assumed to facilitate systematic processing through cognitive tuning (Clore et al., 1994), in particular through enhancing bottom-up processing with reliance on the information given (see Friedman & Förster, 2001). This is because in problematic situations, novel and untested alternatives can make a bad situation even worse, and individuals feel they should rely on well-tested conventional solutions.

Evidence for this reasoning is provided by Friedman and Förster (2000, Experiment 7) as part of the above-mentioned series of studies on the effects of approach and avoidance motor actions.

These authors found that when a motor action associated with avoidance orientation—arm extension—is executed, more correct solutions for four problems from the analytic section of the Graduate Record Examination (GRE) were given than when an approach-related bodily action—arm flexion—was executed. Therefore, for performance in analytic tasks, improved performance under negative stereotypes is predicted. To summarize, we reasoned that when under stereotype threat, analytical problem solving should be enhanced, whereas creative thinking should be impeded. Conversely, when positive stereotypes are activated, analytical problem solving should be impeded, whereas creative thinking should be enhanced. Studies 4 and 5 tested these predictions with induced stereotypes and different creativity tasks.

### Summary

In sum, we conducted a series of five studies to test our prediction that regulatory focus influences task performance when self-stereotypes are activated. We predicted that the activation of a negative self-stereotype induces the strategic avoidance motivation of a prevention focus. This entails a vigilant, risk-averse processing style that diminishes creativity and speed while bolstering analytic thinking and accuracy (in the service of avoiding false alarms). Correspondingly, we predicted that the activation of a positive self-stereotype induces the strategic approach motivation of a promotion focus. This entails a riskier, more explorative processing style that bolsters creativity and speed (in the service of maximizing hits) while diminishing analytic thinking and accuracy.

The first study tested whether negative stereotypes lead to better recall of avoidance-related information and positive stereotypes lead to better recall of approach-related information (Higgins et al., 1994). Studies 2 and 3 tested the hypothesis that negative stereotypes lead to slower and positive stereotypes to faster performance as compared with a control group. These studies further tested the hypothesis that negative stereotypes lead to more and positive stereotypes to less accuracy of performance. Studies 4 and 5 tested the prediction that activated negative stereotypes inhibit creative processes whereas positive stereotypes enhance them. Second, they tested the prediction that activated negative stereotypes enhance analytic thinking whereas positive stereotypes inhibit it. All studies controlled for a possible mediation of results by general motivation and general performance expectancies in order to empirically distinguish between expectancy-based and regulatory focus-based mechanisms.

### Study 1

In the first study, we asked participants to proofread a text containing 20 episodes in the life of a student. Each episode exemplified self-regulation using either approach or avoidance means. Before doing the task, participants were exposed to negative or positive stereotype ascriptions. For this purpose, two groups were defined: the group of psychology students and the group of nonpsychology students. However, only nonpsychology students were tested. Participants were randomly assigned to the positive or the negative stereotype condition. This was done to manipulate the valence of the stereotype without changing the content of the stereotype or the reference group. We assume that in terms of

regulatory processes, these ascribed stereotypes have equivalent effects to using existing stereotypes in that they establish valenced reference points. It should be noted, however, that although “stereotypes are typically viewed as cognitive structures that contain our knowledge, beliefs, and expectations about a social group” (Kunda, 1999, p. 315), the ascribed stereotypes we used in our studies were not available knowledge structures for the participants but rather a newly learned piece of information. Therefore, we are using the term *ascribed stereotype* to refer to this situationally induced belief about the in-group.

After some filler tasks, a surprise recall task asked participants to recall the text. As our dependent measures, we used the number and the rater-coded detailedness of the recalled events per category. The episodes we presented described goal-directed behavior of a protagonist. More specifically, the episodes specified an end state and avoidance means to reach it or an end state and approach means to reach it. As Higgins et al. (1994) found, a prevention focus leads to better recall of episodes stating avoidance means, whereas a promotion focus leads to better recall of episodes stating approach means. On the basis of this finding, we predicted that participants exposed to the negative stereotype would recall more episodes involving avoidance behavior and recall more details about them than participants exposed to the positive stereotype. Participants with a positive stereotype, on the other hand, would recall more episodes involving approach behavior as well as more details about them.

### Method

#### Design and Participants

The study took the form of a  $2 \times 2$  factorial design. The factors were ascribed stereotype (positive vs. negative) and direction of stated means (approach vs. avoidance). Correct recall was the primary dependent measure. Thirty-two German students not majoring in psychology from the city of Bremen (16 women and 16 men) were recruited out of a subject pool for a battery containing several unrelated experiments at the International University of Bremen (IUB) and were randomly assigned to one of the two experimental groups. The battery lasted about 90 min, and participants were paid €10 plus €5 (a total of approximately \$15 US) travel money for participation. The experiment took place at IUB with experimenters from different nations. At IUB, English is the language of any social interaction, and thus some experimenters were not fluent in German. To overcome this problem, there was always a German experimenter present in case our German participants had any questions. One participant had to be excluded from the analyses because he left the experiment before finishing.

#### Procedure

After arrival, participants were asked to do unrelated tasks for about 30 min. Then, the present experiment was introduced as a cognitive performance test for students. As Aronson et al. (1999) showed, domain identification moderates stereotype effects in that only highly identified individuals show performance decrements because of negative stereotypes. Therefore, the introductory sheet stated that the measure was diagnostic of verbal intelligence.

Then, participants in the negative-stereotype condition were told that psychology students usually do very well on the task whereas students majoring in other disciplines do badly. Participants in the positive-stereotype condition were told that psychology students usually do badly on the task whereas students majoring in other disciplines usually do very well. Directly afterward they were asked to write down their major.



Then, we assessed expectations ("What do you think, how well will you perform in the task?" on a scale from 1 = *very bad* to 9 = *very good*), liking of the task ("How much do you like the task?" on a scale from 1 = *not at all* to 9 = *very much*), and general motivation ("How motivated are you to do well on the task?" on a scale from 1 = *not at all* to 9 = *very motivated*). These questions were included to test the supposed independence of the predicted effects from general motivation and performance expectation.

Then, participants were asked to proofread a text for which no time limit was given. This task was used to assess the incidental memory for approach- and avoidance-related information. All participants finished the task within 5 min. The text was entitled "Information on Amelie K." and contained 20 sentences with misprints, 10 reflecting approach behavior (e.g., "Because I wanted to buy something nice for my best friend, I went shopping to town and searched for presents"), and 10 specifying avoidance behavior (e.g., "Because I did not want to say something stupid I did not say anything in class") in an intermixed fashion. After the task, a 15-min filler task was given and then a surprise recall task followed, asking participants to recall as many episodes of the story as possible.

## Results

### Number and Detailedness of Recalled Items

Two experts (blind to condition) rated the correctness of recalled items and in how much detail they were remembered. Both judges were familiar with regulatory focus theory. The judges were asked to categorize the recalled items as correct when the item contained the gist of the episode and when the direction of the stated mean (approach vs. avoidance) was correctly recalled. In cases of disagreement, the raters were asked to discuss the cases and to agree on them until total agreement was reached. A second judgment was made concerning the detailedness of the recalled items, on a scale from 1 (*not detailed at all*) to 5 (*very detailed*). The interrater reliability was  $r = .87$  for approach and  $r = .84$  for avoidance items. The two ratings were summarized and divided by 2.

The mean number of correctly recalled items is summarized in Table 1. An analysis of variance (ANOVA) for repeated designs was conducted, which yielded the predicted interaction between direction of stated means and ascribed stereotype,  $F(1, 29) = 11.87$ ,  $p = .002$ , showing that a positive stereotype led to better memory for approach items ( $M = 1.56$ ) than for avoidance items ( $M = 1.19$ ),  $t(29) = 1.35$ ,  $p = .094$  (one-tailed), whereas a negative stereotype led to better memory for avoidance items ( $M = 1.80$ ) than for approach items ( $M = 0.80$ ),  $t(29) = 3.49$ ,  $p = .001$  (one-tailed).<sup>1</sup> Comparisons within items revealed that the difference between conditions was significant for approach items,  $t(29) = 2.32$ ,  $p = .014$  (one-tailed), as well as for avoidance items,  $t(29) = 3.09$ ,  $p = .002$  (one-tailed). No other effect reached significance (all  $F$ s < 2.5, all  $p$ s > .12).

The ratings on the detailedness of the correctly recalled information showed a similar pattern (see Table 1). A second ANOVA was conducted on these ratings, which yielded the predicted two-way interaction between direction of stated means and ascribed stereotype,  $F(1, 29) = 16.75$ ,  $p < .001$ , showing that under the positive stereotype, recall of approach items was more detailed ( $M = 3.91$ ) than recall of avoidance items ( $M = 2.77$ ),  $t(29) = 2.87$ ,  $p = .005$  (one-tailed), whereas under the negative stereotype, recall of avoidance items was more detailed ( $M = 4.04$ ) than recall of approach items ( $M = 2.83$ ),  $t(29) = 2.92$ ,  $p = .004$  (one-tailed).

Table 1

*Study 1: Mean Number and Detailedness of Correctly Recalled Items as a Function of Ascribed Stereotype and Direction of Stated Means*

| Ascribed stereotype                      | Direction of stated means |           |
|--|---------------------------|-----------|
|  | Approach                  | Avoidance |
| No. of correctly recalled items          |                           |           |
| Positive                                 |                           |           |
| <i>M</i>                                 | 1.56                      | 1.19      |
| <i>SD</i>                                | 1.21                      | 0.54      |
| Negative                                 |                           |           |
| <i>M</i>                                 | 0.80                      | 1.80      |
| <i>SD</i>                                | 0.41                      | 0.56      |
| Detailedness of correctly recalled items |                           |           |
| Positive                                 |                           |           |
| <i>M</i>                                 | 3.91                      | 2.77      |
| <i>SD</i>                                | 0.67                      | 0.93      |
| Negative                                 |                           |           |
| <i>M</i>                                 | 2.83                      | 4.04      |
| <i>SD</i>                                | 1.50                      | 0.59      |

Comparisons within item type revealed that the difference between conditions was significant both for approach items,  $t(29) = 2.35$ ,  $p = .014$  (one-tailed) as well as for avoidance items,  $t(29) = 16.41$ ,  $p < .001$ . No other effects were found (all  $F$ s < 1).

### Motivation, Expectancy, and Liking of the Task

Stereotype condition did not influence motivation, expectancy, or liking (all  $F$ s < 1).

## Discussion

The results from Study 1 show that when a positive stereotype is induced, incidental memory for information focusing on approach behavior is better than that for information on avoidance behavior. When a negative stereotype is induced, information on avoidance behavior is better remembered than information on approach behavior. These findings parallel the results by Higgins et al. (1994) and support our claim that one of the effects of activating stereotypes concerning performance is the induction of regulatory foci. The procedure used was a rather direct test of focus induction, because one of the central tenets of regulatory focus theory concerns the different strategic inclinations toward approach versus avoidance means due to promotion-respective prevention focus (Higgins, 1997). Building on this evidence, we can now explore some of the implications of the regulatory model for performance under activated stereotypes.

### Study 2

To test the implications of our model for the speed and accuracy of performance, we chose the stereotype that women have better

<sup>1</sup> The overall low recall performance might be due to the fact that the task used focused attention on the individual words while impeding effective encoding of the meaning of the entire phrases presented.

verbal abilities than men. This stereotype was activated by introducing the task as one measuring "verbal skills of men and women." The explicit reference to gender was introduced because there is reason to believe that White men are less often confronted with negative stereotypes about their performance than other groups (see Aronson et al., 1999), so a more subtle manipulation might not activate the stereotype. Additionally, participants in the stereotype activation condition were instructed to indicate their gender on the task sheet immediately before solving the verbal task. This was done to make the gender category salient, which in the context of the verbal skill task should heighten the chance of activating the corresponding self-stereotype. In the no-activation condition, the task was also introduced as measuring verbal abilities; however, no reference to gender was made.

Our central hypothesis was that activating the stereotype concerning gender differences prior to presenting a verbal ability task should have opposite effects on the performance of female and male participants. More specifically, it was predicted that women work faster and less accurately when the (positive) stereotype is activated, whereas men should work slower and more accurately when the (negative) stereotype is activated. As a dependent measure, we used a task based on the *Wortauswahl* (Word Selection) scale of the *Intelligenz-Struktur-Test 70* (Intelligence Structure Test 70; Amthauer, 1973). This task was constructed and pretested with respect to its sensitivity to differential concerns with speed and accuracy. Speed was operationalized as the number of items attempted and inaccuracy as the number of mistakes.

## Method

### Design and Participants

The study took the form of a  $2 \times 2$  factorial design. The factors were gender of the participant (male vs. female) and condition (stereotype vs. control). Speed and accuracy were the primary dependent measures. Sixty students at the University of Würzburg (31 women and 29 men) were recruited for a psychological study for which a chocolate bar was offered as compensation. Of the 29 men, 14 were in the stereotype condition and 15 in the control condition. Of the 31 women, 15 were in the stereotype condition and 16 in the control condition.

### Stimulus Material and Pretest

In the word-selection task (Amthauer, 1973), each item consists of five words, and participants have to select the one word that does not belong to the category of the majority of words. To give an example, for *banana*, *apple*, *carrot*, *cherry*, and *strawberry*, the right answer is *carrot* because it is the only vegetable in the group. In order to obtain a verbal ability task that is sensitive to variations in speed versus accuracy concerns, we constructed and pretested items with a low item difficulty when solved under an accuracy instruction but producing many mistakes when presented with time pressure under a speed instruction.

Fifty-eight students from the University of Würzburg participated in the pretest (30 in the speed and 28 in the accuracy condition). This sample was representative with respect to the sample used in the main study. The word-selection task we used consisted of 13 items. Under accuracy instruction, participants were given enough time to solve all the items, whereas under speed instruction, they were interrupted after 30 s. The accuracy instruction stated the following: "For this task, you are given unlimited time. Please think carefully and mark the best solution." Under speed instruction, participants read the following paragraph: "For this task you are given 30 seconds. Please do not think for too long and mark the best

solution. Try to solve all items in the given time!" Because participants solved fewer items in one condition than in the other and because we did not assume item difficulty to be constant across items, we calculated the ratio of wrong answers over total number of answers given per item and per instruction. The average percentage of wrong solutions per item was 19.5% under accuracy instruction and 30.6% under speed instruction. A *t* test for dependent variables on these ratios revealed that this difference was reliable,  $t(12) = -3.25$ ,  $p = .007$ . Thus, the task is sensitive to differences in speed and accuracy concerns.

## Procedure

Participants reported to the laboratory in mixed male and female groups of 3 to 5. Two female experimenters conducted the study. When participants arrived at the laboratory, they were told to read the introductory page of the questionnaire where the anonymity of the study was stressed. Participants in the stereotype condition then read that the upcoming task was part of a series of tasks measuring mathematical, artistic, and verbal abilities and that in their group we were investigating verbal abilities of men and women. In the control condition, participants read that they were asked to solve a verbal task. Then, all participants indicated motivation and expectancy as in Study 1. Afterward, participants indicated their major, age, and gender. Participants in the stereotype activation condition were further told that for data-processing reasons, they would be asked to indicate their age and gender repeatedly on each sheet. The next sheet was either titled "Verbal Ability Test" (stereotype condition) or had no title (control condition) and contained the instruction for the main task. Participants read an example of the task and were then instructed to be as accurate and as fast as possible. They were told that both speed and accuracy were equally important. Once all participants understood the instruction, they started working on the task. The time for solving the task was limited to 30 s and was stopped by trained experimenters with a stopwatch. In the end, participants were interviewed about their hypotheses concerning the purpose of the study. No one discovered its real purpose. After this, they were debriefed, given the chocolate bar, and thanked for taking part.

## Dependent Measures

For each participant, a measure of the speed of performance and an inaccuracy measure were calculated. Speed was measured as the number of items for which a (right or wrong) solution was given; inaccuracy was assessed as the total number of wrong solutions.

## Results

### Test Performance

**Speed measure.** As can be seen in Figure 1, an examination of the means indicates that women were faster in the stereotype condition ( $M = 5.67$ ,  $SD = 1.63$ ) than in the control condition ( $M = 5.00$ ,  $SD = 1.37$ ), whereas men were faster in the control condition ( $M = 5.13$ ,  $SD = 1.73$ ) than in the stereotype condition ( $M = 4.07$ ,  $SD = 1.07$ ),  $F(1, 56) = 5.13$ ,  $p = .027$ , for the Gender  $\times$  Stereotype Activation interaction. For men, this difference was significant,  $t(56) = 1.94$ ,  $p = .029$  (one-tailed), whereas for women, it was not,  $t(56) = 1.26$ ,  $p = .107$  (one-tailed). When the stereotype was activated, a clear gender difference emerged,  $t(56) = 2.91$ ,  $p = .002$ , whereas no gender difference emerged in the control condition ( $t < 1$ ). There was also a tendency for women to be faster than men,  $F(1, 56) = 3.671$ ,  $p = .060$ . For stereotype activation, no main effect emerged ( $F < 1$ ).

*Inaccuracy measure unadjusted for speed.* As Figure 2 shows, women made more mistakes in the stereotype condition ( $M = 1.47$ ,  $SD = 0.74$ ) than in the control condition ( $M = 1.19$ ,  $SD = 0.66$ ), whereas men made fewer mistakes in the stereotype condition ( $M = 0.71$ ,  $SD = 0.73$ ) than in the control condition ( $M = 1.60$ ,  $SD = 0.83$ ),  $F(1, 56) = 9.28$ ,  $p = .004$ , for the Gender  $\times$  Condition interaction. For men, this difference was significant,  $t(56) = 3.22$ ,  $p = .001$  (one-tailed), whereas for women it was not,  $t(56) = 1.05$ ,  $p = .149$  (one-tailed). When no stereotype was activated, a marginal effect for gender emerged (men making more mistakes than women),  $t(56) = 1.55$ ,  $p = .063$ . When the stereotype was activated, however, men committed significantly fewer errors than women,  $t(56) = 2.74$ ,  $p = .004$  (one-tailed). No other effects reached significance.

*Inaccuracy measure adjusted for speed.* To determine whether inaccuracy was influenced independently of speed, the recommendations by Baron and Kenny (1986) were followed. According to this approach, three relationships between the target variables must be demonstrated in a series of regression analyses to establish a basis for testing mediation. The independent variable must predict both the dependent and the mediator variable, and the mediator must predict the dependent variable. Once these conditions are established, the dependent variable is regressed onto the independent variable and mediator in a final regression analysis. Support for mediation is obtained by demonstrating that the effect of the independent variable (the Gender  $\times$  Condition interaction) on the dependent variable (inaccuracy) is significantly reduced when accounting for the effect of the hypothesized mediator (speed). When the independent variable in a mediational analysis is an interaction, the question is one of mediated moderation.

As the above analyses revealed, the Gender  $\times$  Condition interaction was significant for both the dependent variable (inaccuracy) as well as for the potential mediator (speed). Furthermore, speed was a significant predictor of inaccuracy on the verbal test,  $\beta = .27$ ,  $t(58) = 2.10$ ,  $p = .040$ . In the final analysis, the independent variables as well as speed were entered jointly to predict inaccuracy. The effect of speed was not significant in this analysis,  $\beta = .15$ ,  $t(55) = 1.16$ ,  $p = .252$ . Furthermore, the predicted interaction of gender and condition remained significant,  $\beta = .57$ ,  $t(55) = 2.59$ ,  $p = .012$ . A Sobel test of the reduction in the direct stereotype activation effect was not significant ( $Z = 1.04$ ,  $p = .30$ ). It

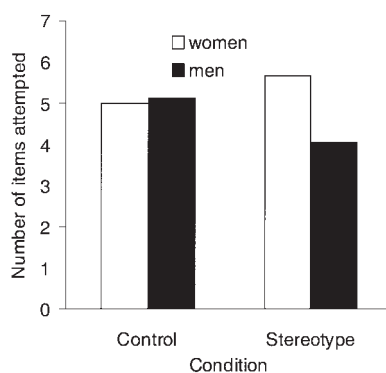


Figure 1. Mean number of items attempted as a function of gender and condition in Study 2.

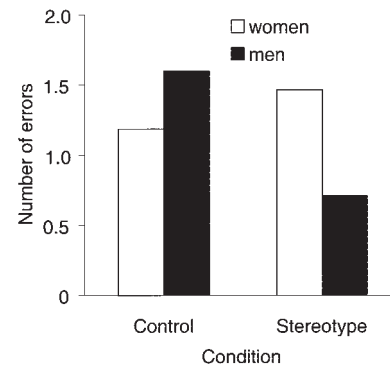


Figure 2. Mean number of errors as a function of gender and condition in Study 2.

can thus be concluded that the predicted effect on inaccuracy does not seem to be mediated by speed.<sup>2</sup>

### Motivation and Expectancy

For motivation, a marginal condition main effect emerged, with those in the stereotype condition being more motivated than those in the control condition,  $F(1, 56) = 3.92$ ,  $p = .053$ . This main effect, however, was qualified by a significant Gender  $\times$  Condition interaction,  $F(1, 56) = 4.58$ ,  $p = .037$ , indicating that men in the stereotype condition ( $M = 6.29$ ,  $SD = 1.64$ ) tended to be more motivated than those in the control condition ( $M = 4.47$ ,  $SD = 2.13$ ), whereas no such difference was found for women ( $M = 5.87$ ,  $SD = 1.55$ , and  $M = 5.94$ ,  $SD = 1.44$ , respectively). No gender main effect emerged,  $F(1, 56) = 1.42$ ,  $p = .23$ .

To determine whether motivation mediated the results, first it was determined whether motivation covaried with the dependent measures. Although it was unrelated to speed ( $\beta < .01$ ), there was a significant relationship with inaccuracy such that participants who indicated more motivation committed fewer mistakes,  $\beta = -.28$ ,  $t(58) = -2.20$ ,  $p = .032$ . Therefore, inaccuracy was regressed onto gender, condition, and their interaction as well as on motivation as the potential mediator. This analysis yielded only the predicted Gender  $\times$  Condition interaction,  $\beta = .56$ ,  $t(55) = 2.57$ ,  $p = .013$ . The effect of motivation on inaccuracy was not significant,  $\beta = -.18$ ,  $t(55) = -1.38$ ,  $p = .173$ . To conclude, general motivation did not seem to mediate the expected interaction effect.

<sup>2</sup> As the previous analyses showed, activating the stereotype influenced men's speed as well as their accuracy whereas it did not reliably influence women's speed or accuracy. Therefore, a second analysis was conducted to determine whether speed might have mediated the effect of stereotype activation on inaccuracy for men alone. However, when inaccuracy was regressed onto both stereotype activation and speed, stereotype activation remained significant in the equation,  $\beta = .43$ ,  $t(26) = 2.45$ ,  $p = .021$ , whereas speed was no longer a significant predictor of inaccuracy,  $\beta = .21$ ,  $t(26) = 1.22$ ,  $p = .235$ . This indicates that the effect of stereotype activation on men's accuracy was independent of its effect on their speed of performance as also evidenced by a nonsignificant Sobel test ( $Z = 1.03$ ,  $p = .30$ ).

A nonsignificant Sobel test corroborated this conclusion ( $Z = 1.16, p = .25$ ).<sup>3</sup>

Analyses with performance expectancy as the dependent variable revealed a tendency for men ( $M = 6.10$ ) to expect a better result than women ( $M = 5.35$ ),  $F(1, 56) = 3.24, p = .077$ . No difference in performance expectancy was found for condition ( $F < 1$ ), nor for the Gender  $\times$  Condition interaction,  $F(1, 56) = 1.31, p = .25$ . Therefore, no mediational analysis was conducted.

### Discussion

We have argued that activated stereotypes, by inducing a regulatory focus, have differential effects on different aspects of performance. In line with this reasoning, we found in Study 2 that the activation of a negative stereotype concerning ability leads to slower performance in that domain whereas the activation of a positive stereotype leads to faster performance. Furthermore, we found that the activation of a negative stereotype leads to more and the activation of a positive stereotype to less accurate performance compared with a control group. Thus, whether stereotype activation leads to performance increments or decrements depends on the match between the requirements of the task and participants' motivational orientation.

Contrast analyses revealed that the effects were only reliable for men or for the activation of the negative stereotype. This could mean that we were less successful in activating the positive stereotype than the negative one. Another possible explanation would be that the performance of women in the control condition was already almost as fast as possible, taking into account the time that it takes to read the items. Note that this asymmetry can be observed in most of the studies on stereotype threat, such as in the original studies by Steele and Aronson (1995), for example, where White men's performance did not improve when the racial stereotype was activated.

The fact that men in the stereotype condition indicated higher motivation than those in the control condition points to the possibility that the rather explicit activation of the stereotype in this study might have motivated male participants to try to disconfirm the stereotype. Reported motivation, however, was not responsible for the differential stereotype activation effects on men and women, as the mediational analysis indicates. Thus, it seems that the specific motivational orientations of prevention vigilance versus promotion eagerness and not the general motivation of the participants per se were responsible for the obtained pattern.

### Study 3

In Study 2, we found evidence that performance is faster under a positive than under a negative stereotype and that it is more accurate under a negative than under a positive stereotype. The overall pattern of results obtained in this study thus evidently confirms our hypotheses. However, it should be noted that some of the predicted effects did not emerge as clearly as desired. Several factors might have attenuated the experimental power for uncovering the predicted effect. First, stereotype-threat effects can best be diagnosed when individual differences in ability are taken into account. Therefore, in most of the studies on stereotype threat, participants' general ability level for certain tasks (e.g., such as reflected in Scholastic Aptitude Test scores) is assessed, which

then serves as a statistical control for actual performance. Because we predicted differential effects on different aspects of performance, we decided not to use such control measures for Study 2. Furthermore, the fact that the positive and the negative stereotype applied to different groups (women and men, respectively) gave us little experimental control over potential sources of variance, a problem that other studies on the effects of regulatory focus induction do not face. Additionally, by not using one of the tasks that had been extensively pretested to be responsive to regulatory focus strategic concerns, we explored the scope of tasks that are affected by the phenomenon but lost some of the power of these well-established measures. And finally, the usage of preexisting stereotypes helped establish the external validity of the observed effects but entailed the problem of asymmetric stereotypes, that is, differential stereotype accessibility for the positively and the negatively stereotyped group.

Therefore, we conducted a further study where we used a task that had been pretested to be sensitive to regulatory focus concerns (Förster et al., 2003) and experimentally manipulated the stereotype condition. By doing this, we could make sure that the positive stereotype would be equally applicable and strong as the negative stereotype. Even though some external validity might get lost when experimentally inducing stereotypes instead of using preexisting ones, such a procedure can reduce variance and can rule out other potential explanations for the effects, such as personality differences or habitual responses to activation of a particular stereotype. This should enable us to obtain clearer evidence concerning the effects of positive stereotypes on speed and accuracy. We used the same stereotype ascription procedure as in Study 1. The task consisted of connecting numbered dots as quickly and as accurately as possible.

### Method

#### Design and Participants

The experiment took the form of a one-factorial design with the ascribed stereotype factor comprising three levels (negative vs. control vs. positive). The speed and accuracy with which participants connected numbered dots in three different figures served as primary dependent measures. Fifty-six students at the University of Würzburg were recruited for a battery containing several unrelated experiments and were randomly assigned to one of the three experimental groups. None of the students were majoring in psychology. There were 20 participants in the positive stereotype condition, 19 in the negative stereotype condition, and 16 in the control condition. The battery lasted about 1 hr, and participants were paid DM 12 (at the time approximately \$6 US) for participation.

#### Stimulus Material

To measure speed and accuracy, we used three pictures from a children's drawing book. The task was to draw a picture by connecting numbered dots

<sup>3</sup> Again, because the contrast for stereotype activation was only significant for men, a second analysis was conducted to determine whether general motivation might have mediated the effect of stereotype activation on inaccuracy for men alone. However, when inaccuracy was regressed onto both stereotype activation and motivation, stereotype activation remained significant in the equation,  $\beta = .42, t(26) = 2.28, p = .031$ , whereas motivation did not,  $\beta = -.19, t(26) = -1.04, p = .31$ .



in the order of their number. When correctly completed, all pictures depicted cartoon animals whose names were printed on the bottom of the page (e.g., "Kamel" [camel]). This labeling was done to control for higher order cognitive mediation, such as ability to imagine the animal behind the dots (see Förster et al., 2003).

### Procedure

On arrival at the laboratory, participants were told that they were participating in several unrelated studies, conducted together for economic purposes. Then, they completed tasks unrelated to the present experiment for about 30 min. The present experiment was introduced as a cognitive performance test for students. To heighten motivation to do well on the test, the following paragraph<sup>4</sup> explained the alleged purpose of the test:

Contemporary research shows that individuals who do well on the task are frequently found to have excellent "abstract imagination." Abstract imagination is necessary to make the right decisions in a world that is becoming more and more interconnected. More and more complex and global problems and tasks call for new strategies of problem solving that draw essentially on abstract imagination. The present test is designed to investigate which students possess this ability to a high degree and which ones do not.

Then, the positive and negative stereotypes were ascribed the same way as in Study 1. For participants in the control condition, the sentence inducing the stereotype was left out. Afterward, participants were instructed to be both fast and accurate. They read that "accurate" meant to actually hit the dot they were connecting, and "fast" meant to connect as many dots as possible. They further read that both results would be used to calculate a compound measure that could predict success in their studies as well as other cognitive achievements. Participants were again reminded that they should be both accurate and fast and that they had 30 s for each of the three pictures. Pretesting indicated that nobody completed a picture within the time period.

Next, participants were asked to indicate their major, followed by ratings of performance expectancy, motivation, and liking for the task as in Study 1. The experimenter started and stopped the time for each picture with a stopwatch. The pictures were always presented in the same order. At the end of the experimental session, participants were debriefed, thanked, and paid.

## Results

### Test Performance

**Speed measure.** The speed measure, the mean number of dots connected in the three pictures, showed the expected main effect for ascribed stereotype,  $F(2, 52) = 11.66, p < .001$ . Participants who had been exposed to the positive stereotype ( $M = 56.42, SD = 8.74$ ) were faster than those in the neutral condition ( $M = 51.02, SD = 6.51$ ),  $t(52) = 2.16, p = .018$  (one-tailed), whereas those exposed to the negative stereotype ( $M = 44.88, SD = 6.69$ ) were slower than those in the neutral condition,  $t(52) = -2.43, p = .009$  (one-tailed).

**Inaccuracy measure unadjusted for speed.** For the inaccuracy measure, the average number of dots left out or not crossed correctly across the three pictures, the expected main effect of ascribed stereotype emerged,  $F(2, 52) = 12.51, p < .001$ . Participants in the negative stereotype condition were the least inaccurate ( $M = 3.07, SD = 2.93$ ) whereas those in the positive stereotype condition were the most inaccurate ( $M = 8.95, SD = 4.52$ ); the result for the control condition fell in between ( $M = 5.94,$

$SD = 3.24$ ). Contrast analyses for the predicted effects show that both the positive stereotype condition,  $t(52) = 2.45, p = .009$  (one-tailed), and the negative stereotype condition,  $t(52) = 2.30, p = .012$  (one-tailed), differed significantly from the neutral condition.

**Inaccuracy measure adjusted for speed.** It was then determined whether the effect of the manipulation on inaccuracy was mediated by its effect on the speed of performance. The coefficient of speed predicting inaccuracy was  $\beta = .62, t(53) = 5.82, p < .001$ . Given this significant result, a mediational analysis for the effect of stereotype ascription (positive vs. negative stereotype) on inaccuracy with speed as a mediator was conducted. This analysis showed that partial mediation had occurred. The regression coefficients were significant for stereotype condition predicting accuracy,  $\beta = .62, t(37) = 4.79, p < .001$ ; stereotype condition predicting speed,  $\beta = .60, t(37) = 4.61, p < .001$ ; and speed predicting accuracy when stereotype condition was controlled,  $\beta = .49, t(36) = 3.45, p = .001$ . This significant mediation was confirmed in a significant Sobel test ( $Z = 2.75, p = .006$ ). Furthermore, stereotype condition still significantly predicted accuracy when speed was controlled,  $\beta = .32, t(36) = 2.26, p = .030$ , indicating that the mediation was only partial. The treatment thus did have an independent effect on inaccuracy, which, however, was considerably weaker than its effect when speed was not controlled.

### Motivation, Liking of the Task, and Expectancy

There was a tendency for those in the control condition to expect a better performance ( $M = 5.47, SD = 1.30$ ) than those in the negative stereotype condition ( $M = 4.70, SD = 1.34$ ) or the positive stereotype condition ( $M = 4.30, SD = 1.53$ ),  $F(2, 52) = 3.00, p = .058$ . For liking of the task and motivation, no group effect emerged (all  $F_s < 1$ ). Therefore, no mediational analyses were conducted.

## Discussion

Study 3 provided further support for the hypothesis that positive stereotypes enhance the speed of performance whereas negative stereotypes enhance accuracy. Consistent with the self-regulation approach proposed in this article, we were able to show that speed and accuracy are influenced in different directions by positive and negative stereotypes. Given that we found evidence for regulatory focus induction with the same manipulation on a different dependent variable in Study 1, these results point to the validity of our regulatory focus account for the effects of positive as well as negative stereotypes. The fact that in Study 2 as well as in Study 3 the effects of stereotype condition on accuracy remained significant when controlling for its effect on speed justifies the separate consideration of both aspects of performance. This conclusion is also supported by four studies by Förster et al. (2003), who found strong independent effects of regulatory focus on both aspects of performance.

Study 3 extended the findings from Study 2 in two important aspects. First, we obtained conclusive evidence not only for

<sup>4</sup> This was originally in German. The German version can be obtained from Beate Seibt.

heightened vigilance under negative stereotypes (slower and more accurate performance) but also for heightened eagerness motivation under positive stereotypes (faster and less accurate performance). Second, the task used in Study 3 has already been shown to be influenced by regulatory focus (Förster et al., 2003). Finding the predicted effects in this task can thus give us more confidence that regulatory focus did indeed mediate them.

One could argue that the results by Leyens, Desert, Croizet, and Darcis (2000), who found that men exposed to the stereotype of less emotional sensitivity than women produced more false alarms when judging the emotionality of verbal material, contradict our reasoning. It seems to us, though, that the specific content of the stereotype of men being less emotional than women implies specific strategies to avoid confirming the stereotype. The stereotype of not being emotional can be counteracted by trying to detect as much emotional information as possible, which in turn can be accomplished by lowering the response criterion for saying that a word does indeed have emotional content. This specific strategy would then override the prevention strategy of avoiding false positives. Most of the stereotypes considered in the stereotype-threat literature, however, do not contain specific strategies that might help to disconfirm them ("bad in mathematics," "low intellectual abilities," "bad driver," etc.). Furthermore, in most performance tasks used, it is not clear to the participant which strategies would help or hinder their solution. In these kinds of tasks, it should be the regulatory focus strategic concerns that determine the probability of the different kinds of errors (see Förster et al., 1998).

Furthermore, assuming that careful and accurate behavior facilitates golf playing, one might expect individuals in a prevention focus to be better at golf. The results by Stone et al. (1999), however, showed that Black as well as White males needed more strokes to finish the course when under stereotype threat. We argue that this is not inconsistent with a regulatory focus model. Beilock and Carr (2001) found that skilled performance suffers when the reward of another person is contingent on the participant's success (for similar results, see also Butler & Baumeister, 1998). They argued that such a condition produces extra performance pressure, which in turn leads to explicit monitoring and thereby hurts performance when proceduralized skills are involved. Thus, it is conceivable that prevention focus vigilance interferes with automatized complex psychomotoric skills. One possible mechanism is that automatic processes like eye-hand coordination are disrupted by conscious control, which can result from a state of high vigilance. Further, the kind of accuracy typical for prevention focus (avoiding mistakes) might not be the kind of accuracy necessary for a good golf performance (approaching the hole). The latter might, on a conceptual level, indeed be more compatible with promotion focus eagerness for hits than with prevention focus vigilance. Research on regulatory focus strategies has begun to systematically investigate the interaction of task characteristics and regulatory foci for a variety of tasks (i.e., Crowe & Higgins, 1997; Förster et al., 2003; Friedman & Förster, 2001; Liberman, Molden, Idson, & Higgins, 2001). However, the results do not yet allow prediction of the effects of regulatory foci for all possible tasks. Therefore, this article relies on types of tasks that already have been shown to be influenced by regulatory focus concerns, because the main aim is to show that stereotypes can influence performance by activating regulatory foci.

Regulatory focus has been shown to not only affect the speed and accuracy of performance but also creativity and analytic thinking. As Friedman and Förster (2001) showed, a promotion focus enhances different aspects of creative thinking whereas a prevention focus impedes them. In the context of the present regulatory model of stereotype-threat effects, it is therefore predicted that the activation of positive stereotypes enhances creativity and the activation of negative stereotypes impedes it. Study 4, then, explores whether the activation of a negative stereotype leads to performance decrements relative to the activation of a positive stereotype in a task requiring creative thinking. In Study 5, the effects of ascribed stereotypes on creative thinking are assessed using typicality judgments as dependent measure (see Friedman & Förster, 2000, Experiment 6).

For analytic thinking, on the other hand, a different pattern of effects is expected. A prevention focus enhances performance in analytic tasks whereas a promotion focus impedes it (Friedman & Förster, 2000). Consequently, we expected participants under negative stereotypes to show a better performance in an analytic task than participants under positive stereotypes. Studies 4 and 5 therefore also contain a task assessing analytic thinking to test this prediction. If the studies confirm these hypotheses, this will be first evidence for a reversal of stereotype-threat effects in specific kinds of tasks.

## Study 4

In the following study, we gave participants under positive or negative stereotypes a task capturing aspects of creativity and one measuring analytic thinking. We assumed that positive stereotypes would enhance creativity and undermine analytic thinking and that negative stereotypes would enhance analytic thinking and diminish creative processes. To assess creativity, we used a task measuring creative generation (Friedman & Förster, 2001), namely, the *brick task*, where participants have to generate as many creative uses for a brick as they can think of. For the analytic task, we used items from the analytic section of the GRE. Whereas the brick task profits from breaking mental set, reencoding of the stimulus, and unconscious broader memory search (creative heuristics according to Amabile, 1996; Schooler & Melcher, 1995), the analytic tasks rely on adherence to well-known algorithms as opposed to heuristics (Amabile, 1996; Friedman & Förster, 2000). The former has been found to profit from a promotion focus (Friedman & Förster, 2002) in which people broaden their minds to come up with novel ideas and solutions. The latter, however, profited from a prevention focus (Friedman & Förster, 2000), which enables people to carefully process the information given and to choose the appropriate solution path without getting distracted by more remote information. Note, however, that not every task called "analytical" requires purely analytical strategies. We therefore chose a process-pure task that has been successfully used in other research on motivation-induced cognitive tuning processes (Friedman & Förster, 2000).

## Method

### Design and Participants

The experiment took the form of a one-factorial design with the ascribed stereotype factor comprising two levels (negative vs. positive). The depen-

dent measures were the brick task (Friedman & Förster, 2001) and four translated GRE tasks (Friedman & Förster, 2000). Twenty-eight German students from the city of Bremen were recruited for a battery containing several unrelated experiments at IUB and were randomly assigned to one of the two experimental groups. The battery lasted about 1 hr, and participants were paid €7 (at the time equivalent to \$7 US) for participation. The experiment took place at IUB with experimenters from different nations. As in Study 1, there was always a German experimenter present in case our German participants had any questions. Three participants did not follow instructions and had to be omitted from the analyses below.

### Stimulus Material

*The brick task.* Participants were instructed to generate as many creative uses for a brick as they could think of. They were asked to refrain from listing typical uses or from listing uses that are virtually impossible. The dependent variable of interest was the creativity of the uses for a brick generated by participants as rated by independent experts.

*The analytic task.* Participants also worked on four logic problems from the analytical reasoning section of the GRE, translated into German. These problems involve evaluating the truth value of a number of propositions given an initial set of basic facts. This initial information was the same for all four problems. Within each problem, the task consisted of choosing which of four given propositions was true. The problems had been carefully chosen with the premise that they demand little in the way of mental restructuring and do not require cognitive search for information beyond that provided but instead require deductive reasoning and careful application of the laws of logic.

### Procedure

Participants were told that they were participating in several unrelated studies, conducted together for economic purposes. They then completed tasks unrelated to the present experiment for about 60 min. Then, they were asked to participate in a cognitive performance test for students. To make the task more relevant, participants were told that the task is highly correlated with intelligence scores. The same instructions as in Study 1 were used to induce positive and negative stereotypes. Then, participants answered the questions on performance expectancy, liking of the task, and motivation. Furthermore, we also assessed their mood ("How do you feel right now?" on a scale from 1 = *very bad* to 9 = *very good*), because mood is known to affect creativity as well as analytic thinking. Then, the analytic task was introduced as a test of verbal abilities. Participants were given 5 min to solve this task. On completion, participants were asked to solve the creativity task. They were interrupted after 1 min and told to stop generating uses. At the end of the experiment, participants were debriefed, thanked, and paid.

## Results

### Data Coding

In former research using the brick task (Friedman & Förster, 2001), 12 independent scorers (members of the University of Würzburg psychology department) were asked to rate the creativity of the 117 different uses participants generated on a 9-point scale ("How creative is this response?"), anchored at 1 (*very uncreative*) and 9 (*very creative*), with an explicit midpoint of 5 (*neither creative nor uncreative*). The means of these ratings were used to score the creativity of each response given by our participants. An example of a creative response was "to use it as a makeup for carnival" or "to show that I am just another brick in the wall"; an example of an uncreative solution was "to throw it" or

"to build a house." For this new sample, 25 solutions were reported that were not part of the original expert ratings. Therefore, 5 new independent experts rated these new solutions along the same rating scale. Before they rated the new solutions, they had been asked to read the old solutions and their ratings carefully to become familiar with the scale. Then, we calculated the mean creativity score for each participant (summed ratings for each response rendered divided by the total number of responses).

### Test Performance

As expected, the positive stereotype group had a higher creativity score ( $M = 4.44$ ,  $SD = 0.49$ ) than the negative stereotype group ( $M = 3.87$ ,  $SD = 0.48$ ),  $t(23) = 2.93$ ,  $p = .008$ , reflecting the typical finding that stereotype threat undermines task performance. However, for the analytic task, the reverse pattern was found. Here, a negative stereotype led to better performance ( $M = 1.46$ ,  $SD = 0.78$ ) than a positive one ( $M = 0.83$ ,  $SD = 0.72$ ),  $t(23) = -2.10$ ,  $p = .047$ .

### Motivation, Expectancy, Mood, and Liking of the Task

There was no effect of stereotype condition on either mood, motivation, or performance expectancy (all  $t$ s  $< 1.5$ , all  $p$ s  $> .17$ ). We found, however, a tendency for those in the positive stereotype group ( $M = 6.33$ ,  $SD = 1.50$ ) to like the task better than those in the negative stereotype group ( $M = 5.15$ ,  $SD = 1.68$ ),  $t(23) = 1.85$ ,  $p = .077$ . When regressing the performance measures on liking, we found a marginal effect on the creativity score,  $\beta = .39$ ,  $t(23) = 2.01$ ,  $p = .056$ , and none on analytic performance ( $\beta = -.03$ ,  $t < 1$ ). When stereotype condition was entered along with liking in the regression analysis, however, liking no longer predicted creativity,  $\beta = .23$ ,  $t(22) = 1.21$ ,  $p > .23$ , whereas stereotype condition did,  $\beta = .44$ ,  $t(22) = 2.33$ ,  $p = .030$ . Thus, liking does not seem to mediate the effects of stereotype condition on performance.

### Discussion

As predicted, the effect of ascribed stereotype on task performance differed as a function of the requirements of the tasks. For tasks requiring creativity, the usual stereotype-threat effect was observed. For tasks requiring analytical thinking and great attention to detail, on the other hand, the reverse pattern emerged. These results, then, provide evidence for improvement as well as undermining of task performance due to negative stereotypes. Thus, regulatory focus theory provides a framework to generate specific predictions of how task characteristics and activated stereotypes jointly determine task performance. These findings can help to judge when performance decrements due to activated negative stereotypes should be expected and when they are less likely to occur.

Even though the brick task has high face validity, it is influenced by many different processes and bound to rather subjective ratings. Therefore, we replicated Study 4 with a different creativity task tapping regulatory focus strategies more directly. One reason why creativity profits from a promotion focus eagerness is broadening conceptual attention or, in other words, the more extensive activation of unrelated networks (cf. Friedman & Förster, 2000, 2002).



A solution like using the ground brick as makeup for a carnival requires the activation of at least two quite unrelated mental networks, the network of a brick and the network of a carnival. One task that captures this aspect of creativity and that profited from approach cues (Friedman & Förster, 2000) was used in Study 5. In addition, the motivational experiences of vigilance and eagerness are assessed, because these specific experiences should be affected by a regulatory focus manipulation.

### Study 5

Study 5 was conducted to obtain further evidence that a positive stereotype leads to more creativity and a negative stereotype to less whereas a negative stereotype leads to better analytic performance and a positive stereotype to worse. In a task inspired by Isen and Daubman (1984), participants rated the goodness of fit of weak exemplars (e.g., "camel") of a given category (e.g., "vehicle"). This task captures two main classic components of creative thinking: unconscious mental search and breaking context-induced mental set (Schooler & Melcher, 1995). As proposed by Isen (1987), more inclusive categorization (i.e., higher goodness-of-fit ratings) may reflect more extensive underlying search for shared features between weak and prototypical exemplars. The tendency to include atypical exemplars as members of a given category may also be seen as set breaking, inasmuch as it may entail overcoming preexisting assumptions regarding the criteria for category membership; this capacity is another important ingredient of creative processes. For instance "foot" may be included as a member of the category "vehicles" if the common assumption that body parts do not represent vehicles is called into question or momentarily held aside. Thus, it was expected that atypical exemplars would be seen as better category members when participants were told that their in-group could do the task better than the out-group than when they were told that the out-group could do the task better than the in-group. The judgments of the control group were expected to fall in between.

Furthermore, we wanted to assess more directly the mediating processes. We reasoned that promotion eagerness should be enhanced when participants think that their in-group is better and that prevention vigilance is enhanced under stereotype threat. These motivational experiences could then mediate the different tasks: Cognitive flexibility as measured by the categorization task should be mediated by eagerness, whereas analytical reasoning should be mediated by vigilance. Mediation by current mood and general motivation, on the other hand, was not expected. We therefore assessed these variables and tested for mediation.

### Method

#### Design and Participants

The experiment took the form of a one-factorial design with the ascribed stereotype factor comprising three levels (negative stereotype vs. control vs. positive stereotype). The dependent measures were a German adaptation of the categorization task (Friedman & Förster, 2000; Isen & Daubman, 1984) and the analytical task from Study 4. Forty-two German students from the city of Bremen were recruited for a battery containing several unrelated experiments and were randomly assigned to one of the three experimental groups. The battery lasted about 1 hr and participants were paid €7 (at the time equivalent to \$7 US) for participation. One

participant was excluded from the experiment because he threatened the female experimenter.

#### Stimulus Material

A paper-and-pencil version of the categorization task (Isen & Daubman, 1984) was used to measure participants' creativity. Participants were provided with four category names ("furniture," "vehicle," "vegetable," and "clothing"), each followed by nine items. Three of these items were good exemplars, three were moderately good exemplars, and three were poor exemplars for the category (see Friedman & Förster, 2000). Participants were asked to rate the typicality of each exemplar for the relevant category on a 10-point scale anchored at 0 (*very poor example*) and 9 (*very good example*). Category order was randomly assigned.

#### Procedure

On arrival at the laboratory, participants were told that they were participating in several unrelated studies, conducted together for economical purposes. They then completed tasks unrelated to the present experiment for about 20 min. Then, they were asked to participate in a cognitive performance test for students. The same instructions as in Study 1 were used to convey information about the performance of the in-group. To make the tasks self-relevant, participants were told that the tasks correlate highly with intelligence. Both tasks were introduced as tests of verbal abilities that can predict people's performance in various other intellectual tasks, such as abstract thinking. Then, participants were instructed to indicate their major. Afterward, performance expectancies and liking of the task were assessed as in Study 1. In addition, we checked for the regulatory focus-specific motivational experiences of vigilance ("How careful are you now?" on a scale from 1 = *not at all* to 9 = *very careful*) and eagerness ("How eager are you right now?" on a scale from 1 = *not at all* to 9 = *very eager*). Moreover, general motivation ("How motivated are you in general to do the task?" on a scale from 1 = *not at all* to 9 = *very much*) and current mood (as in Study 4) were measured. Participants worked on the task for a self-paced period of time. After they completed the task, the analytic task was introduced as another test of verbal abilities. Here, participants were instructed to complete the task within 5 min and were stopped after that period of time by the experimenter. At the end of the experiment, participants were debriefed, thanked, and paid.

### Results

#### Creativity and Analytic Performance

*The categorization task.* We hypothesized that the negative stereotype diminishes the inclusiveness of categorization whereas the positive stereotype augments it. To assess these predictions, we separately calculated the average typicality ratings for both the 12 atypical (i.e., poor) exemplars and the 24 typical (i.e., good and moderately good) exemplars, indexed by experimental condition and typicality (see Table 2).

Table 2 reveals that as expected, participants in the positive stereotype condition rated the poor exemplars as better members of their respective categories than did participants in the control condition or participants in the negative stereotype condition. To assess the reliability of these differences, the typicality ratings for the atypical exemplars were subjected to an ANOVA with the factor stereotype comprising three levels. This analysis revealed that the differences among experimental conditions were reliable,  $F(2, 38) = 9.86, p < .001$ . Planned contrasts between the cell means supplemented this ANOVA. Here, the difference between average ratings of the positive stereotype and the control group



Table 2

*Study 5: Mean Typicality Ratings for Poor and Good Exemplars in the Categorization Task and Mean Number of Solutions in the Analytic Task by Condition*

| Measure            | Ascribed stereotype |          |      |
|--------------------|---------------------|----------|------|
|                    | Positive            | Negative | None |
| Atypical exemplars |                     |          |      |
| <i>M</i>           | 2.99                | 1.46     | 2.13 |
| <i>SD</i>          | 1.03                | 0.64     | 0.97 |
| Typical exemplars  |                     |          |      |
| <i>M</i>           | 6.93                | 7.29     | 6.94 |
| <i>SD</i>          | 0.64                | 0.96     | 0.94 |
| Analytic task      |                     |          |      |
| <i>M</i>           | 0.79                | 2.23     | 1.50 |
| <i>SD</i>          | 0.70                | 1.36     | 0.94 |

was reliable,  $t(38) = 2.53$ ,  $p = .008$  (one-tailed), as was the difference between positive stereotype and negative stereotype,  $t(38) = 4.42$ ,  $p < .001$  (one-tailed) and the difference between the negative stereotype and the control condition,  $t(38) = 1.94$ ,  $p = .030$  (one-tailed). As an inspection of the means in Table 2 suggests, there were no reliable differences found for the good exemplars ( $F < 1$ ). This pattern reveals that the group differences found for atypical members did not result from a simple response bias.

**The analytic task.** We calculated the number of GRE problems solved out of the four presented. The mean number of problems solved, indexed by experimental condition, are summarized in Table 2. As predicted, participants with a negative stereotype solved more GRE tasks correctly than did participants in the control group or participants with a positive stereotype. The ANOVA testing the reliability between the differences was significant,  $F(2, 38) = 6.66$ ,  $p < .005$ . Planned contrasts revealed a significant difference between the positive stereotype and the control condition,  $t(38) = 1.83$ ,  $p = .037$  (one-tailed); a significant difference between the negative stereotype and the control condition,  $t(38) = 1.85$ ,  $p = .036$  (one-tailed); and a significant difference between the positive and the negative stereotype condition,  $t(38) = 3.65$ ,  $p < .001$  (one-tailed).

#### Motivation, Expectancy, Mood, and Liking of the Task

ANOVAs for mood, vigilance, eagerness, general motivation, performance expectancy, and liking of the task by stereotype condition were conducted. The only reliable differences were found for vigilance and eagerness. Here, as predicted, participants with the negative stereotype reported higher vigilance ( $M = 7.77$ ,  $SD = 1.24$ ) than participants in the control condition ( $M = 6.14$ ,  $SD = 1.83$ ), followed by participants with the positive stereotype ( $M = 4.71$ ,  $SD = 2.16$ ). The differences were reliable,  $F(2, 38) = 9.73$ ,  $p < .001$ . Planned contrasts revealed that the difference in vigilance between positive stereotype and control group was reliable,  $t(38) = 2.10$ ,  $p = .021$ , as well as between negative stereotype and control group,  $t(38) = 2.35$ ,  $p = .012$  and between positive and negative stereotype groups,  $t(38) = 4.41$ ,  $p < .001$ . This pattern is consistent with our hypothesis that negative stereotypes induce a prevention focus of vigilance. Furthermore, vigi-

lance motivation decreased when the positive stereotype was induced.

In addition, as predicted, for eagerness the reverse pattern emerged. Here, participants with a positive stereotype reported higher eagerness ( $M = 7.93$ ,  $SD = 1.38$ ) than participants in the control group ( $M = 6.21$ ,  $SD = 1.67$ ), followed by participants with negative stereotypes ( $M = 4.69$ ,  $SD = 1.44$ ). These differences were reliable,  $F(2, 38) = 15.64$ ,  $p < .001$ , and planned contrasts revealed that the positive stereotype condition differed from the control condition,  $t(38) = 3.01$ ,  $p = .002$  (one-tailed); that the negative stereotype condition differed from the control condition,  $t(38) = 2.63$ ,  $p = .006$  (one-tailed); and that the negative stereotype condition differed from the positive stereotype condition,  $t(38) = 5.58$ ,  $p < .001$ . There were no reliable differences for mood, general motivation, performance expectancy, and liking of the task (all  $F$ s  $< 2.1$ , all  $p$ s  $> .12$ ).

Next, we tested whether eagerness or vigilance mediated the effect of experimental condition (positive vs. negative stereotype) on typicality ratings for atypical exemplars (see Figure 3). These analyses revealed that stereotype condition predicted typicality ratings,  $\beta = .68$ ,  $t(25) = 4.62$ ,  $p < .001$ , as well as eagerness ratings,  $\beta = .77$ ,  $t(25) = 5.96$ ,  $p < .001$ . When entered jointly, eagerness had a significant effect on the typicality ratings,  $\beta = .56$ ,  $t(24) = 2.73$ ,  $p = .011$ , and the effect of stereotype condition was reduced to nonsignificance,  $\beta = .25$ ,  $t(24) = 1.22$ ,  $p = .233$ . The corresponding Sobel test corroborated this conclusion ( $Z = 2.51$ ,  $p = .012$ ). Thus, eagerness motivation mediated the effect of ascribed stereotype on inclusiveness of categorization.

The effect of stereotype condition on vigilance motivation was also significant,  $\beta = -.66$ ,  $t(25) = -4.46$ ,  $p < .001$ . When used along with stereotype condition to predict typicality ratings, however, vigilance motivation did not have a unique effect,  $\beta = .20$ ,  $t(24) = 1.04$ ,  $p > .30$ . Stereotype condition did not predict general motivation, expectancy, or liking (all  $\beta$ s  $> .24$ , all  $t$ s  $< 1.3$ , all  $p$ s  $> .22$ ).

We then repeated the mediational analyses for the score in the analytical task as dependent variable. As already reported, stereotype condition predicted the analytical score,  $\beta = -.57$ ,  $t(25) = -3.50$ ,  $p = .002$ . In a next step, the analytic score was regressed onto stereotype condition and eagerness. Eagerness did not have a unique effect over and above ascribed stereotype ( $\beta = -.08$ ,  $t < 1$ ). Then, the analytic score was regressed onto stereotype condition and vigilance. Here, vigilance had a marginal unique effect on analytic performance,  $\beta = .41$ ,  $t(24) = 1.98$ ,  $p = .060$ . In this

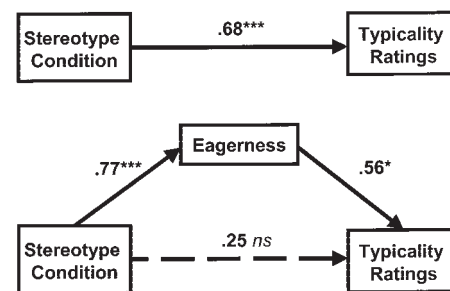


Figure 3. Mediation analysis for eagerness mediating the stereotype condition effect on typicality ratings in Study 5. \* $p < .05$ . \*\*\* $p < .001$ .

analysis, the effect of ascribed stereotype was reduced considerably,  $\beta = -.30$ ,  $t(24) = 1.45$ ,  $p = .161$  (see Figure 4). However, because the expected mediator did not have a significant unique effect, the data do not allow a clear conclusion as to whether mediation did indeed occur or not. Again, the Sobel test corroborated the results of the regression analyses ( $Z = 1.81$ ,  $p = .071$ ). It can thus be concluded that there is some evidence of vigilance motivation mediating the effects of ascribed stereotype on performance in analytical tasks.

### Discussion

Results from Study 5 replicated the findings from Study 4 with a different creativity measure and the same analytic task. It is important to note that the inclusion of a control group allowed comparison of the effects of positive and negative stereotypes. The results provide evidence for more creative thinking (i.e., more inclusive categorization as indicated by higher typicality ratings for atypical exemplars) under positive stereotypes as well as less creative thinking (i.e., less inclusive categorization) under negative stereotypes as compared with a control group without stereotype activation. Furthermore, we found better analytic performance under negative stereotypes as well as worse analytic performance under positive stereotypes compared with the control group.

The results for the categorization task replicate results by Friedman and Förster (2000, Experiment 6) who found comparable evidence in a study where motor actions induced regulatory foci. The present study found evidence that the effect of activated stereotypes was mediated by the self-judged eagerness of participants. These findings imply that positive self-stereotypes lead to the experience of eagerness and to an approach orientation. Negative self-stereotypes lead to an inhibition or lack of approach means and eagerness. The approach orientation, in turn, is characterized by a riskier, more explorative processing style as evidenced by more inclusive categorizations and the generation of more creative uses for a brick (Study 4). We thus assume that the motivational orientation of approach drove the effect of stereotype activation on the inclusiveness of categorization. The experience of eagerness, then, is an expression of this approach orientation rather than the cause of the effect.

The pattern of results for the analytic task points to a processing advantage for the negative stereotype group and a disadvantage for the positive stereotype group. The finding that the experience of vigilance motivation might have mediated the results implies that

it is mainly the absence or presence of avoidance motivation that determines the amount of analytic processing. As was the case for eagerness and approach motivation, the experience of vigilance motivation can thus be seen as an indicator of the presence or absence of avoidance motivation and the risk-averse, systematic, and vigilant processing style required for solving these kinds of analytic tasks. Taken together, the results of Study 5 point once more to the importance of taking into account the specific task requirements when making predictions about the effects of stereotype activation on performance.

### General Discussion

Five studies provided evidence that active self-stereotypes can change the strategic means of task performance by inducing regulatory foci. In general, we found that negative self-stereotypes fostered a risk-averse, vigilant processing style as indicated by higher performance accuracy, diminished creativity, and enhanced analytic thinking whereas positive self-stereotypes fostered a risky, explorative processing style with enhanced speed and creativity and diminished analytic thinking. In the first study, active positive stereotypes led to better recall for episodes describing approach strategies of goal pursuit whereas active negative stereotypes led to better recall for episodes describing avoidance strategies (see Higgins et al., 1994). In Study 2, men performed more slowly and also more accurately when the stereotype about lower verbal abilities of men was activated than when it was not. For women, stereotype activation tended to have the opposite effect. In the third study, individuals who were told that their group underperformed in a certain domain performed slower and more accurately than individuals who were not told this. Conversely, individuals who were told that their group performed better in this domain showed faster and less accurate performance than those who were not told this. Study 4 showed that the activation of a negative stereotype led to poorer performance in a creative generation task than the activation of a positive stereotype. At the same time, performance in an analytic task was improved under negative stereotypes. Finally, Study 5 provided evidence that the activation of a positive stereotype leads to more inclusive categorization as measured by a categorization task (Isen & Daubman, 1984) whereas the activation of a negative stereotype leads to less inclusive categorization. This effect was mediated by eagerness motivation. Furthermore, Study 5 replicated the effect of stereotype activation on analytic performance from Study 4 with a control group and found that both experimental conditions contributed to the effect. Additionally, some evidence for a mediation of this effect by vigilance motivation was obtained. These findings support our hypothesis that activating a negative stereotype induces a prevention-focus state of vigilance and that activating a positive stereotype induces a promotion focus state of eagerness.

### The Mediation of Stereotype-Threat Effects

The detrimental effects of negative self-stereotypes on performance might be regarded as the most disturbing and saddest phenomenon discovered by social psychologists over the last 2 decades of research on stereotypes. It looks like a tragic vicious circle: The victims of discrimination actually underperform as expected by prejudiced beliefs and thereby confirm and perpetuate

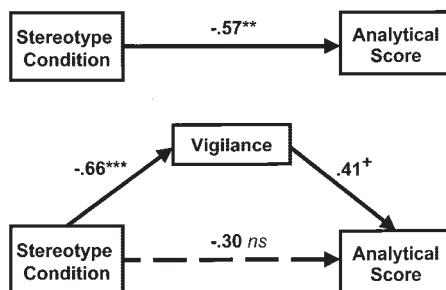


Figure 4. Mediation analysis for vigilance mediating the stereotype condition effect on analytical performance in Study 5.  $^+p < .10$ .  $^{**}p < .01$ .  $^{***}p < .001$ .

the stereotype. In our studies and in line with former assumptions (Steele, 1997), we could show that negative stereotypes created within the experimental situation had the same effects as those that are culturally inherited: In Studies 3 and 4 we could demonstrate inhibition of speed or creative generation just by creating a group of “nonpsychologists” who we described to be worse in a certain task than the group of psychologists. Our studies tested some of the limits of stereotype threat, and the bad message appears to be that the phenomenon is even broader than formerly expected.

This makes it necessary to understand the psychological processes elicited by self-stereotype activation and the boundary conditions under which their activation does not impede performance or can even improve it. Yet, because stereotypes about one's own group can have various motivational, emotional, and cognitive implications depending on the specific content of the stereotype and the context in which it is activated, it also appears to us that there might be a multitude of diverse mechanisms operating in these situations, all contributing to the effects on performance. These mechanisms can work in parallel, in sequence, or can even be conditional on aspects of the situation. Thus, we do not think that we captured all aspects of stereotype threat in the studies above, by far; rather, our intention was to identify one class of motivational mechanisms that might have been overlooked: the motivational strategies determined by individuals' regulatory focus. Regulatory focus theory (Higgins, 1997) can explain why some aspects of performance can be enhanced or impeded when self-stereotypes are activated. To make this very clear, we do not think that these mechanisms are the only ones necessary to produce the effects. We do argue, however, that in order to fully understand stereotype-threat effects in everyday situations, it is necessary to take into account the implications that the stereotypes have for the regulatory focus of a person and thereby for her or his strategic inclinations.

To give just one example, in our studies, we could not find a mediation by expectancies; however, we still believe that lowered expectancies can undermine performances and do not doubt that stereotype threat can lower expectancies (see Stangor et al., 1998, and our results in Study 3). However, it seems that negative stereotypes can also, quite automatically and prior to any experience with the task, elicit certain strategic orientations by inducing a regulatory focus. These strategic orientations then directly influence task performance. Recent research also shows that positive and negative feedback, presumably more directly eliciting different expectancies, and regulatory foci have an interactive effect on performance (Förster, Grant, Idson, & Higgins, 2001). That is, vigilance motivation is higher when in a prevention focus and receiving negative feedback, and eagerness motivation is higher when in a promotion focus and receiving positive feedback. Future research must show whether the same mechanisms hold when regulatory foci are induced by negative or positive self-stereotypes. Here, it would be especially intriguing to see whether positive feedback under activation of positive stereotypes even enhances speed and creativity and reduces accuracy and analytic performance.

It is important to note that the mechanism proposed here does not imply that a valenced stereotype always and automatically leads to a regulatory focus. Rather, the consequences of stereotype activation should depend on the self-relevance of the performance domain. If a person's self-worth is contingent on her or his

performance in a particular domain, not failing in a test will become more like a necessity (Crocker, Karpinski, Quinn & Chase, 2003). Hence, introducing a negative reference point through a negative stereotype or through negative feedback should induce a strong prevention focus. Conversely, if a person does not care about her or his performance in a domain, only a weak motivation results. If this person in addition has a self-concept of low ability in the domain, the activation of a group stereotype might even alleviate some of the performance pressure, leading to less prevention focus. In other words, when domain identification is low, confirming the stereotype is not necessarily represented as a loss (see also Steele, 1997). As a consequence, it should not induce a prevention focus. Taken together, then, the situations that elicit strong stereotype-threat effects (high domain relevance plus activation of a negative stereotype) should be the ones that elicit a strong prevention focus. In our studies we framed the tasks as highly relevant. Consequently, we assume that in the present studies, negative stereotypes could be expected to induce a prevention focus. Further research must show whether relevance moderates the observed effects.

### *Positive Stereotypes and Performance*

The results obtained in these studies do not only address the question of when to expect performance improvement or impairment due to activated negative stereotypes; they also show some conditions under which activated positive stereotypes influence performance. As such, these results are among the first to look at effects of positive stereotypes. Of the other studies, some have found positive effects of activating positive stereotypes, some negative effects, and some null effects (e.g., Spencer et al., 1999; Steele & Aronson, 1995). Positive effects of positive stereotypes were found, for example, in a study by Levy (1996), where the subliminal activation of the stereotype of wisdom led to better memory performance in elderly participants compared with a control group. Recently, Ambady, Shih, Kim, and Pittinsky (2001) found that Asian girls as young as 5 years performed better than a control group on a math test when their Asian identity was made salient (and worse when their female identity was made salient). Conversely, performance of Asian boys of the same age was boosted by activation of both their Asian and their male identity. These findings show that positive stereotypes can boost performance (for similar findings, see Kray, Thompson, & Galinsky, 2001, Experiment 2; Shih et al., 1999).

However, fulfilling a positive stereotype can sometimes turn into a necessity. In a study by Baumeister, Hamilton, and Tice (1985), privately expecting success improved performance, but performance dropped when participants were told that their audience expected them to succeed. Apparently, the performer felt the necessity of living up to the expectations of the audience. In this case, a stereotype is not only a reference point but also a standard of comparison. Accordingly, the situation of confirming the (positive) stereotype would be a nonloss and the situation of not living up to it a loss. This should be more likely to happen when one is aware that one's performance is judged against the high standard of the positive stereotype (as when the test administrator wants to find out whether gay men really are so much more creative than heterosexuals). In line with this reasoning, positive stereotypes have also been shown to impair performance.

Brown and Josephs (1999) examined whether only negative stereotypes can undermine task performance. In their studies, they framed a math test as either diagnostic of weak but not exceptional abilities or as diagnostic of exceptional but not weak abilities. As expected, women underperformed when the test was framed as diagnostic of weak abilities. It is interesting, however, that male performance dropped when the test was presented as diagnostic of exceptional abilities. It seems that male participants in such a test are afraid of not conforming to the positive stereotype that exists about their group. Accordingly, when they could attribute failure to a computer crash impeding their alleged practice opportunity, performance increased (Brown & Josephs, 1999, Experiment 3). We argue that in these studies, the instruction plus the stereotype both probably converged to the impression that an exceptional performance was expected of men. Thus, meeting this expectancy or norm would mean not losing, whereas not meeting it would mean losing. Such a situation should induce a prevention focus. The stereotype induction methods used in the present experiments, although explicit, did not communicate a norm to fulfill. Unlike in the studies by Brown and Josephs, these methods did not convey the impression that a certain very high cutoff point had to be reached in order to pass the test. Therefore, we expected the positive stereotype to induce a promotion focus. In sum, the effects of activated positive self-stereotypes seem to depend on subtle variations in the situation or the item content as well as on individual differences. The reviewed evidence is in line with the assumption that positive stereotypes sometimes lead to setting maximal goals and perceiving the situation as a gain-nongain situation that induces a promotion focus, and sometimes act as obligations, leading to minimal goals and a perception of the situation as a nonloss-loss situation that induces a prevention focus. However, more research is needed to assess how aspects of the situation and requirements of the task determine the effect of positive stereotypes on performance.

### *Measuring Regulatory Focus*

The present studies used several methods to measure the effects of regulatory foci. According to Higgins (1997), the effects of a prevention focus are (a) sensitivity to absence or presence of negative outcomes, (b) avoidance as strategic means, (c) a concern with ensuring correct rejections and ensuring against errors of commission, and (d) emotions of quiescence versus agitation (anxiety). Correspondingly, a promotion focus leads to (a) sensitivity to absence or presence of positive outcomes, (b) approach as strategic means, (c) a concern with ensuring hits and ensuring against errors of omission, and (d) emotions of cheerfulness versus dejection. Although we measured the first three outcomes, we did not measure emotions. Two problems make the assessment and interpretation of emotion reports difficult. First, emotion reports are highly subjective self-report data. Therefore, they are subject to self-presentational concerns. Second, it is hard to know what the reported emotions are about unless performance feedback is given.

In a recent study, Keller and Dauenheimer (2003) manipulated stereotypes in a performance situation and measured dejection, agitation, cheerfulness, and quiescence after test completion. The authors proposed that "stereotype threat leads to reduced performance expectations . . . and . . . to an experience of dejection in the participants when a promotion focus is activated in a test situation"

(p. 373). Although expectations were not assessed in their study, they found a significant influence of the stereotype-threat manipulation on dejection. Furthermore, dejection was significantly related to performance, and it mediated the relationship between the stereotype-threat manipulation and task performance. Thus, a mediation by dejection might point more to an activated promotion focus. Other studies, however, have found elevated levels of agitation, anxiety, and nervousness in the stereotype-threat condition (e.g., Josephs, Newman, Brown, & Beer, 2003; Osborne, 2001; Spencer et al., 1999), which points rather to an activated prevention focus, and Spencer et al. (1999) also found evidence that the performance deficits might have been mediated by anxiety. The reason Keller and Dauenheimer found stereotype threat to show a relationship with promotion focus emotions might be that they specifically instructed participants to solve as many items as possible (but did not call for correct answering). Furthermore, the Keller and Dauenheimer instructions may have reduced potential self-threat by not describing the test as diagnostic of ability. In contrast, several studies have found that stereotype-threat effects only occurred when the test was self-relevant and described as diagnostic of abilities (Aronson et al., 1999; Croizet & Claire, 1998; Steele & Aronson, 1995; Stone et al., 1999). Given this atypical picture in the study by Keller and Dauenheimer, it seems likely that mediation by dejection is restricted to circumstances when a promotion focus is activated before the stereotype-threat manipulation occurs. Indeed, Keller and Dauenheimer provided instructions that "were introduced to heighten the activation of promotion concerns and to lower possible prevention concerns" (p. 375), which created "specific circumstances eliciting promotion concerns" (p. 378), or, in other words, that induced a promotion focus before stereotype threat was manipulated. Thus, it might well be that the results by Keller and Dauenheimer point to yet another mechanism through which negative stereotypes can lead to performance deficits, which, however, are not typical of stereotype-threat effects as originally described by Steele and colleagues (e.g., Spencer et al., 1999; Steele & Aronson, 1995). To advance understanding of self-regulation under stereotype threat, more research should be addressed to the question of how regulatory strategies and emotions interrelate. There is little research looking at both outcomes simultaneously (but see Leung & Lam, 2003; Roney, Higgins, & Shah, 1995).

### *Mode of Stereotype Activation*

With regard to the mode of stereotype activation, from a regulatory focus perspective, the explicit induction of stereotypic beliefs used in this series of experiments should not yield different effects from a more implicit procedure. However, for these effects to occur it seems important that the induction procedure and the stereotype itself do not suggest the use of any specific strategies that are apt to override regulatory focus concerns. Therefore, a general stereotype (good vs. bad performance) was used in all of the studies. It should be noted that in some studies, the explicit activation of stereotypes did yield different results from an implicit one (e.g., Kray et al., 2001). Although strategic tendencies elicited by regulatory focus hinge only on the kind of outcome considered (negative vs. positive), other strategic choices seem to depend on the amount of elaboration on the stereotype and ways to counteract it.



Kray et al. (2001, Experiment 3), for example, found that women negotiated better than men when the stereotype of women being bad negotiators was activated explicitly but not when it was activated implicitly. In both conditions, task instructions stated that the perfect negotiator is rational and assertive and pursues his or her own interest throughout the negotiation. In the explicit condition, participants were further told that these personality characteristics tend to vary across gender. We think that this statement made women aware of their possible deficits and motivated them to form compensatory action intentions. It is important to note that in the context of this experiment, the compensatory (i.e., assertive) strategies were both available and achievable (Kray et al., 2001). These action intentions in turn might very well raise the chance of successfully negotiating and overriding prevention concerns because of their concrete nature. We thus assume that such strategic decisions can override the strategic inclinations elicited by regulatory focus in cases where such strategies follow from the contents of the stereotype, can be inferred from general knowledge, or are suggested by task instructions. In more academic performance tasks, on the other hand, like solving math problems, the explicit statement that women are known to perform worse than men does not point to any particular solution. In this case, according to our prediction, women will exhibit a tendency to use avoidance strategies like slowness, accurateness, avoidance of false alarms, and so on. It is possible that the explicit reference to the stereotype heightens the motivation to disconfirm it relative to an implicit one, but as long as the person does not possess the knowledge on the best strategy, this motivation does not necessarily lead to better results. To the contrary, if the task requires mostly approach strategies like creativity or speed, a strong prevention motivation will likely hinder its solution.

### *Lowered Accuracy Under Stereotype Threat*

If negative stereotypes can enhance or deteriorate performance, as is suggested in this article, the question arises why, in the studies investigating stereotype threat, negative stereotypes always had deteriorating effects on performance. For example, in the studies by Steele and Aronson (1995), Black participants in the diagnostic condition spent more time working on each individual task and gave fewer solutions than participants in the other conditions. However, lowered speed did not pay off in terms of heightened accuracy for them. They displayed both lowered accuracy and lowered speed. Of course, it is conceivable that mechanisms other than regulatory focus were responsible for these results. It should be noted, however, that they are not incompatible with a regulatory focus account.

One mechanism by which regulatory focus might lead to lowered speed and lowered accuracy in difficult tasks results from the finding in Studies 4 and 5 that negative stereotypes impede creative generation and breaking context-induced mental set. These effects should negatively impact performance in a great variety of tasks. Most difficult performance tasks require application of knowledge to some novel problem. Thus, they usually imply the transfer of knowledge and procedures from one context to another. Furthermore, an initial solution often needs to be abandoned in favor of a different one. Our results suggest that these aspects of task solution might be especially impeded by activated negative stereotypes, because they refer to task characteristics usually cap-

tured by creativity tasks (Amabile, 1996). Taken together, the vigilant, risk-averse processing style displayed by individuals in a prevention focus does not necessarily foster accuracy on difficult items. Whether prevention- or promotion-focus strategic means lead to more success on a task depends on the match, or regulatory fit (Higgins, 2000), between these means and the specific requirements of the task. For the kind of tasks used in the present studies, prevention means led to more accuracy. For more difficult tasks, this might not be the case.

The idea that generating different ways of solving a task is hindered by a prevention focus is also supported by Liberman et al. (2001), who found that a prevention focus leads to generating fewer hypotheses than a promotion focus. In addition, individuals in a promotion focus were faster to generate each hypothesis than individuals in a prevention focus. Furthermore, Crowe and Higgins (1997) found that individuals in a promotion focus used more different sorting criteria and listed more different characteristics for objects than did individuals in a prevention focus, who, in turn, were more likely to repeat sorting criteria and characteristics. Thus, individuals in a prevention focus seem to confine the search for a solution to one or a few hypotheses.

In sum, these findings show that individuals in a prevention focus compared with individuals in a promotion focus generate less creative solutions, produce fewer hypotheses, use categories less inclusively, tend to repeat criteria more, persevere more on initial responses, and perform worse on tasks of creative insight. It is easily conceivable that such a risk-averse, perseverant processing style can interfere with the search for the solution in difficult word problems. Yet these are exactly the kinds of problems used in stereotype-threat research. In such problems, usually, several hypotheses must be tested and several strategies must be tried out before a successful strategy is found. If a prevention focus limits the number of hypotheses generated and the creativity of ideas produced, the chances of finding a successful strategy are severely limited. Thus, to the extent that negative stereotypes indeed induce a prevention focus, they should lead to lowered accuracy on such tasks.

What can be learned from these results about the mediation of stereotype-threat effects? The findings indicate that the occurrence of stereotype-threat effects not only depends on the difficulty of the task and on the amount of time pressure but more specifically on the kind of process necessary to solve the task: Tasks that require going beyond the information given, testing several hypotheses, drawing new connections, and generating many solutions in a short time should generally suffer from the activation of a negative self-stereotype, whereas tasks that require the stringent application of criteria to material given (as in analytic tasks) should generally profit from the activation of negative stereotypes. Assuming that the tasks used in most stereotype-threat studies are of the former rather than the latter type, it is thus possible to explain stereotype-threat effects with the suggested regulatory processes.

Our results also imply that GRE tasks such as the quantitative word problem tasks often used in stereotype-threat research (e.g., Quinn & Spencer, 2001) and analytical tasks such as the ones used in our research differ with regard to the strategies necessary for successful performance and thus with regard to the effect stereotype activation has on them. It is interesting that Quinn and Spencer (2001) found stereotype-threat effects only when the

problems were presented as word problems that had to be transformed to their correct mathematical form to solve them. When instead the equation was given, no difference due to stereotype activation emerged. Thus, the amount of restructuring and mental search for applicable strategies necessary to solve a task seems to determine whether negative stereotypes have detrimental effects on performance. This is consistent with our regulatory model because restructuring and mental search processes have been shown to suffer from an activated prevention focus (Friedman & Förster, 2001).

The analytical tasks used in our study, on the other hand, made the consequent application of a few simple logical rules necessary. Such a task affords a thorough, careful processing style rather than a mental search for the correct strategy. In addition, the loosened time constraints allow this thorough processing style. In sum, to alleviate stereotype-threat effects, not only measures aimed at reducing the stereotype activation such as claiming that the test is gender neutral (see Spencer et al., 1999) but also measures aimed at changing the task requirements should be considered. These could be, for example, reducing time constraints or choosing tasks requiring more analytical and less creative strategies.

### *Reduction of Stereotype-Threat Effects*

As overwhelming and sad as stereotype-threat effects are and as important as it is to fully understand them, we believe that it is also important to identify conditions under which active negative stereotypes might have more positive than negative effects. This could eventually help in designing test situations that are fairer to stereotyped groups. One implication of the present research would be to investigate whether relaxed time constraints would yield the same results. It would also be interesting to see whether performance in tasks where accuracy is inherently more important or relevant, such as in medical diagnoses or court judgments, is equally impeded by negative stereotypes or if it can even be enhanced by prevention focus vigilance. Moreover, Studies 1 and 5 provide first evidence that the processes elicited by self-stereotypes can extend to tasks that are not achievement related, such as incidental memory or typicality judgments. Thus, it seems worthwhile to see when risky versus careful processing styles improve behavior unrelated to performance and task completion in general and when they impede it. This can be done by drawing on the existing and immensely growing body of literature on regulatory focus effects on a number of different outcomes. Thus, we think that a regulatory focus perspective can generate diverse research programs on stereotype threat that might not have been considered under other perspectives.

Our results also mean that the concept of performance deficits should be more thoroughly conceptualized. "Deficits" might be construed by certain biases of evaluation, which might then work against stereotyped groups. In most of the studies on stereotype threat, the tasks used and their administration might have been biased toward promotion strategies (being fast, working under time pressure) and against prevention-focus strategies (being thorough and careful). This mirrors the situation in everyday life where the concept of an exam is to measure achievement within a given time. It might be worthwhile to rethink this concept in order to give people using prevention-focus means a better chance. For teaching and students' evaluation this implies that self-paced projects might

more diagnostically reflect an applicant's capability to solve difficult problems than exams under time constraints, whereas exams would reflect efficient work on problems. In terms of research, it would be worthwhile to compare educational programs that endorse self-paced projects with those that endorse typical exams to see whether the latter produce even worse performances for stereotyped groups.

We are well aware of the fact that the complex phenomenon of stereotype threat can be caused by multiple psychological processes. Our research points to motivational principles of self-regulation, especially people's regulatory focus. Examining the processes that lead to stereotype threat is an important endeavor, which can ultimately lead to designing intervention strategies to stop the vicious circle that perpetuates the disadvantages of stereotyped groups.

### References

- Amabile, T. M. (1996). *Creativity in context*. Boulder, CO: Westview Press.
- Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science*, 12, 385–390.
- Amthauer, R. (1973). *Intelligenz-Struktur-Test (IST 70)* (4th ed.). Göttingen, Germany: Hogrefe.
- Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, 35, 29–46.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Baumeister, R. F., Hamilton, J. C., & Tice, D. M. (1985). Public versus private expectancy of success: Confidence booster or performance pressure? *Journal of Personality and Social Psychology*, 48, 1447–1457.
- Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General*, 130, 701–725.
- Brendl, C. M., & Higgins, E. T. (1996). Principles of judging valence: What makes events positive or negative? In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 95–160). San Diego, CA: Academic Press.
- Brown, R. P., & Josephs, R. A. (1999). A burden of proof: Stereotype relevance and gender differences in math performance. *Journal of Personality and Social Psychology*, 76, 246–257.
- Butler, J. L., & Baumeister, R. F. (1998). The trouble with friendly faces: Skilled performance with a supportive audience. *Journal of Personality and Social Psychology*, 75, 1213–1230.
- Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European Journal of Social Psychology*, 33, 267–285.
- Clare, G. L., Schwarz, N., & Conway, M. (1994). Affective causes and consequences of social information processing. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition: Vol. 2. Applications* (2nd ed., pp. 323–417). Hillsdale, NJ: Erlbaum.
- Crocker, J., Karpinski, A., Quinn, D. M., & Chase, S. K. (2003). When grades determine self-worth: Consequences of contingent self-worth for male and female engineering and psychology majors. *Journal of Personality and Social Psychology*, 85, 507–516.
- Croizet, J. C., & Claire, T. (1998). Extending the concept of stereotype and threat to social class: The intellectual underperformance of students from

- low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, 24, 588–594.
- Crowe, E., & Higgins, E. T. (1997). Regulatory focus and strategic inclinations: Promotion and prevention in decision-making. *Organizational Behavior and Human Decision Processes*, 69, 117–132.
- Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Förster, J., Grant, H., Idson, L. C., & Higgins, E. T. (2001). Success/failure feedback, expectancies, and approach/avoidance motivation: How regulatory focus moderates classic relations. *Journal of Experimental Social Psychology*, 37, 253–260.
- Förster, J., Higgins, E. T., & Idson, L. C. (1998). Approach and avoidance strength during goal attainment: Regulatory focus and the “goal looms larger” effect. *Journal of Personality and Social Psychology*, 75, 1115–1131.
- Förster, J., Higgins, E. T., & Taylor Bianco, A. (2003). Speed/accuracy decisions in task performance: Built-in trade-off or separate strategic concerns? *Organizational Behavior and Human Decision Processes*, 90, 148–164.
- Friedman, R. S., & Förster, J. (2000). The effects of approach and avoidance motor actions on the elements of creative insight. *Journal of Personality and Social Psychology*, 79, 477–492.
- Friedman, R. S., & Förster, J. (2001). The effects of promotion and prevention cues on creativity. *Journal of Personality and Social Psychology*, 81, 1001–1013.
- Friedman, R. S., & Förster, J. (2002). The influence of approach and avoidance motor actions on creative cognition. *Journal of Experimental Social Psychology*, 38, 41–55.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, 52, 1280–1300.
- Higgins, E. T. (1999). Promotion and prevention as a motivational duality: Implications for evaluative processes. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 503–525). New York: Guilford Press.
- Higgins, E. T. (2000). Making a good decision: Value from fit. *American Psychologist*, 55, 1217–1230.
- Higgins, E. T., Roney, C. J. R., Crowe, E., & Hymes, C. (1994). Ideal versus ought predilections for approach and avoidance distinct self-regulatory systems. *Journal of Personality and Social Psychology*, 66, 276–286.
- Higgins, E. T., & Tykocinski, O. (1992). Self-discrepancies and biographical memory: Personality and cognition at the level of psychological situation. *Personality and Social Psychology Bulletin*, 18, 527–535.
- Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, 11, 365–371.
- Isen, A. M. (1987). Positive affect, cognitive processes, and social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, pp. 203–253). New York: Academic Press.
- Isen, A. M., & Daubman, K. A. (1984). The influence of affect on categorization. *Journal of Personality and Social Psychology*, 47, 1206–1217.
- Josephs, R. A., Newman, M. L., Brown, R. P., & Beer, J. M. (2003). Status, testosterone, and human intellectual performance: Stereotype threat as status concern. *Psychological Science*, 14, 158–163.
- Keller, J., & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin*, 29, 371–381.
- Kray, L. J., Thompson, L., & Galinsky, A. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology*, 80, 942–958.
- Kunda, Z. (1999). *Social cognition: Making sense of people*. Cambridge, MA: MIT Press.
- Leung, C. M., & Lam, S. F. (2003). The effects of regulatory focus on teachers' classroom management strategies and emotional consequences. *Contemporary Educational Psychology*, 28, 114–125.
- Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. *Journal of Personality and Social Psychology*, 71, 1092–1107.
- Leyens, J. P., Desert, M., Croizet, J. C., & Darcis, C. (2000). Stereotype threat: Are lower status and history of stigmatization preconditions of stereotype threat? *Personality and Social Psychology Bulletin*, 26, 1189–1199.
- Liberman, N., Molden, D. C., Idson, L. C., & Higgins, E. T. (2001). Promotion and prevention focus on alternative hypotheses: Implications for attributional functions. *Journal of Personality and Social Psychology*, 80, 5–18.
- Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, 110, 472–489.
- Osborne, J. W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology*, 26, 291–310.
- Quinn, D. M., & Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues*, 57, 55–71.
- Roney, C. J. R., Higgins, E. T., & Shah, J. (1995). Goals and framing: How outcome focus influences motivation and emotion. *Personality and Social Psychology Bulletin*, 21, 1151–1160.
- Schooler, J. W., & Melcher, J. (1995). The ineffability of insight. In S. M. Smith & T. B. Ward (Eds.), *The creative cognition approach* (pp. 97–133). Cambridge, MA: MIT Press.
- Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, 10, 80–83.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4–28.
- Stangor, C., Carr, C., & Kiang, L. (1998). Activating stereotypes undermines task performance expectations. *Journal of Personality and Social Psychology*, 75, 1191–1197.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613–629.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797–811.
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). San Diego, CA: Academic Press.
- Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology*, 77, 1213–1227.
- Tanner, J., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, 61, 401–409.

Received April 22, 2002

Revision received February 3, 2004

Accepted February 12, 2004 ■