

**MIND OVER MATTER: ACTION IDENTIFICATION AS A MEDIATOR OF  
UNDERPERFORMANCE IN STEREOTYPE THREAT**

**Nikisha P. Williams**

Thesis Submitted in Partial Fulfillment of the Requirements  
for the Bachelor of Arts Degree at Hamilton College

May 4, 2001

Advisor: Jennifer Borton

### Abstract

Although several factors have been suggested as potential mediators of stereotype threat-related underperformance, a precise mediator remains elusive. In the current study, I tested the hypothesis, based on action identification theory (Vallacher & Wegner, 1986), that stereotype threat would alter the level at which threatened individuals identified a relevant task, such that they would not attend to the basic actions required to successfully complete it. Math-identified women and men completed a difficult test of math ability, described as either having previously produced gender differences or not. Results showed that, contrary to previous research, women in the threat condition did not underperform relative to other participants. In addition, level of action identification did not significantly correlate with task performance. Possible explanations for these findings are discussed.

## Mind Over Matter: Action Identification as a Mediator of Underperformance in Stereotype Threat

Relative to men, women do very poorly in math and the quantitative sciences (Brown & Josephs, 1999). Academic records, the sheer disparity in the number of women and men in math and science classes, and the gender gap evident on nearly every standardized test all point to the simple inability of women to perform in these domains. Female students have scored as many as 50 points lower than male students on the quantitative section of the SAT, and an astonishing 80 points lower on the same section of the Graduate Record Exam (Sadker & Sadker, 1993). It has been widely speculated that these gaps in test scores, as well as the relative absence of women in math and the quantitative sciences, is due to women's natural lack of ability in these domains, or to an internalization of this stigma (Brown & Josephs, 1999; Steele & Aronson, 1995). However, recent investigations of the underperformance of women in math and the quantitative sciences have revealed an alternative explanation that focuses not on some innate inferiority of women in these domains, but on situational factors that interfere with their performance. The goal of the present study is to investigate how these situationally induced factors hinder the performance of even the most talented women in math-oriented domains.

### What is Stereotype Threat?

Stereotype threat is the situational threat experienced when a negative stereotype about one's group exists in a valued domain (Steele & Aronson, 1995). Generally, it is unconsciously experienced as a self-evaluative threat, such that any action in the domain may be interpreted as indicating the validity of the stereotype in relation to the self (Steele, 1997, 1999; Steele & Aronson, 1995). For instance, the possibility that answering a question incorrectly or performing poorly on a quiz might lead others to believe that their math ability is inferior leaves women in

math classes vulnerable to stereotype threat. The vast majority of research focuses on groups considered the most susceptible to stereotype threat, including women in math-intensive domains and African American students in general (Aronson, Quinn, & Spencer, 1998; Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995).

Typically, stereotype threatened individuals perform significantly worse in the threatened domain relative to their own potential ability, as predicted by standardized test scores and past domain performance, as well as relative to their non-threatened peers, and individuals to whom the stereotype does not apply, an effect known as underperformance (Spencer et al., 1999; Steele, 1997; Steele & Aronson, 1995). Steele and Aronson (1995) found that making race salient before a test or describing a test as diagnostic of ability led African American students to perform significantly worse than both their White counterparts and other African American students not told the test was diagnostic of ability. Similarly, Spencer, Steele, and Quinn (1999) found that when a difficult math test was presented as having previously produced gender differences, women underperformed relative to men of equal skill, as well as relative to women who were not told the test had previously produced gender differences. Other studies of stereotype threat have replicated these results. Inducing threat by making a negative stereotype salient has consistently led stigmatized participants under threat conditions to underperform in comparison to participants in control conditions (Aronson, Lustina, Good, Keough, Steele, & Brown, 1999; Croziet & Claire, 1998; Stone, Lynch, Sjomeling, & Darley, 1999; Walsh, Hickey, & Duffy, 1999). One major goal of stereotype threat research has been to explain how the feelings of evaluative threat and pressure stemming from the stereotype lead to underperformance in critical domains.

Steele (1997) has identified four general features of stereotype threat that are useful in the investigation of underperformance. First, although the research focuses on the most noticeably affected groups, women and minorities, the effects of stereotype threat are not limited to individuals stigmatized on the basis of their race or gender. Steele (1997) describes stereotype threat as a general threat that may affect the performance of a member of any group about whom a negative stereotype exists. Empirical evidence supports this point. Croizet and Claire (1998) have found that students stigmatized on the basis of their socioeconomic status experience a form of stereotype threat similar to the threat experienced by African American students. In addition, the findings that athletes and even White men may experience some form of stereotype threat support the general nature of threat (Aronson et al., 1999; Spencer et al., 1999; Stone et al., 1999).

Second, because the content of the stereotypes about each group is unique, different groups experience stereotype threat in different forms and to varying degrees (Steele, 1997). For instance, although African American and White athletes are both stigmatized in the domain of sports performance, they experience stereotype threat in this domain differently because a distinct stereotype exists for each group (Stone et al., 1999). In one study (Stone et al., 1999), White participants experienced threat and underperformed on a golf task only when negative stereotypes about their lack of natural athletic ability were primed, whereas African American participants experienced threat and underperformed only when negative stereotypes about their sports intelligence were primed.

Third, one need not believe the relevant stereotype, or have prolonged exposure to it, to be vulnerable to stereotype threat (Steele, 1997). Aronson et al. (1999) induced stereotype threat in White men, a group for which no established negative stereotype exists. In this study, White

men took a very challenging math test while either primed to believe that Asian men usually outperformed White men on the test, or not. As in previous stereotype threat studies, participants in the threat condition underperformed significantly compared to participants in the control condition. Although the stereotype used was relatively uncommon, the implication that the task might show a race-related discrepancy in ability was enough to trigger stereotype threat.

Finally, the domain in which the stereotype exists must be self-relevant and important to the individual's self-identity; that is, the individual must be identified with the domain (Steele, 1997). Domain-identified individuals care deeply about their performance in the given domain, expect to do well on domain-related tasks, and perceive strong potential to succeed in the domain (Aronson et al., 1998; Steele, 1997). Their self-investment, strong ability, and high prospects in the domain leave domain-identified individuals extremely vulnerable to stereotype threat and, consequently, underperformance (Aronson et al., 1998, 1999; Spencer et al., 1999; Steele & Aronson, 1995).

Underperformance is one of the most apparent and detrimental effects of stereotype threat (Aronson et al., 1998), especially for domain-identified individuals. The self-threatening nature of the stereotype leads domain-identified individuals to question whether the stereotype accurately describes their own ability and performance (Aronson et al., 1998). Because of their strong ability, these individuals expect to do well. As a result, any stereotype that implies a lack of ability in the relevant domain is particularly distressing to them. Such stereotypes often lead domain-identified individuals to attempt to avoid poor performance, and thereby confirmation of the stereotype. This attempt to avert poor performance often has the opposite effect, serving only to depress performance even further (Aronson et al., 1998).

### Potential Mediators of Stereotype Threat-Related Underperformance

How do the elements of stereotype threat lead to underperformance? That is, what specifically does stereotype threat do that results in underperformance? Several factors have been suggested as mediators of the underperformance associated with stereotype threat. However, two factors frequently emerge from the literature as having the strongest effects on underperformance: anxiety and cognitive distraction (Aronson et al., 1998; Aronson, Quinn, & Spencer, 1998; Spencer et al., 1999; Steele & Aronson, 1995; Stone et al., 1999).

Anxiety is a response to a perceived danger or challenge and the inability to handle it in a satisfactory way. Generally, in order to arouse feelings of anxiety, a situation must be viewed as threatening and difficult, and must lead to thoughts of inadequacy and concern about the undesirable consequences of that inadequacy. Individuals in anxiety-provoking situations typically anticipate failure and loss of regard as a result of their inability to cope with the given situation (Sarason, 1980). Because stereotype threat presents a situational threat that, like anxiety, arouses concerns about ability and adequacy, anxiety makes sense as a possible cause of underperformance. Furthermore, previous research has shown that anxiety is consistently negatively correlated with performance (Sarason, 1972,1980). It seems logical that stereotype threat could arouse feelings of anxiety that in turn depress performance.

Indeed, past research shows that participants performing a task under stereotype threat report a range of anxious reactions. Steele and Aronson (1995) found that participants under stereotype threat reported experiencing significantly higher levels of anxiety than did participants in a control condition. Similarly, Spencer et al. (1999) found that when other factors were ruled out, participants under stereotype threat reported feeling marginally more anxious than did participants in a no threat condition. In addition, Stone et al. (1999) found that although anxiety

did not correlate significantly with performance, African American participants threatened on the basis of their sports intelligence showed a significant increase in anxiety after threat was induced.

Although the relationship between underperformance and anxiety has not been absolute or stable across different situations, results of stereotype threat studies strongly suggest that participants under threat experience a form of anxiety akin to test anxiety (Aronson et al., 1998; Spencer et al., 1999; Steele & Aronson, 1995; Stone et al., 1999). Sarason (1972) defined test anxiety as a susceptibility to self-centered thoughts that interfere with performance under evaluative conditions. Test anxiety may be divided into two separate components: emotionality and worry (Deffenbacher, 1980; Libert & Morris, 1967; Sarason, 1984; Wine, 1971). The emotionality component of test anxiety consists of one's awareness of physical arousal and tension (Deffenbacher, 1980). It may be triggered by an autonomic response to uncertainty in a testing situation (Libert & Morris, 1967). The worry component of test anxiety consists of a cognitive focus or preoccupation with the consequences of failure and concern about one's ability in relation to others (Libert & Morris, 1967; Morris, Davis, & Hutchings, 1981; Sarason, 1984). The emotionality and worry components both significantly contribute to test anxiety; however, past studies have found that only worry consistently and significantly correlates with test performance (Deffenbacher, 1980; Morris et al., 1981, Sarason, 1972; 1984).

Steele (1999) has suggested that worry may play a larger role than overall test anxiety in underperformance. Like stereotype threat, worry is induced by situational cues related to self-relevant performance evaluation and possible failure (Deffenbacher, 1980; Libert & Morris, 1967). Furthermore, the more cues a situation presents that imply poor performance, the more worry one experiences, and the more performance is depressed (Libert & Morris, 1967). By



definition, stereotype threat provides cues that lead to performance concerns and, therefore, worry. Some studies also suggest that evaluation apprehension, a more general form of anxiety triggered by a real or imagined evaluative audience, may also act as a mediator of underperformance in stereotype threat (Spencer et al., 1999). Research suggests that because both worry and evaluation apprehension debilitate task performance by distracting individuals from the task at hand (Sarason, 1984; Spencer et al., 1999; Wine, 1971), worry and evaluation apprehension may also act as cognitive distracters. Cognitive distracters direct attention away from task-relevant information, preoccupying individuals with thoughts and worries about the consequences of their performance. Other sources of cognitive distraction that may affect stereotype threat include frustration, stereotype activation, and negative thoughts (Stangor, Carr, Kiang, 1998; Steele, 1999; Steele & Aronson, 1995; Walsh et al., 1999).

Frustration provides a negative cue about ability. If one becomes frustrated with a difficult task, then one's ability to perform that task is called into question. For stereotype threatened individuals, this doubt about their capability raises the possibility that the negative stereotype is applicable to their performance in the given domain (Steele, 1997; Steele & Aronson, 1995). Frustration, like all sources of cognitive distraction, triggers performance concerns that divert attention away from the task. In addition, frustration may also interfere with performance by leading to a distracting emotional reaction (Walsh et al., 1999). However, neither of these notions has been formally investigated.

Individuals performing a task under stereotype threat have been found to experience significantly higher levels of stereotype activation than control participants. Steele and Aronson (1995) found that on a post-test measure of stereotype activation, Black participants in the threat condition completed significantly more word fragments with race-related words than did White

participants in the same condition. Stereotype activation may lead to performance deficits by lowering performance expectancies. Stangor, Carr, and Kiang (1998) showed that activation of negative stereotypes about ability led to uncertainty about task performance, even in the presence of positive feedback. If a negative stereotype is salient as other cues are raising concerns about one's capacity to perform a task, it is more likely that performance difficulties will be interpreted in relation to the stereotype. For threatened individuals, stereotype activation serves as one more negative cue that depresses performance by activating performance concerns.

Finally, distracting thoughts essentially encompass any negative thought that directly interferes with performance. Distracting thoughts and other possible mediators of underperformance compete with the relevant task for cognitive resources (Steele, 1999; Steele & Aronson, 1995). For instance, instead of concentrating on reading each question carefully, stereotype threatened test takers may be preoccupied by worries like the possibility that the stereotype is true.

Stereotype threat provides a serious intimidation such that threatened individuals experience a great deal of pressure in the given domain. Although this pressure is likely unconscious, it leads to cognitive distraction that may impair individuals' best thinking (Steele, 1999). Steele and Aronson (1995) found that although stereotype-threatened participants did not report experiencing significantly greater numbers of distracting thoughts than did participants in the control condition, the former spent more time and energy answering fewer questions incorrectly. This decreased speed and accuracy on the relevant task seems typical of participants working under stereotype threat. Aronson et al. (1999) found that stereotype-threatened participants spent more time per question, despite their depressed task performance.

Why do individuals performing a task under stereotype threat underperform if they spend at least as much, and usually more, effort on the task than control participants? Research suggests that the negative effects of stereotype threat lead to a state of impaired efficiency (Aronson et al., 1998; Steele, 1999; Steele & Aronson, 1995). Individuals experiencing impaired efficiency tend to exert a great deal of effort while still performing poorly (Aronson et al., 1998; Steele, 1999; Steele & Aronson, 1995). These participants work harder, but not necessarily smarter (Steele, 1999). Individuals working under threat may engage in actions such as re-reading questions and re-checking answers. On tasks such as standardized tests, where speed counts, these actions present a serious hindrance to performance. Threatened individuals likely underperform because they misdirect their effort, wasting time and energy ensuring that every response is correct, instead of concentrating on completing as many questions as possible (Steele, 1999).

Impaired efficiency may act as an umbrella mediator that incorporates the effects of the other possible mediators of stereotype threat. Some studies suggest that possible mediators such as worry and stereotype activation create an additive effect that leads to impaired efficiency (Steele, 1999; Steele & Aronson, 1995). Typically, all the proposed mediators trigger performance concerns. The subsequent need to relieve these concerns diverts attention from the task (Aronson et al., 1998; Steele, 1999; Steele & Aronson, 1995) and may result in inefficiency. However, even if impaired efficiency does act as a superordinate factor that incorporates the effects of multiple mediators, it is still unclear how the effects of mediators like evaluation apprehension, frustration and distracting thoughts interfere with performance efficiency.

### Possible Role of Action Identification

The theory of action identification (Wegner & Vallacher, 1986) offers one explanation of how the proposed mediators of underperformance lead to impaired efficiency at cognitive tasks. Action identification explains how we define, perform, and maintain specific actions. A main tenet of the theory is that for every action, an optimal way of thinking, or “identity” for that action exists (Wegner & Vallacher, 1986; Vallacher, Wegner, & Somoza, 1989). For example, an optimal identity for the act of walking might be “putting one foot in front of the other”. When the identity for a given act moves away from its optimal level, performance suffers (Seidel, Stasser, & Collier, 1998; Vallacher & Wegner, 1987; Vallacher et al., 1989). In principle, as long as it operates within the constraints of reality, the identity for an action may be defined in any way that the actor deems appropriate (Vallacher & Wegner, 1987). For instance, one might also define the act of walking as “going from here to there”. It is possible that stereotype threat leads individuals to define relevant domain-related actions in a non-optimal manner, leading to underperformance.

A single action may have multiple identities that each operate at different levels of comprehension (Wegner & Vallacher, 1986). For example, the action of “brushing my teeth” may be defined with both high level identities such as “caring for my dental hygiene” and “preventing tooth decay” and low level identities like “using toothpaste” and “moving the toothbrush”. At high levels of action identification, the act is usually defined in terms of a broad understanding of the action. High level identities focus on why the action is done, and what its possible consequences or implications may be. In addition, the action can be explained with relatively few details. In contrast, at low levels of action identification, the act is typically defined in terms of a comprehensive understanding of the action. Low level identities tend to

provide information about how the action is done and require that it be explained with a large amount of detail (Seidel et al., 1998; Vallacher & Wegner, 1987; Wegner & Vallacher, 1986). Whatever identity is most accessible when the action begins determines how the action will be defined and maintained. This initial identity for an action is known as its prepotent identity (Vallacher & Wegner, 1987; Wegner & Vallacher, 1986).

When multiple identities for a given action are available, the highest level is generally adopted as the prepotent identity (Vallacher & Wegner, 1987; Wegner, Vallacher, Macomber, Wood, & Arps, 1984). According to the theory, individuals tend to perform an action exactly as they define it. If one defines the act of “brushing my teeth” as “caring for my dental hygiene”, one will pay special attention to the quality and frequency of one’s tooth-brushing. Because high level identities provide the most information about the structure and consequences of an action, typically they present the best way to perform the action effectively (Vallacher & Wegner, 1987; Vallacher et al., 1989; Wegner, et al., 1984).

However, identifying an action at its highest level may not always be the most effective way to perform the act successfully. If an action proves too difficult to maintain at the level of its prepotent identity, then the actor may move toward a lower level of identification for that action (Wegner & Vallacher, 1987; Wegner et al., 1984). For example, if one adopts “getting exercise” as the prepotent identity for the act of swimming, but has difficulty executing the action, one may adopt the lower level identity of “kicking and floating,” which provides a more detailed description of the action, as the new prepotent identity (Seidel et al., 1998; Vallacher & Wegner, 1987; Wegner et al., 1984). This process of balancing high and low level identities with respect to performance success to yield a new identity is known as emergence. Several different identities for an action may emerge during the course of performing the act (Vallacher &

Kaufman, 1996; Wegner et al., 1984). Ultimately, the level at which an action is maintained represents a compromise between high level identities that describe why the action is performed and low level identities that describe how the action is performed (Vallacher & Kaufman, 1996; Vallacher et al., 1989). Ideally, this compromise results in an optimal level of identification for the given action.

The difficulty or familiarity of an action often determines the optimal level of identification for that action (Vallacher & Wegner, 1986; Vallacher et al., 1989). Action identification theory suggests that actions that are familiar and personally easy are best maintained at high levels, whereas more difficult actions are best maintained at lower levels (Vallacher, & Wegner, 1985; Vallacher et al., 1989). Easy and familiar actions tend to become automatic. One does not therefore require a high level of detail to perform these actions effectively. In fact, describing a personally easy action in low level terms may actually disrupt the action. Lower than optimal levels of action identification deprive the action of its automaticity and fluidity by unnecessarily promoting basic elements of the act (Vallacher et al., 1989; Vallacher, Wegner, McMahan, Cotter &, Larsen, 1992). For example, an experienced swimmer who no longer needs to think of swimming in basic terms such as “kicking and floating” may find it difficult to maintain the action if asked to attend to a low level element of the act, like kicking.

In contrast, personally difficult or unfamiliar actions require a high level of detail for effective performance. Higher than optimal levels of action identification lack sufficient information about the basic elements of an action to allow effective performance of that action (Vallacher et al., 1989, 1992). For instance, a beginning swimmer most likely lacks experience with the basic elements of the action needed to successfully maintain it in terms of “getting

exercise”. Although the criteria for when to use high versus low levels of action identification seem relatively intuitive and simple, individuals tend to move toward identities that describe the action in relatively high level terms (Vallacher et al., 1992), leading to a detrimental disparity between an action’s prepotent and optimal identities.

Generally, any disparity between the prepotent and optimal identity for an act stems from the relatively unstable nature of low level identities. Low level identities are relatively constrained; unlike high level identities, they lack the flexibility to be extended to include a new understanding of a given act. At low levels of action identification, if a higher level identity for the act emerges it simply replaces the original prepotent identity (Vallacher & Kaufman, 1996; Vallacher & Wegner, 1985). For example, a beginning swimmer who encounters the high level identity “getting exercise” will cease to define the act of swimming in low level terms like “kicking and floating” in favor of the higher level identity. Low level identities are highly susceptible to contextual information (Wegner & Vallacher, 1985); any situational suggestion that a higher level of identification is possible for a given action may lead an actor to adopt the more comprehensive understanding of the act, even if the new identity does not describe the action well enough for effective performance (Vallacher, & Wegner, 1985; Vallacher et al., 1992, Wegner et al., 1984).

For instance, Wegner, Vallacher, Macomber, Wood, and Arps (1984) were able to induce college students to endorse the act of “going to college” in terms of “improving/impairing one’s sex life,” by first asking them to describe the act in low level terms such as “reading textbooks” and “studying”. Clearly, a prepotent identity of “improving/impairing my sex life” is completely unrelated to academic pursuit, and therefore presents a nonoptimal approach to the act of “going to college”. However, because the students adopted a low level prepotent identity for the act,

they were highly susceptible to the emergence of any new high level identity to which they were exposed, regardless of how well it allowed them to maintain the act (Vallacher et al., 1984). This nonoptimal shift in the level of action identity occurs because the lower the level of action identification is, the more unstable it tends to be, and the more readily suggestions of higher level identities are accepted (Vallacher & Wegner, 1985).

The theory of action identification predicts that the level of action identification adopted will affect performance such that the closer the prepotent identity is to the optimal identity, the better one will perform the given action. When situational cues induce actions to shift toward nonoptimal identities, performance will suffer (Vallacher & Wegner, 1987; Vallacher et al., 1989). For personally difficult actions, this nonoptimal shift typically involves moving from a low level of action identity to a higher one, leading to poor performance (Vallacher et al., 1992).

Testing or exam situations are among those acts best identified at low levels (Vallacher & Kaufmann, 1996; Wegner & Vallacher, 1986). Vallacher and Wegner (1985) found that college students who endorsed relatively low levels of action identification performed better on an exam than did those who identified the action at high levels. Because taking an exam usually involves expending high levels of effort, even for individuals skilled in the testing domain, the act of “taking an exam” stands to benefit from detailed descriptions of the elements involved in performing the act (Vallacher & Kaufmann, 1996). Paradoxically, however, because test taking is best identified at these low level identities, it is highly susceptible to situational cues that lead to the emergence of higher level identities (Vallacher & Kaufmann, 1996; Vallacher & Wegner, 1985).

Generally, situational cues that create pressure to do well lead to high level identities for an action. Situations like exams especially favor high level identities that are self-evaluative in



nature (Vallacher et al., 1989, 1992). Instead of allowing an understanding of the basic actions required to successfully complete the task itself, self-evaluative identities shift focus onto the motivation behind, and implications of, the action (Vallacher et al., 1989, 1992). Self-evaluative high level identities can be triggered by situational cues similar to those that produce stereotype threat. Research has suggested that factors such as distracting thoughts, evaluation apprehension, and worry about failure can lead individuals to shift to nonoptimal levels of action identification (Seidel et al., 1998; Vallacher et al., 1992). This shared tendency of stereotype threat and action identification to impair effective performance by inducing individuals to think about an action in a maladaptive manner suggests that action identification level may play a role in the mediation of underperformance in stereotype threat.

#### Overview of the Current Study

The present study investigated the possibility that individuals taking exams under stereotype threat underperform because they identify the action of taking the exam at a higher than optimal level, hence impairing efficiency. Women and men majoring in math and the quantitative sciences took a difficult test of math ability after being told that the test either had or had not previously produced gender differences. Participants were then asked to generate several statements that they felt described the action of “taking an exam,” and to rate how well each statement applied to their personal performance. In addition, participants completed an established measure of action identification for the act of “taking an exam”. Participants’ levels of cognitive interference, evaluation apprehension, worry, frustration, and stereotype activation were also measured.

The primary hypothesis was that level of action identification would mediate stereotype threat-related underperformance. I expected women in the threat condition to endorse higher

levels of action identification than all other participants; these women were expected to focus their limited attention and energy on the ability implications of their performance or on the applicability of the stereotype, rather than the basic actions required to perform well on the test. Higher levels of action identification were expected to be directly related to poorer performance.

## Method

### Participants

In order to maximize the likelihood that participants would be domain-identified with math, students whose field of academic study required a high level of mathematical ability were targeted for recruitment. In addition, only those students who passed the introductory calculus course with a grade of B or better were contacted. Forty-seven sophomore, junior, and senior students majoring or minoring in Math, and/or majoring in Physics, Chemistry or Economics participated in the study. To qualify for inclusion, each participant was also required to agree, by responding with a rating of 5 or higher on an 11-point Likert-type scale, with both of the following statements: “It is important to me that I am good at math” and “I am good at math”. Spencer, Steele and Quinn (1999) have used similar questions to assess participants’ domain identification. Participants were paid \$5 as compensation for their time.

### Design

Participants were randomly assigned to either a stereotype threat or no threat condition, in a 2 (gender) x 2 (threat vs. no threat) between-subjects design. Level of action identification and test performance served as the primary dependent measures. The degree to which participants experienced cognitive interference, evaluation apprehension, frustration, worry, and stereotype activation was also measured.

### Dependent Measures

Test of math performance. Because stereotype threat is most effectively induced when participants are stretched to the limit of their ability, 15 questions drawn from the Graduate Record Examination in Mathematics served as the performance measure. Past research has suggested that domain-identified individuals with strong math skills perceive this exam to be more challenging than the quantitative subsection of the GRE general test (Spencer et al., 1999). The questions used on the test are presented in Appendix A.

Measures of action identification. I included two measures of action identification in the post-experimental questionnaire. First, participants were asked to compose several statements they believed described the action of “taking an exam” on an open-ended questionnaire. Each statement was then rated on a 7-point scale according to how well participants felt the statement applied to their own actions during the exam they had just completed (1 = does not at all apply, 7 = applies very well). Following the open-ended measure, participants were asked to respond to 28 statements on a closed-ended questionnaire established by Vallacher and Wegner (1985). Each statement on the questionnaire was presented as a sentence fragment phrased in the first person, present tense, using the gerund form of the verb (e.g., “thinking about my answers”). These statements were rated on a 7-point scale according to how well participants thought each one described the act of “taking an exam” for the test they had just completed (1 = describes very poorly, 7 = describes very well). Appendices B and C contain both the open- and closed-ended action identification questionnaires, respectively.

Sources of cognitive distraction. A modified version of Steele and Aronson’s (1995) stereotype activation task measured the extent to which negative stereotypes about women’s math ability were active in participants’ thoughts. Participants were asked to complete 12 word fragments by filling in the missing letters (e.g., S T \_ \_ \_ \_; stupid). The target words were

interspersed in columns with 6 distracter words so that each column contained 4 target word fragments and two distracters. The target words were chosen based on the results from a pre-test questionnaire. Approximately 50 introductory level psychology students were asked to list words they thought reflected society's image of women's math ability. The words were then coded into positive, negative, and neutral categories and a single word was chosen to represent each category. The 12 negative words that represented the categories that contained the largest number of words were chosen as the target words. The target words and their corresponding fragments are listed in Appendix D.

To assess the extent to which participants experienced various other sources of cognitive distraction during the performance task, scales measuring frustration, cognitive interference, evaluation apprehension, and worry were included in the post-experimental questionnaire. Participants completed a scale measuring frustration, on which they rated 8 adjectives related to the feeling of frustration according to how well each one described how the exam made them feel (e.g., "irritated"). Participants also completed the Cognitive Interference Questionnaire (Sarason, 1980), which is composed of two parts. On the first part, participants rated 12 statements that described some thoughts they might have experienced during the test (e.g., "I thought about how poorly I was doing"). On the second part, participants rated a global statement of how often they believed their thoughts had wandered during the math test. In addition to the CIQ and frustration scale, participants completed The Evaluation Apprehension Scale (Spencer et al., 1999) and the worry component of the Worry-Emotionality Scale (Morris et al., 1981). On these two scales they rated 5 statements about their feelings and attitudes in relation to the math test (e.g., "I feel others will be disappointed in me") on a Likert-type scale.

Questions from the two scales were randomly intermixed and presented together. Appendix E contains all the measures used to assess these sources of cognitive distraction.

### Procedure

Participants were tested individually in sessions lasting about 40 minutes. After the participants arrived at the laboratory, the experimenter explained the purpose of the study and secured informed consent. All participants were told that the study was an examination of the mental processes underlying strong quantitative ability. However, the experimenter gave different descriptions of the math test to participants in the two conditions. Participants in the stereotype threat condition were told that the math test had previously produced gender differences; in contrast, participants in the no stereotype threat condition were told that the test had not previously produced gender differences. Previous researchers have successfully used a similar manipulation to induce stereotype threat in domain-identified individuals (Spencer et al., 1999; Steele & Aronson, 1995).

After the consent forms were signed and collected, the experimenter escorted participants to a small room where they were seated at a computer used to administer the math test. The experimenter explained the format of the exam, and informed the participants that they should expect it to be difficult. Each participant was given 25 minutes to work on the math test. The experimenter started the computer program, made sure the participant had no questions, and then left the room. After 25 minutes had elapsed, the experimenter returned with the post-test questionnaire. In order to avoid raising participants' suspicions about the true nature of the experiment, the post-test measures of action identification and cognitive distraction were presented as measures of how individuals with strong math ability think.

Participants completed the global measure of thought distraction from the CIQ first, followed by the open- and closed-ended action identification questionnaires, respectively. Participants then completed the frustration scale, followed by the questionnaire portion of the CIQ, and the worry and evaluation apprehension scales. The stereotype activation measure was completed last to prevent it from biasing participants' responses on the other measures. Finally, participants completed a short questionnaire comprised of some basic demographic questions and a manipulation check (see Appendix F). Once participants had finished the questionnaire, they were thoroughly debriefed, thanked, and asked not to divulge the true nature of the experiment to other potential participants.

## Results

### Manipulation Check

Three participants were removed from the analysis because they failed to agree with the domain identification items with a rating of 5 or higher. A chi-square analysis on the manipulation check item confirmed that participants correctly reported whether or not the math test had been described as previously producing gender differences,  $\chi^2(1, N = 42) = 13.74, p < .0001$ . That is, participants in the threat condition were more likely to report that the test had previously been found to produce gender differences, and participants in the no threat condition were more likely to report that the test had not previously been found to produce gender differences. Thus, the manipulation was successful.

### Test Performance

Math test performance served as the primary dependent measure. Using participants' self-reported scores from the quantitative section of the SAT I as a covariate (see Spencer, Steele, & Quinn, 1999), an ANCOVA was performed on the number of math questions participants answered correctly by condition and gender. Five questions were eliminated from the analysis because so few participants were able to answer them correctly. Results from the remaining 10 questions yielded a marginally significant gender x condition interaction,  $F(1, 38) = 3.17, p < .08$ . There were no main effects. To determine the nature of the interaction, two independent groups  $t$ -tests were performed, comparing the mean test scores of women and men within each condition. As illustrated in Figure 1, results from the  $t$ -tests revealed that women in the no threat condition ( $M = 3.50, SD = 1.08$ ) scored marginally lower than did men in the no threat condition ( $M = 4.73, SD = 1.90$ ),  $t(19) = 1.79, p < .08$ . There was no significant difference between the test scores of women and men in the threat condition ( $M = 3.92, SD = 1.97$  for women;  $M = 3.73, SD$

= 1.90 for men). Thus, the direction of the interaction was inconsistent with the prediction that women in the stereotype threat would exhibit the lowest performance.

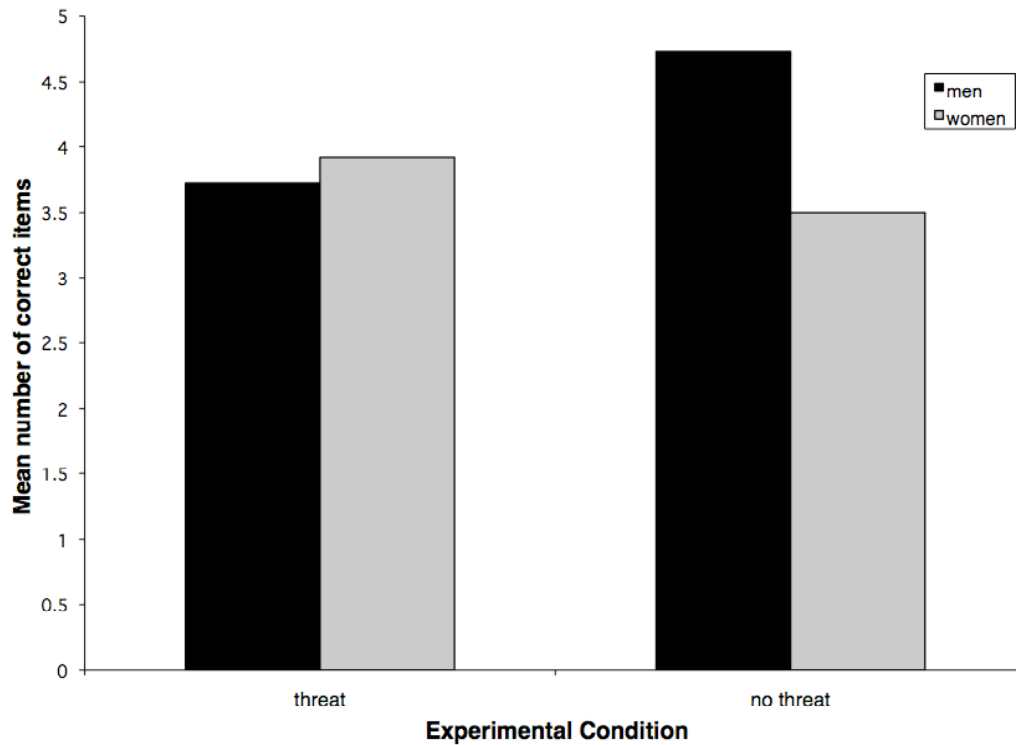


Figure 1. Mean number of math items answered correctly as a function of gender and condition.

### Action Identification

Open-ended measure. Participants' statements on the open-ended measure of action identification were coded into high and low level identities by two independent raters. Inter-rater agreement was calculated to be .81. After the statements had been coded, a mean rating of how strongly participants agreed with their self-generated statements was calculated for both the high and low level identities. The mean ratings were calculated by dividing the sum of participants' ratings for each type of statement (high /low level identity) by the total number of statements of



that type. Thus, a high mean rating indicated that participants believed that their self-generated identity statements described their actions during the exam very well, whereas a low mean rating indicated the opposite.

Two separate 2 (gender) x 2 (condition: threat vs. no threat) analyses of variance were performed to examine mean ratings of high and low levels of action identification. The ANOVA on mean ratings of high level identities produced no significant effects. The ANOVA on mean ratings of low level identities yielded a significant main effect for condition. Participants in the threat condition rated their self-generated low level identities significantly higher ( $M = 6.21$ ,  $SD = .65$ ) than did participants in the no threat condition ( $M = 5.71$ ,  $SD = .76$ ),  $F(1, 39) = 5.14$ ,  $p < .05$ . The ANOVA also revealed a significant gender x condition interaction,  $F(1, 39) = 4.24$ ,  $p < .05$ . A comparison of the means found that women in the no threat condition ( $M = 6.08$ ,  $SD = .79$ ) rated their low level statements significantly higher than did men in the no threat condition ( $M = 5.42$ ,  $SD = .61$ ),  $t(18) = -2.11$ ,  $p < .05$ , whereas no gender differences emerged in the threat condition.

Closed-ended measure. High and low level identity statements from the 28-item questionnaire were analyzed separately. As with the open-ended measure, results from the ANOVA performed on the mean ratings for the high level statements were not significant. The ANOVA on the low level statements revealed a marginally significant main effect for gender. Women ( $M = 4.50$ ,  $SD = .72$ ) rated the low-level items on the scale slightly higher than did men ( $M = 4.12$ ,  $SD = .75$ ),  $F(1, 40) = 3.38$ ,  $p < .08$ . In addition, results from the ANOVA revealed a significant gender x condition interaction,  $F(1, 40) = 4.39$ ,  $p < .05$ . As with the open-ended measure, women ( $M = 4.58$ ,  $SD = .88$ ) in the no threat condition rated the low level statements

significantly higher than did men in that condition ( $M = 3.74$ ,  $SD = .74$ ),  $t(19) = -2.35$ ,  $p < .05$ , whereas there was no gender difference in the threat condition.

### Correlation Between Action Identification and Test Performance

To determine whether higher levels of action identification were associated with lower levels of performance, correlations were conducted between test performance and the open- and closed-ended measures of action identification. No significant correlations emerged (see Table 1), thus no mediational analysis was conducted.

### Sources of Cognitive Distraction

The CIQ, evaluation apprehension, worry, and frustration scales were all moderately to strongly reliable (see Table 2 for the Cronbach's  $\alpha$  for each scale). However, results from the 2(gender) x 2(condition) factorial analyses of variance performed for each scale were not significant. The ANOVA performed on the stereotype activation measure did yield a significant main effect for gender,  $F(1, 40) = 6.27$ ,  $p < .01$ . Women ( $M = 1.18$ ,  $SD = 1.30$ ) completed significantly more word fragments with the target stereotype related words than did men ( $M =$

Table 1

### Correlations Among Ratings of Low Level Action Identities, Stereotype Activation, and Performance

	Math Performance
Low level action identification (Open-measure)	-.23
Low level action identification (Closed -measure)	-.07

Stereotype Activation -.18

---

$p > .05$  for all

Note. None of the correlations is statistically significant

Table 2

Cronbach's  $\alpha$ s for measures of cognitive distraction

---

Dependent Measure	Cronbach's $\alpha$
Cognitive Interference Questionnaire	.82
Evaluation Apprehension	.87
Worry component of Worry-Emotionality Scale	.75
Frustration Scale	.87

---

.36,  $SD = .73$ ). However, as shown in Table 1, stereotype activation did not significantly correlate with performance,  $r(42) = -.18$ ,  $p > .05$ .

Impaired Efficiency

Level of impaired efficiency was evaluated by calculating the mean number of seconds participants spent answering each question. A 2 (gender) x 2 (condition) ANOVA performed on this measure yielded no significant results; however, the pattern indicates a possible trend for a gender x condition interaction. As illustrated in Figure 2, women in the threat condition appeared

to spend more time per question than did men in that condition. However, these findings do not offer any support to the hypothesis that women in the threat condition would experience the highest level of impaired efficiency. The analysis of the mean scores for math test performance found that women in the threat condition performed better than men in the threat condition did, though not significantly so.

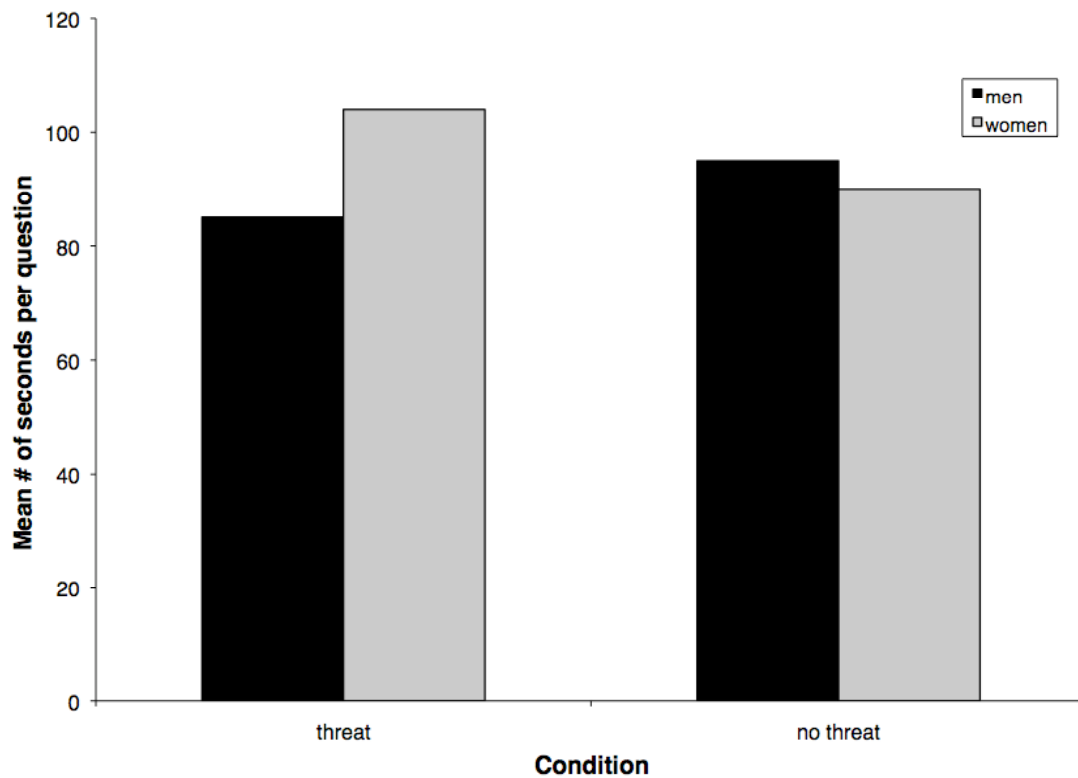


Figure 2. Amount of time spent per question as a function of gender and condition.

Because the results failed to produce a significant stereotype threat effect in the predicted direction, and none of the dependent measures correlated significantly with performance, no mediational models were generated.

## Discussion

Math test performance served as the primary dependent measure. The purpose of the test was to establish that the underperformance effect found in previous stereotype threat literature could be replicated with the current research population so that possible mediators could be investigated. Consistent with past research, I hypothesized that women in the stereotype threat condition would significantly underperform compared to women in the no stereotype threat condition and men in both conditions. However, the data failed to support this hypothesis. There was no significant difference between the test performance of women and men in the threat condition. In addition, contrary to my predictions and previous stereotype threat research, women in the no threat condition performed worse in that condition than men did. Given that the *p*-value for the interaction did not reach conventional significance levels, the data suggest that there was no statistically significant stereotype threat effect.

Typically, domain identified individuals are most susceptible to the performance depressing effects of stereotype threat (Steele, 1997). One possible reason that the test performance of women in the threat condition was not affected by stereotype threat could be that they simply were not domain identified enough to experience threat. However, two findings refute this explanation. First, all the participants who failed to agree with both domain identification questions (“I am good at math” and “It is important to me that I am good at math”) were removed from the analysis. Second, an ANOVA performed to check for differences in domain identification across gender and condition revealed no significant results. Although all the participants should have been domain-identified as a result of the recruiting process, it is possible that participants in the present study were not as domain-identified as were the

participants in past stereotype threat research, preventing them from experiencing the threat effect. However, this explanation is unlikely.

Another possible explanation for the failure of a threat effect to emerge could be differences in ability level. However, this possibility is refuted by the fact that analyses of variance revealed no significant difference between the ability of women and men in the two conditions, as measured by their self-reported SAT scores and the number of college math courses completed.

Overall, results from the open- and closed-ended measures of level of action identification were not consistent with my predictions. Participants' ratings of high level action identities did not differ significantly as a function of condition or gender on either the open-ended or closed-ended measures of action identification. Although I did not make any specific predictions about participants' rating of low level identities, analyses of these data did yield some interesting results. As measured by the both the open- and closed-ended action identification questionnaire, women in the no threat condition rated low-level identities higher than did men in that condition. However, participants' low level action identification ratings did not correlate significantly with performance. This finding is inconsistent with the prediction that performance on a difficult exam would benefit from low-level action identification.

With the exception of stereotype activation, analyses of the sources of cognitive distraction (cognitive interference, worry, evaluation apprehension, and frustration) did not yield any significant results. These data make sense given the failure of the results to produce a significant stereotype threat effect. Furthermore, failure of the scales to show any significant results cannot be attributed to lack of reliability. The Cronbach's  $\alpha$ s for each scale were acceptable. The one significant finding for cognitive distraction was a main effect for the

stereotype activation measure. Women showed significantly higher levels of stereotype activation than men. However, this finding is consistent with the fact that the negative stereotype with which participants were primed referred directly to women's math performance. In addition, level of stereotype activation was not significantly correlated with test performance.

Finally, analyses of participants' levels of impaired efficiency did not produce any significant differences between women and men in the two conditions. Although the results showed that women in the threat condition appeared to spend more time per question than did men in that condition, they also performed better than those men.

One possible explanation for the overall lack of significant findings could lie in the somewhat unique academic environment for math at Hamilton College. The math professors are acutely aware of the negative stereotypes surrounding women's mathematical ability and performance. In recent years they have made an extra effort to personally encourage and attend to women who major in the department, thus creating a supportive learning environment that minimizes gender stigmatization (R. Bedient, personal communication, April 22, 2001). Steele (1997) has suggested that learning environments that allow threatened individuals to perform without being held under the suspicion of a negative stereotype can attenuate the performance-depressing effects of stereotype threat. It is possible that there were some individual differences among the math majors that were not accounted for, which may have impacted the findings. These unexplained individual differences might have contributed to some type of stereotype threat buffer. Perhaps this buffer allowed the women in the threat condition to disregard the stereotype's potential relevance to their performance, negating threat, or at least preventing the stereotype threat from distracting them enough to lead to underperformance.

The current body of research has provided an excellent framework in which to study the many factors contributing to, and the negative consequences of, stereotype threat. However, to date it has not empirically supported a way to attenuate the performance-depressing effects of threat, leaving threatened individuals swimming against the current with no relief in sight. In 2000, 44% of high school graduates, nearly 2 million students, took the SAT (Cloud, 2001). On average, American students will take 12 standardized tests in the course of their years of education. As standardized exams become a more common and prevalent way to measure ability, students taking these exams have begun to give them more credit as accurate measures of their ability (Sadker & Sadker, 1993).

This increase in the use and credibility of standardized tests poses an extremely difficult situation for individuals affected by negative stereotypes. Most standardized tests induce situational cues that are all but guaranteed to trigger stereotype threat, making it almost impossible for threatened individuals to perform to their full potential. Results from the present study suggest that one method of reducing the effects of stereotype threat may be to alter the threatening situation. This finding is especially encouraging given that compared to internal cognitions, situational factors are relatively easy to change. Fostering an environment in which threatened individuals are secure in the knowledge that they will not be evaluated in terms of a negative stereotype may protect them from the underperformance effect. Future researchers should focus on finding empirical evidence for this possibility, and should investigate ways to apply potential data in developing real world solutions for the growing problem of underperformance on standardized tests.



## References

- Aronson, J., Lustina, M. J., Good, C., & Keough, K. (1998). When White men can't do math: Necessary and sufficient factors in stereotype threat. Journal of Experimental Social Psychology, *29*, 29-46.
- Aronson, J., Quinn, D., & Spencer, S. J. (1998). Stereotype threat and the academic underperformance of women and minorities. In J. K. Swim and C. Stangor (Eds.), Prejudice: The target's perspective (pp. 83-103). San Diego, CA: Academic Press.
- Cloud, J. (2001, March 12). Should SATs matter? Time, *157*, 164-170.
- Croizet, J., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic status. Personality and Social Psychology Bulletin, *24*, 588-594.
- Deffanbacher, J. L. (1980). Worry and emotionality in test anxiety. In I. G. Sarason (Ed.), Test anxiety: Theory, research and applications (pp. 111-128). New York: Erlbaum.
- Libert, L. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. Psychological Reports, *20*, 975-978.
- Morris, L. W., Davis, M. A., & Hutchings, C. H. (1981). Cognitive and emotional components of anxiety: Literature review and a revised worry-emotionality scale. Journal of Educational Psychology, *73*, 541-555.
- Sadker, M. & Sadker, D. (1993). Failing at fairness: How America's schools cheat girls. New York: Charles Scribner's Sons.
- Sarason, I. G. (1972). Experimental approaches to test anxiety: Attention and uses of information. In C. D. Spielberger (Ed): Anxiety: Current trends in theory and research: Vol.2 (pp. 381-405). New York: Academic Press.

Sarason, I. G. (1980). Introduction to the study of test anxiety. In I. G. Sarason (Ed.), Test anxiety: Theory, research, and applications (pp. 111-128). New York: Erlbaum.

Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. Journal of Personality and Social Psychology, *46*, 929-938.

Seidel, S. D., Stasser, G. L., & Collier, S. A. (1998). Action identification theory as an explanation of social performance. Group Dynamics: Theory, Research and Practice, *2*, 147-154.

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. (1999, August). Thin ice: Stereotype threat and the Black college student. The Atlantic Monthly, 44-54.

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.

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.

Vallacher, R. R., & Kaufman, J. (1996). Dynamics of action identification: Volatility and structure in the mental representation of behavior. In P. M. Gollwitzer & J. A. Bargh (Eds.): The psychology of action: Linking cognition and motivation to behavior. New York: Guilford.

Vallacher, R. R. & Wegner, D. M. (1985). A theory of action identification. Hillsdale, NJ: Erlbaum.

Vallacher, R. R. & Wegner, D. M. (1987). What do people think they are doing? Action identification and human behavior. Psychological Review, *94*, 3-15.

Vallacher, R. R., Wegner, D. M., & Somoza, M. P. (1989). That's easy for you to say: Action identification and speech fluency. Journal of Personality and Social Psychology, *56*, 199-208.

Vallacher, R. R., Wegner, D. M., McMahon, S.C., Cotter, J., & Larsen, K. (1992). On winning friends and influencing people: Action identification and self-presentational success. Social Cognition, *10*, 335-355.

Walsh, M., Hickey, C., & Duffy, J. (1999). Influence of item content and stereotype situation on gender-differences in mathematical problem solving. Sex Roles, *41*, 219-240.

Wegner, D.M., & Vallacher, R.R. (1986). Action identification. In R. M. Sorrentino, & E. T. Higgins (Eds.): Handbook of motivation and cognition: Foundations of social behavior (pp. 550-582). New York: Guilford.

Wegner, D. M., Vallacher, R. R., Macomber, G., Wood, R., & Arps, K. (1984). The emergence of action. Journal of Personality and Social Psychology, *46*, 269-279.

Wine, J. (1971). Test anxiety and the direction of attention. Psychological Review, *76*, 92-104.

## Appendix A: Math Test

**Each question has five answer choices. Pick the one that you think is correct or that best answers the question. You will have 25 minutes to complete the test.**

1. \_\_\_\_\_ The graph of the arccosine function is the graph of the arcsine function
  - (A) translated horizontally  $\frac{\pi}{2}$  units to the right
  - (B) first reflected in the horizontal axis and then translated vertically  $\frac{\pi}{2}$  units upward
  - (C) first translated horizontally  $\frac{\pi}{2}$  units to the left and then reflected in the horizontal axis
  - (D) first translated vertically  $\frac{\pi}{2}$  units downward and then reflected in the vertical axis
  - (E) translated horizontally  $\frac{\pi}{2}$  units to the left
  
2. \_\_\_\_\_ If  $f(x) = e^x - e^{-x}$ , then  $[f'(x)^2] - [f(x)]^2$  equals
  - (A) 4
  - (B)  $4e^{-2x}$
  - (C)  $2e^{-2x}$
  - (D) 2
  - (E)  $2e^x$
  
3. \_\_\_\_\_ The domain of  $f(x) = \frac{\sqrt[3]{x+2}}{x-6}$  is given by
  - (A)  $(6, +\infty)$
  - (B)  $[-2, +\infty)$
  - (C)  $R \setminus \{-2, 6\}$
  - (D)  $[-2, +\infty) \setminus \{6\}$
  - (E)  $R \setminus \{6\}$
  
4. \_\_\_\_\_ Let  $x_n = \frac{n^n}{n!}$  for  $n = 1, 2, 3, \dots$ . Then  $\lim_{n \rightarrow +\infty} \frac{x_{n+1}}{x_n}$  equals
  - (A)  $\sqrt{e}$
  - (B)  $e$
  - (C)  $\sqrt{e^3}$
  - (D)  $e^2$
  - (E)  $e^{-1}$

5. \_\_\_\_\_ The derivative of  $f(x) = \int_0^x \frac{\cos xt}{t} dt$  is
- (A)  $-\frac{\cos 2x^2}{x}$  (D)  $-\frac{\sin 2x^2}{x}$
- (B)  $\frac{1}{x}[1 + 2 \cos x^2]$  (E)  $\frac{\cos x^2}{x}$
- (C)  $\frac{1}{x}[1 + 2 \sin x^2]$
6. \_\_\_\_\_ Which of the following numbers is divisible by 9?
- (A) 7224466 (D) 5224466
- (B) 9224466 (E) 1224466
- (C) 3224466
7. \_\_\_\_\_ The inflection point for  $f(x) = \frac{\ln x}{x}$  occurs at  $x =$
- (A)  $\sqrt{e}$  (D)  $e^{-1}$
- (B)  $e$  (E)  $\sqrt{e^{-1}}$
- (C)  $\sqrt{e^3}$
8. \_\_\_\_\_ Which of the following is equivalent to  $\sin^3 x \cos^2 x$ ?
- (A)  $\frac{1}{16}[2 \sin x - \sin 3x - 2 \sin 5x]$
- (B)  $\frac{1}{16}[\sin x - 2 \sin 3x - \sin 5x]$
- (C)  $\frac{1}{16}[2 \sin x - \sin 3x - \sin 5x]$
- (D)  $\frac{1}{16}[2 \sin x + \sin 3x - \sin 5x]$
- (E)  $\frac{1}{16}[\sin x + \sin 3x - \sin 5x]$

9. \_\_\_\_\_ The value of  $\int_0^4 |x - 2| dx$
- (A) 3 (D)  $\frac{3}{2}$   
 (B)  $\frac{5}{2}$  (E)  $\frac{7}{2}$   
 (C) 2
10. \_\_\_\_\_ The absolute maximum of  $f(x) = \cos 2x - 2 \cos x$  on  $[0, 2\pi]$  occurs at  $x =$
- (A)  $\frac{\pi}{3}$  (D)  $\frac{5\pi}{3}$   
 (B)  $\frac{\pi}{2}$  (E)  $\frac{3\pi}{4}$   
 (C)  $\pi$
11. \_\_\_\_\_ What is the maximum perimeter of all rectangles that can be inscribed in  $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$
- (A)  $4\sqrt{a^2 + b^2}$  (D)  $a^2 + b^2$   
 (B)  $\frac{8}{\sqrt{a^2 + b^2}}$  (E)  $2(a^2 + b^2)$   
 (C)  $2\sqrt{a^2 + b^2}$
12. \_\_\_\_\_ The solution set for the inequality  $x - \frac{3}{x} > 2$  is given by
- (A)  $(0, +\infty)$  (D)  $(-\infty, 0) \cup (3, +\infty)$   
 (B)  $(3, +\infty)$  (E)  $(-\infty, 3)$   
 (C)  $(-1, 0) \cup (3, +\infty)$
13. \_\_\_\_\_ If  $i = \sqrt{-1}$ , then  $\sum_{j=0}^{10} (-i)^j$  is
- (A)  $i$  (C)  $-i$  (E)  $1 - i$   
 (B)  $-1$  (D)  $1 + i$

14. \_\_\_\_\_ Let  $R$  be the region defined by  $y = x - 1$ ;  $x = 1$ ;  $y = -x + 3$

Find the maximum value of  $f(x, y) = -2x + 3y$  on  $R$

- (A)  $-2$  (D)  $4$   
 (B)  $1$  (E)  $-1$   
 (C)  $2$

15. \_\_\_\_\_ Let  $T$  be a linear transformation of the plane such that  $T(1,1) = (-1,1)$  and  $T(2,3) = (1,2)$ . Then  $T(2,4)$  equals

- (A)  $(4,2)$  (D)  $(2,4)$   
 (B)  $(2,-4)$  (E)  $(-3,2)$   
 (C)  $(3,-2)$

**Answer Key:**

- |      |       |       |
|------|-------|-------|
| 1. B | 6. C  | 11. A |
| 2. A | 7. C  | 12. C |
| 3. E | 8. D  | 13. C |
| 4. B | 9. B  | 14. D |
| 5. A | 10. C | 15. A |

## Appendix B: Open-Ended Action Identification Measure

**What do you do in taking a math exam?**

Do not worry about grammar or spelling; just write whatever comes to mind. For example, if asked what you are doing when brushing your teeth, you might write “caring for my dental hygiene,” or “moving the toothbrush”. Please write as many answers as you can think of, but do not worry about filling all the spaces. Then, using the scale below, rate how well you think each answer applies to the exam you just completed.

	1	2	3	4	5	6	7
	not at all applies		somewhat applies			applies very well	
1.							Rating
2.							Rating
3.							Rating
4.							Rating
5.							Rating
6.							Rating
7.							Rating
8.							Rating
9.							Rating
10.							Rating
11.							Rating
12.							Rating
13.							Rating
14.							Rating
15.							Rating



## Appendix C: Closed-Ended Action Identification Measure

Using the scale provided, please respond to the statements below according to how well each one applies to **the exam you just completed**.

1

2

3

4

5

6

7

Describes very poorly

Describes very well

- H 1. \_\_\_\_\_ Demonstrating my intelligence  
L 2. \_\_\_\_\_ Sitting in a classroom  
H 3. \_\_\_\_\_ Learning something about math  
L 4. \_\_\_\_\_ Answering multiple-choice questions  
L 5. \_\_\_\_\_ Marking wrong answers  
L 6. \_\_\_\_\_ Trying to get comfortable in my chair  
H 7. \_\_\_\_\_ Competing with other students  
H 8. \_\_\_\_\_ Performing poorly  
H 9. \_\_\_\_\_ Showing what I know  
H 10. \_\_\_\_\_ Enjoying myself  
H 11. \_\_\_\_\_ Thinking about my answers  
L 12. \_\_\_\_\_ Making marks on my paper  
H 13. \_\_\_\_\_ Demonstrating my ability  
L 14. \_\_\_\_\_ Moving my hands  
L 15. \_\_\_\_\_ Reading questions  
H 16. \_\_\_\_\_ Spending time  
H 17. \_\_\_\_\_ Taking an exam  
L 18. \_\_\_\_\_ Failing to remember things  
H 19. \_\_\_\_\_ Being anxious  
L 20. \_\_\_\_\_ Circling correct answers  
L 21. \_\_\_\_\_ Guessing at some answers



Appendix D: Stereotype Activation Measure

- |             |               |                 |                 |
|-------------|---------------|-----------------|-----------------|
| 1. Unusual  | U N _ _ _ _ _ | 7. Incapable    | I N _ _ _ _ _   |
| 2. Minority | M _ _ _ _ _ Y | 8. Stupid       | S T _ _ _ _ _   |
| 3. Inferior | _ _ _ _ _ O R | 9. Timid        | T I _ _ _ _ _   |
| 4. Weak     | W E _ _ _     | 10. Discouraged | D _ _ _ _ _ E D |
| 5. Slow     | S _ _ _ W     | 11. Bad         | B A _ _ _       |
| 6. Hard     | H _ _ R _ _   | 12. Bias        | B _ _ _ S       |

Filler words:

Change

Plane

Quick

Space

Tower

Button

## Appendix E: Measures of Cognitive Distraction

Cognitive Interference Questionnaire (Sarason, 1980)

I. The following list includes some thoughts you might have had during the exam. Using the scale below, please indicate approximately how often each thought occurred to you while you were working on the exam.

1	2	3	4	5
never	once	a few times	often	very often

1. \_\_\_\_\_ I thought about how poorly I was doing.
2. \_\_\_\_\_ I wondered what the experimenter would think of me.
3. \_\_\_\_\_ I thought about how I should work more carefully.
4. \_\_\_\_\_ I thought about how much time I had left.
5. \_\_\_\_\_ I thought about how others have done on this task.
6. \_\_\_\_\_ I thought about the difficulty of the problems.
7. \_\_\_\_\_ I thought about my level of ability.
8. \_\_\_\_\_ I thought about the purpose of the experiment.
9. \_\_\_\_\_ I thought about how I would feel if I were told how I performed.
10. \_\_\_\_\_ I thought about how often I got confused.
11. \_\_\_\_\_ I thought about things completely unrelated to the experiment.

II. Please circle the number on the following scale which best represents the degree to which you felt your mind wandered *during the math test you have just completed*.

1	2	3	4	5	6	7
Not at all						very much

Evaluation Apprehension Scale (Spencer, Steele, & Quinn, 1999)

1                      2                      3                      4                      5  
strongly disagree                      agree                      strongly agree

1. \_\_\_\_\_ People will think I have less ability if I do not do well on this test.
2. \_\_\_\_\_ People will look down on me if I do not do well on this test.
3. \_\_\_\_\_ If I do not do well on this test, others may question my ability.
4. \_\_\_\_\_ If I do poorly on this test, people will look down on me.
5. \_\_\_\_\_ I feel self-confident.

Worry component of the Worry-Emotionality Scale (Morris, Davis, & Hutchings, 1981)

To the left of each of the following statements, indicate your feelings, attitudes, or thoughts as they are right now *in relation to the exam you just completed*.

1                      2                      3                      4                      5  
strongly disagree                      agree                      strongly agree

1. \_\_\_\_\_ I feel regretful.
2. \_\_\_\_\_ I am afraid that I do not know as much as I should.
3. \_\_\_\_\_ I feel others will be disappointed in me.
4. \_\_\_\_\_ I feel I may not do as well on this test as I could.
5. \_\_\_\_\_ I do not feel very confident about my performance on this test.

Frustration scale

How did the exam make you feel? Using the scale below, please rate the extent to which you felt each emotion.

1                      2                      3                      4                      5  
not at all                      somewhat                      extremely

1. \_\_\_\_\_ Agitated
2. \_\_\_\_\_ Irritated
3. \_\_\_\_\_ Confident
4. \_\_\_\_\_ Competent
5. \_\_\_\_\_ Frustrated
6. \_\_\_\_\_ Capable
7. \_\_\_\_\_ Aggravated
8. \_\_\_\_\_ Intelligent

Appendix F: Demographics and Manipulation Checks

Please rate the following statements.

1. I am good at math.  
 1 2 3 4 5 6 7 8 9 10 11  
 strongly disagree strongly agree

2. It is important to me that I am good at math.  
 1 2 3 4 5 6 7 8 9 10 11  
 strongly disagree strongly agree

Please provide the following information.

1. Gender  Male  Female
2. Class year  Sophomore  
 Junior  
 Senior
3. Age \_\_\_\_\_
4. Academic Major: \_\_\_\_\_
5. How many math courses have you taken while enrolled at Hamilton College? \_\_\_\_\_  
 Of these, how many were Calculus classes? \_\_\_\_\_
6. SAT Quantitative score: \_\_\_\_\_  
 If you can not remember your score, provide a probable range for your score \_\_\_\_\_  
 How sure are you that this was your actual score? 1 2 3 4 5  

not at all sure very sure
7. The math test you completed :  
 has been found to show gender differences  
 has been found to show NO gender differences