CHOKING UNDER PRESSURE AND WORKING MEMORY: INFLUENCE OF COGNITIVE APPRAISAL

By

YE WANG

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS IN EDUCATION

UNIVERSITY OF FLORIDA

2010

To Jie Zou

ACKNOWLEDGMENTS

I wish to thank several individuals for their help and support concerning my thesis study. First, I thank Dr. David Therriault, my advisor and the chair of my thesis committee, for providing me with generous support throughout the process. His constructive and prompt feedback on my study design, data analyses, and writing was greatly appreciated. I also thank the other member of my thesis committee, Dr. David Miller, for his service and helpful comments about my study. Last but not the least, I acknowledge Jie Zou, my husband, for supporting and encouraging me as well as helping me with the writing style. I thank my parents for their love and encouragement, and unconditional support for my interests and dreams. All these individuals have helped me achieve a goal that seemed like a very daunting challenge.

TABLE OF CONTENTS

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Arts in Education

CHOKING UNDER PRESSURE AND WORKING MEMORY:
INFLUENCE OF COGNITIVE APPRAISAL

By

Ye Wang

August 2010

Chair: David Therriault
Major: Educational Psychology

Eysenck (1992) proposed that anxiety could influence cognition by occupying cognitive processing resources, jeopardizing working memory capacity by pressure via increasing state anxiety. Recent findings (Beilock & Carr, 2001, 2005; Gimming, Huguet, Caverni, & Cury, 2006) have demonstrated that only individuals with high working memory (HWM) capacity "choke under pressure" on cognitive tasks with high working memory demands; while individuals with low working memory (LWM) capacity increased their performance under such conditions. Gimming and his colleagues suggested that LWM and HWM individuals interpreted stress differently, but research has not explored how these groups may be different. Our present study examined how LWM and HWM individuals interpreted stress differently by applying cognitive appraisal theory (Lazarus and Folkman, 1984). Our results suggested that HWM group experienced highest threat appraisal and thus choked under pressure.

CHAPTER 1
OVERVIEW

Working memory plays an essential role in everyday life; it helps us to keep track of where we are and what we are doing moment to moment. It aids us in holding information long enough to make a decision. Storing while simultaneously processing information is fundamental to many cognitive tasks such as reasoning, reading comprehension and learning (Baddeley, 2002). Working memory capacity (WMC) is also limited. Individuals with high working memory have greater resources at their disposal to control processing and maintain information. Controlled processing is required for gating, blocking, or suppressing distracting information in the face of distraction and interference in order to maintain temporary goals (Engle, Kane, & Tuholski, 1999).

There are many potential influences on working memory. For example, working memory is susceptible to interference and intrusions. One source of interference and intrusion comes from anxiety, or stress. Eysenck (1992) proposed that anxiety could influence cognition by occupying cognitive processing resources and WMC was jeopardized by pressure via increasing state anxiety (Beilock & Carr, 2001, 2005; Beilock, Tugade, & Engle, 2004). People exposed to stress showed significant working memory impairments (Schoofs, 2008). Ashcraft (2002) found math anxiety disrupted cognitive processing by compromising ongoing activity in working memory. Ashcraft and Kirk (2001) found individuals with high math anxiety demonstrated smaller working memory spans. They considered disruption of central executive processes as a possible mechanism and suggested individual emotional difference variables like math anxiety deserve greater attention, especially on assessments of working memory capacity and functioning.

One research approach to induce anxiety and stress in laboratory settings is by using stereotype threat. According to stereotype threat theory (Steele & Aronson, 1995), performing in

a negatively stereotyped group produces feelings of anxiety, uncertainty, and discomfort. Stereotype threat often impairs performance (Aronson, Lustina, Good, Keough, Steele, & Brown, 1999; Croizet & Claire, 1998; Steele & Aronson, 1995). Considerable research has indicated that cognitive and affective consequences of activating a negative self-relevant stereotype of race (Cheryan & Bodenhausen, 2000), social economic statues (Steele, 1997), or gender (Spencer, Steele, & Quinn, 1999; Steele, 1997), including decreased performance expectations (Cadinu, Maass, Frigerio, Impagliazzo, & Latinotti, 2003; Stangor, Carr, & Kiang, 1998), increased anxiety (Bosson, Kaymovitz, & Pinel, 2004; Osborne, 2001), and reduced working memory capacity (Beilock, Rydel, & McConnell, 2007; Schmader & Johns, 2003).

Stress can lead to "choking." "Choking under pressure" is defined as performing more poorly than expected given one's skill and occurs in situations where the pressure to perform is high (Beilock & Carr, 2001). Having examined the relationship between choking under pressure and individual differences in working memory capacity, some authors concluded that the individuals who were most likely to fail under performance pressure were those who, in the absence of pressure, had the highest potential for success (i.e., high working memory individuals). Beilock and Carr (2005) found that only high working memory (HWM) individuals were harmed by performance pressure in math problems. Gimmig, Huguet, Caverni and Cury (2006) also found that individuals with high working memory were more likely than their low-working-memory (LWM) counterparts to choke under pressure on working memory and fluid intelligence tests. In Lai's (2008) results, individuals low in working memory capacity (when given negative feedback about their performance) actually improved on working memory tasks. In contrast, HWM individuals who were given the same feedback demonstrated choking.

Several potential reasons for choking under pressure have been proposed. Anxiety is the one that has been most explored (Beilock & Carr, 2001; Beilock & Carr, 2005; Gimming et al., 2006). Beilock and Carr (2001) explained that heightened anxiety induced self-focused attention, thus producing intrusive thoughts. Beilock and Carr (2005) explained that performance pressure increased reported state anxiety. Intrusive thoughts, such as worries, reduced the storage and processing capacity of working memory, anxiety-providing situations reduced attentional resources, resulting in performance decrements. They explained that HWM individuals' usual working memory advantage might make them susceptible to failure when pressure was added, if pressure-induced consumption of working memory denied them the capacity they normally relied on to produce their superior performance. In line with Beilock and Carr (2005), Gimming and his colleagues (2006) found anxiety mediated choking among HWM individuals, and choking under pressure was due to increased state anxiety for HWM individuals. Interestingly, they found under same pressure situation as HWM individuals' experience, LWM individuals reported decreased anxiety and increased performance. They suggested that LWM and HWM individuals might interpret high-stress situations differently. Anxiety increased in HWM individuals under pressure, but decreased in LWM individuals. They explained that only HWM individuals choked because of their anxious perception of high-stakes situations at the onset.

In sum, reactions to pressure appear to depend on WMC. It appears that choking only occurs in HWM individuals. Previous studies are less clear about the specific mechanism driving these results. According to Eysenck (1992), anxiety occupies cognitive processing resources. The question remains, why is this effect only occur for HWM individuals. In fact, LWM individuals demonstrated decreased anxiety (Gimming, et al, 2006) and even increased performance (Lai, 2008) under high pressure situation. Gimming and his colleagues (2006) suggested that the

pressure also affected LWM individuals, but their performance increased in pressure situation. If anxiety compromises working memory capacity, why doesn't the same pattern of choking occur for LWM individuals? The purpose of our study was to further explore the relationship of anxiety and working memory. Specifically, we examined the link between working memory and stress by employing cognitive appraisal theory.

In cognitive appraisal theory, stress can have differential effects. The effect of stress can be either positive or negative. Negative effect of stress causes anxiety and decreases performance, referred as threat appraisal. Positive effect of stress increases performance on a task causing what is referred as challenge appraisal (Lazarus & Folkman, 1984). For example, Vick, Seery, Blascovich and Weisbuch (2008) found an effect of gender stereotype activation related to challenge or threat cognitive appraisal states. Girls, who were told that experiment was gender biased, would feel gender stereotype threat; therefore they exhibited the threat state and underperformed. If girls were told that experiment was gender fair, they would not feel stereotype threat; instead, they exhibited challenge appraisal and perform better. In contrast, if the experimenter told boys that the experiment was gender fair, they showed threat and underperformed. If boys were told that experiment was gender biased (i.e., females would underperform), they would show the challenge state and increase performance. Thus, studies regarding cognitive appraisal theory suggested that interpretation of stress could either increase or decrease performance, depending on the appraisal of the stress.

In cognitive appraisal theory, Lazarus and Folkman (1984) proposed two major components of the processes: primary appraisal and secondary appraisal. Primary appraisal reflects perceptions of the nature and degree of risk in the situation. Secondary appraisal reflects perceptions of resources or abilities to cope with the situation. According to appraisal theories,

before performing a task, the task is appraised on two criteria: “how difficult the task will be” and “how well I could cope with the task demands.” Appraisals can then be classified as either “threatening” or “challenging.” Threat appraisals are those in which the perception of danger exceeds the perception of abilities or resources to cope with the stressor, and result in underperformance. Challenge appraisals, in contrast, are those in which the perception of danger does not exceed the abilities to cope, and possibly increase performance. In a serial subtraction tasks (i.e., the participant was asked to repeatedly subtract a small number, like 7 or 13, from a running total that starts as a four-digit number, such as 5,342), threatening appraisals leaded to performances with more mistakes and about a 25% slower rate than the challenging appraisal conditions (Tomaka, Blascovich, Kelsey, & Leitten, 1993).

Threat appraisals are more strongly associated with negative emotional reactions, high rates of cognitive interference, and poor performance levels than challenge appraisals are. Threatened individuals may be less task-focused and have intrusive cognitions, whereas challenged individuals may be less distracted by negative emotions and more task-focused. Folkman and his colleagues (1986) also found that participants who appraised self-reported stressful events as changeable (high secondary appraisal of coping efficacy) used task-focused coping strategies (e.g., planed problem solving), and were more motivated and more satisfied with encounter outcomes than participants who appraised the events as unchangeable. These participants (low secondary appraisal of coping efficacy) used more emotion-focused coping strategies (e.g., distancing and escape-avoidance). These emotional coping strategies left participants less motivated to finish the task, and more likely to try to find self-handicapping excuses. Skinner and Brewer (2002) demonstrated that challenge appraisals were associated with more confident coping expectancies, lower perceptions of threat, higher positive emotion, and

more beneficial perceptions of the effects of appraisal and emotion on performance. Generally, higher performance was associated with challenge appraisals. Gaab, Rohleder, Nater, and Ehlert (2005) linked appraisal with psychobiological stress change and found that appraisal helped to predict performance. Vick and his colleagues (2008), as we reviewed above, found an effect of gender stereotype activation due to challenge/threat states. Females in the gender-biased condition exhibited more threatened state, whereas women in the gender-fair condition exhibited more challenge state than they did in the gender-biased condition. Thus, challenge and threat states can be activated by anxiety of negative information of oneself and change performance.

Threatened individuals may be less task-focused, and more self-emotion focused, whereas challenged individuals may be less distracted by negative emotions, and more task-focused. As reviewed above, in working memory research, Beilock and Carr (2001) found self-consciousness training, which trained participants not to focus on thoughts about self but focus on task, eliminated choking. Intrusive thoughts, which are less task-focused, can cause choking (Beilock and Carr, 2005). Heightened anxiety induces self-focused attention and causes choking. Given our review of the literature, it was reasonable to hypothesize that choking is related to cognitive appraisal. Cognitive appraisal might help to explain why choking works differently in high or low working memory individuals. It is possible that HWM people are more likely to demonstrate threat appraisal and choke under pressure; whereas LWM people are more likely to demonstrate challenge appraisal and avoid choking.

Furthermore, because individuals with threat appraisal lose the confidence to cope with the problems they will be more likely to try to employ excuses to avoid the stress, resulting in high self-handicapping scores. In contrast, for challenge appraisal individuals, they are likely to be more motivated and demonstrate low self-handicapping scores under stressful situation.

Gimming et al. (2006) found differences in the self-handicapping measurements for HWMs and LWMs. Choking was not only related to state anxiety, but also self-handicapping. Self-handicapping was used to describe the tendency of trying not to focus on task in order to use this as an excuse to avoid pressure. Gimming et al. (2006) reported that LWM people had less anxiety score on state anxiety scale, and low self-handicapping score on Rhodewalt's scale (1990), explaining why LWMs did not choke under pressure. In contrast, HWM individuals displayed a tendency to excuse their poor performance under pressure (they scored higher in self-handicapping than LWM individuals).

Thompson (2007) found that poor performance feedback could produce self-handicapping behavior in people with high ability, which is analogous to the HWM individuals in our study. In our study, we predicted that for LWMs, challenge appraisal would be linked to lower self-handicapping scores; and for HWMs, threat appraisal would be linked to higher scores of self-handicapping. Therefore, we incorporated in our study a self-handicapping scale (Rhodewalt's scale, 1990) at the end of second task and Gimming et al. used the same scale in their study.

Previous studies focused on reasons why HWM individuals' choke. Relatively few studies explored why LWM individuals perform better and seem to be under less anxiety when placed in high-pressure situations. In Lai's (2008) study, she argued that LWM individuals did not use full cognitive resources under low stressful task, saving cognitive resources for more stressful tasks. In our study, we would also explore this hypothesis by giving a self-report effort scale at the end of each task. On the other hand, this self-report effort scale could also give us important information about whether HWM individuals' self-handicapping behavior would cause a withdrawal of effort under high pressure.

The current study extends previous research by examining a potential cognitive appraisal consequence in HWMs and LWMs. The current study was a 2 (WMC: HWMs vs. LWMs) x 2 (stress: treatment vs. control) design. Our study adopted Lai's (2008) approach of using feedback to induce stress. Participants were categorized into high, middle or low working memory span (while dropping the middle span group in the paired comparison analyses because it was not of interest to our hypotheses). Participants were administered a working memory test, then given feedback on how well they did by assigning them randomly into one of two groups/conditions: 40th percentile, or no-feedback condition. As noted by Beilock and Carr (2005), pressure that caused people to choke included the consequences of suboptimal performance, especially on examinations, including poor evaluations by mentors, teachers and peers. The 40th percentile condition was expected to create stress for the second task (i.e., the participants were informed that they received a failing score of 60%). Then participants then completed a second working memory task. Participants in the no-feedback condition continued on to the second test without any explicit feedback.

We predicted that people with HWM would be more likely to interpret feedback as threat appraisal and choke under pressure while LWM people would not. Our specific hypotheses regarding individuals in the 40th percentile feedback treatment condition were as follows: (1) challenging participants' views of their own competency would result in choking, especially in HWM individuals, not in LWM ones; (2) HWM people would demonstrate more state anxiety than LWM people; This hypothesis could potentially provide confirmatory evidence for Beilock and Carr's (2005) and Gimming et al.'s (2006) speculations that choking is mediated by anxiety; (3) HWM individuals would show more threat appraisal in the second task than in the first task; (4) HWMs would feel more stressful, put less effort and express high self-handicapping in the

second task than in the first task; (5) LWMs would show more challenge appraisal in the second task than in the first task; (6) LWMs would feel less stressful, put more effort and express lower self-handicapping in the second task than in the first task. In no feedback condition, we hypothesized: (7) the control groups would have no significant changes in state anxiety, cognitive appraisal, RSPAN performance, stress, effort in first and second task.

CHAPTER 2
METHODS

Sample

There were 157 participants in this study. They were undergraduate educational psychology students from the University of Florida who participated for credits toward a course requirement.

Materials

Bell Curve Survey

Upon signing the informed consent, participants were administered a Bell Curve Survey asking them to indicate where they believed their academic competency (percentile) fell on the curve, comparing to their peers. Because in our experiment, participants were informed that this experiment correlated well with academic competency, the survey served to set the stage for the manipulation of giving negative feedback treatment by highlighting social comparisons between individuals.

Working Memory Test

We employed a dual-processing test called the reading-span task (RSPAN) to measure working memory capacity (Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005). The idea is that high working memory capacity individuals do better job in dual tasks because they have more resources to maintain and keep track of multiple sources of information. Daneman and Carpenter (1980) originally developed the RSPAN task to assess working memory during reading. In our study, participants evaluated sentences while memorizing letters for later recall so it required dual processing. Participants were asked to read a sentence presented on the computer screen, and decided whether or not that sentence made sense in a real world scenario by clicking either “yes” or “no” button on the screen. A letter that required participants to recall later was

presented for 2 seconds after each sentence decision and a blank screen lasting 1 second separated the presentation of each sentence and letter. At the end of a series of sentence/letter combination trials, participants were asked to recall, as best as they could, the letters in the order presented. Each set was separated by a screen said “next set”, which was displayed for 3 seconds. The computer recorded the number of letters recalled in the correct order, and participants’ correct or incorrect sentences decision. One set ranged from three to six sentences, so at the end of a set participants were asked to recall the letters he/she could recall in the set, and then came to the next set.

In our study, the RSPAN task included 15 sets of sentence/letter trials that contained three to six sentence/letter combinations. The sets were presented in random order so participants could not anticipate how many sentences and letters they were required to evaluate and recall at the beginning of each set. This was a dual task because participants made judgments whether the sentence made sense, and at the same time remembered the letters presented after each sentence. Working memory capacity was scored by the total number of letters that participants recalled correctly from each set while maintaining at least an 85% accuracy rate on judging the sensibility of sentences.

The second RSPAN task mirrored the first. Both tasks had equal number of sets and equal number of nonsense and sense sentences. The letters to be recalled were also taken from the same sampling of letters as the first RSPAN task. The sentences from the second RSPAN were of equal length to the first task and same number of words per sentence. And the sentences are also of similar difficulty levels using the Flesch-Kincaid Grade Level readability scale.

We chose RSPAN task because it was stressful and difficult, as research reports had indicated (Conway et al., 2005; Lai, 2008). Stress is most likely to occur when the task is

sufficiently difficult to frustrate participants' abilities (Steele, Spencer, & Aronson, 2002). Therefore, this task is likely to induce threat and cognitive appraisal of stress.

Self-Handicapping Scale

Participants were administered a self-handicapping scale after completing the second task. They rated statements about how handicapped they had felt during the task performance on a 5-point scale ranging from 0 (totally disagree) to 5 (totally agree) (e.g., "I slept well last night"; adapted from Rhodewalt, 1990). This served to index self-handicapping (a=0.69).

Demographic Questionnaire

Participants were administered a brief Demographic Questionnaire (Appendix A) at the end of the study. This was included to insure that there was no accidental priming of a gender or race stereotype threat that could affect performance. The questionnaire asked questions regarding age, gender, and ethnicity. The questionnaire was included as an additional control measure.

State Appraisal Measure

Skinner and Brewer's (2002) reports provided an example of this scale used. Four threat appraisal items addressed concerns about personal evaluation ("I'm concerned that others will be disappointed in my performance") and performance ("I worry that I may not be able to achieve the grade I am aiming for"; "I'm concerned about my ability to perform under pressure") and the tendency to focus on negative outcomes ("I am thinking about the consequences of performing poorly"); The other four challenge appraisal items addressed participants' enthusiasm toward testing skills and abilities in demanding situations (" I am looking forward to testing my knowledge, skills, and abilities") and the anticipation of positive benefits (" I am focused on the positive benefits I will obtain from this situation"; " I am looking forward to the rewards of success"; "I am thinking about the consequences of performing well"). Intensity and frequency of state appraisals were measured on separate 6-point response scales (1 = strongly disagree, 6 =

strongly agree and 1 = this thought hardly ever occurred, 6 = this thought occurred almost constantly).

Self-Reported Appraisal and Coping Expectancies

This scale was similar to state appraisal scale because they both measure cognitive appraisals. However, the self-reported appraisal scale asked about "the current state" (i.e., our manipulation) because it asked participants about the upcoming task. In contrast, the state appraisal scale is a more global assessment of general state rather than participants' feelings about the specific task. Tomaka et al. (1993) assessed primary and second appraisal by asking subjects, "How stressful do you expect the upcoming task to be?" Secondary appraisal was assessed by asking: "How able are you to cope with this task?" before the task; and assessed subjective experience of stress by asking, "How stressful was the task you just completed?" after the task. Skinner and Brewer (2002) also measured coping expectancies by self-reported anticipated coping ability and expected level of performance. Coping expectancies were assessed by two 10-point scales tapping anticipated coping ability (1 = very confident can cope effectively, 10 = very concerned whether can cope effectively) and the expected level of performance (1 = poor, 10 = high quality). The second coping item was reverse coded. Kelsey et al. (2000) also asked five questions immediately before performing task, which were (a) "How threatening do you think the upcoming task will be?" (b) "How demanding do you think the upcoming task will be?" (c) "How stressful do you think the upcoming task will be?" (d) "How able are you to cope with this task?" and (e) "How well do you think you will perform this task?" Obviously (a) was related to threat appraisal; (b) was related to challenge appraisal; (c) was the primary appraisal; (d) and (e) were like what Skinner and Brewer used as secondary appraisal coping expectancy.

We adopted the five questions from Kelsey's (2002) report and we used 10-point scales for each question. These appraisal items were presented to subjects through computer after they had read the instructions of the task. Finally we asked participants "How stressful was the task you just completed?" after they had completed each of the tasks.

State Anxiety Scale

Following each task, participants completed an anxiety scale that has several items related to perceptions of the test and testing situation. Anxiety during the tests was measured using the questions adapted from the Spielberger State Anxiety Scale (1970) (see Schmader & Johns, 2003). Participants rated on a 7-point scale (ranging from 1 not at all to 7 very much) how much they felt anxious, comfortable, jittery, worried, at ease, nervous, relaxed, and calm while taking the test. The items on the scale were summed (after reverse scoring comfortable, at ease, relaxed, and calm) to form an index of state anxiety, where higher numbers indicated more anxiety (a= 0.90).

Procedure

After seating a participant by a computer in a quiet laboratory room, a research assistant first asked the participant to read and sign an informed consent form. The research assistant then presented the Bell Curve Survey in which participants indicated their academic competency in percentile compared to their peers on the bell curve. Participants were told that the experiment consisted two tests and took about 1 hour long. The research assistant highlighted the importance and credibility of the test by saying that the test was created by an authority and was related to the academic performance. The research assistant also described working memory capacity so that the participants were made fully aware of the task. The detailed script was included in Appendix C. The first RSPAN test instruction was then presented on the computer. Immediately after viewing and fully understanding the instruction, participants were measured on the state

appraisal measure and self-report appraisal and coping expectancies. Then the first RSPAN test was administered.

Following the first test, participants were required to answer, “How stressful was the task you just completed?” “How much effort did you put in?” (both questions were 10-point scale), and the state anxiety scale. After completing the questions, the research assistant told them their “performance” and indicated the percentage in the Bell Curve. Participants were randomly assigned into one of two conditions/groups, which was no feedback, or negative (stress) feedback. We chose to use percentile ranks as it highlighted the participant’s own performance compared to others. Even though our participants should have had also familiar with this type of reporting which they experienced in standardized tests, like the SAT, the research assistant still explained the percentile rankings and informed participants how their scores were compared to others. Participants in the 40th percentile were told that based on their rank they were expected to have difficulty on the subsequent test. The 40th percentile was used because it was considered below average so as to create stress, but it was still a realistic achievement group. In the no-feedback condition, participants completed the questionnaires and questions but continued onto the second test without any score or expectation feedback. Using the no-feedback condition as the control group, we would learn whether the two different RSPAN tasks were of equal difficulty level.

The second RSPAN test was then administered. After the computer displayed the instructions, participants were assessed on the same self-report appraisal and state appraisal scales. Then the second RSPAN test started. After finishing the second RSPAN test, participants were assessed on the same questions and state anxiety scales as they did in the first RSPAN test. Upon completing above tests and scales for all participants, the experimenter let participants fill

out the self-handicapping scale, and a questionnaire asking how they felt after they received the feedback, and how they perceived their performance, expectancy, effort and motivation changed in the second task (Appendix B). Participants were then asked to fill out the demographic questionnaire.

Finally, participants in the treatment condition were debriefed by informing participants that their percentile rankings feedbacks for the first test were fictitious and their percentile group was assigned at random.

CHAPTER 3
RESULTS

Reading Span Task

Participants who made errors more than 15% in deciding whether a sentence made sense were excluded from the analyses. We used an existing standard of 85% accuracy criterion for all participants (Conway et al., 2005). Eleven participants were excluded using this standard. Additional participants were excluded because they missed one page of the state appraisal scale or because their native language was not English (i.e., the RSPAN task requires the participants to be a native English speaker). In sum, we excluded 13 participants' data from the 157 participants. This resulted in removing 8% of the data.

We used working memory total scores in the analyses. Total scores were calculated by summing the total number of letters recalled in the correct position (First RSPAN total scores: $M = 54.60$, $SD = 10.78$; Second RSPAN total scores: $M = 54.92$, $SD = 11.86$). The correlation between these two tasks was 0.728 ($p < 0.01$), demonstrating parallel-forms reliability between the two RSPAN tasks.

Participants were split into three working memory capacity (WMC) groups based on the distribution of the first RSPAN scores (LWM group's scores ranged from 26 to 50, $M = 42.48$, $SD = 6.35$, $n = 50$; Medium WM group's scores ranged from 51 to 60, $M = 55.91$, $SD = 2.71$, $n = 46$; HWM group's scores ranged from 61 to 73, $M = 65.98$, $SD = 3.85$, $n = 48$). This classification resulted in approximately equal numbers of individuals with low, medium, and high spans in the working memory capacity. Only LWMs and HWMs were included in the analysis of variance (ANOVA). Participants who were categorized into middle WM group were dropped in the analyses because they were not of theoretical interest to our specific hypotheses.

We conducted a 2 (stress: treatment vs. control) x 2 (WMC: HWM vs. LWM) x 2 (RSPAN: first vs. second) repeated measures analysis of variance (ANOVA) with the experimental condition of stress (40th negative feedback treatment group and no feedback control group) and WMC (high and low) as between-subjects factors and the repeated measures on the RSPAN scores (first and second) as a within-subjects factor. The three-way interaction was not significant [$F(1, 94) = 0.489$, n.s.]. A 2 (WMC: HWMs vs. LWMs) x 2 (RSPAN: first vs. second) interaction was significant [$F(1, 94) = 10.83$, $p < 0.01$]. A similar ANOVA with the treatment groups revealed a significant WMC x RSPAN interaction [$F(1, 42) = 8.549$, $p < 0.01$]. For the control groups we did not obtain a significant WMC x RSPAN interaction effect. We could conclude there was a significant difference on the first and the second RSPAN scores between HWM and LWM individuals under the treatment condition, but not under the control condition. HWM and LWM in the treatment groups performed differently on the first and the second tasks, whereas the control groups did not.

Planned comparisons derived from our hypotheses showed that for the LWM treatment group, the RSPAN scores increased 2.90 points out of 75 ($SE = 1.39$) on the second task compared to the first task [$t(21) = -2.09$, $p < 0.05$]; for the HWM treatment group, the RSPAN scores decreased 3.59 points out of 75 ($SE = 1.74$) [$t(21) = 2.07$, $p < 0.05$ (one-tailed)]. This confirmed our hypothesis that HWMs choked in the treatment condition, while LWMs did not. In addition, we found that for LWM control group, the RSPAN scores increased 3.10 points out of 75 ($SE = 1.52$) on the second task compared to the first task [$t(27) = -2.05$, $p = 0.050$], which was quite close to significant. This surprising result showed that LWM in both stress conditions had a trend to increase their performance, which suggested a statistical regression or a practice

effect existed. In this case, we could not conclude that the increasing performance for LWM treatment group was due to our manipulation.

State Appraisal Score

The state appraisal score was calculated by summing the scores for each item in the state appraisal scales after reversing the challenging appraisal items (first state appraisal scores: M = 22.38, SD = 4.64; second state appraisal scores: M = 23.92, SD = 6.33). The higher the state appraisal scores, the higher the threat appraisal.

We conducted a 2 (stress: treatment vs. control) x 2 (WMC: HWM vs. LWM) x 2 (state appraisal scores: first vs. second) repeated measures ANOVA, with the state appraisal scores (first and second) as a within-subjects factor. The three-way interaction was not significant [$F(1, 94) = 1.673$, n.s.]. A 2 (stress: treatment vs. control) x 2 (state appraisal scores: first vs. second) interaction was significant [$F(1, 94) = 7.88$, $p < 0.01$]. A similar ANOVA with the HWM groups revealed a significant stress x state appraisal interaction [$F(1, 46) = 7.136$, $p < 0.05$]. For the LWM groups we did not obtain a significant stress x state appraisal interaction effect; there was also no evidence of any main effects. We could conclude that among the individuals with HWM, there was a significant difference on the first and the second state appraisal scores between the treatment and the control conditions, but not among the individuals with LWM.

Planned comparisons derived from our hypotheses showed that for the HWM treatment group, the state appraisal scores increased 2.86 points out of 48 (SE = 0.98) on the second task comparing to the first task [$t(21) = -2.932$, $p < 0.01$]; for the HWM control group, no difference was observed [$t(21) = 0.738$, n.s.]. This confirmed our hypothesis that in the treatment condition HWMs demonstrated more threat appraisal from first task to the second task.

Self-Reported Appraisal Scale

Following the procedure that Kelsey, Blascovich, Leitten, Schneider, Tomaka, and Wiens (2000) employed, we summed the scores for the five self-report appraisal items after reversing some items as needed (first self-reported appraisal scale: $M = 19.30$, $SD = 6.55$; second self-reported appraisal scale: $M = 26.78$, $SD = 7.21$). The higher the self-reported appraisal score, the higher the threat appraisal.

We conducted a 2 (stress: treatment vs. control) x 2 (WMC: HWM vs. LWM) x 2 (self-reported appraisal scores: first vs. second) repeated measures ANOVA, with the self-reported appraisal scores (first and second) as a within-subjects factor. The three-way interaction was not significant [$F(1, 94) = 0.656$, n.s.]. A 2 (stress: treatment vs. control) x 2 (state appraisal scores: first vs. second) interaction was significant [$F(1, 94) = 8.35$, $p < 0.01$]. A similar ANOVA with the HWM groups revealed a significant stress x state appraisal interaction [$F(1, 46) = 6.75$, $p < 0.05$]. For the LWM groups we did not obtain a significant stress x state appraisal interaction effect; or evidence of any main effects. We could conclude that among the individuals with HWM, there was a significant difference on the first and the second self-reported appraisal scores between the treatment and the control condition, but not among the individuals with LWM.

Planned comparisons derived from our hypotheses showed that for the HWM treatment group, the self-reported appraisal scores increased 8.36 points out of 50 ($SE = 1.58$) on the second task compared to the first [$t(21) = -5.28$, $p < 0.01$]. This confirms our hypothesis that in the treatment condition, HWMs increased in threat appraisal from the first task to the second task. For the HWM control group, a similar difference of 3.38 points ($SE = 1.15$) was also observed [$t(21) = -2.953$, $p < 0.01$]. This was not consistent with our hypothesis, that there would

be no differences in the control groups. However, we could see that HWMs increased much more under the treatment condition than under the control condition (8.36 comparing to 3.38).

State Anxiety Scale

This scale was calculated by summing the scores for the eight items after reversing some items as needed (first state anxiety scale: M = 33.62, SD = 8.31; second state anxiety scale: M = 33.80, SD = 9.36). The higher the score, the higher the state anxiety. We did not get any significant interactions or main effects in the ANOVA analyses (stress x WMC x two repeated measures on state anxiety). Therefore we could conclude that there were no changes in the anxiety for either group. This result was not consistent with our hypotheses and findings from Beilock and Carr's (2005) and Gimming et al.'s (2006) that HWMs became more anxious whereas LWMs became less anxious under stress.

Stress Score

The self-reported ratings on the question "How stressful was the task you just completed?" were used to calculate the stress scores (first stress scores: M = 5.91, SD = 2.04; second stress scores: M = 6.01, SD = 2.35). We did not get any significant interactions or main effects in the ANOVA analyses (stress x WMC x two repeated measures on stress). Therefore we could conclude that there were no changes in the stress scores for either group.

Effort Score

The self-reported ratings on the question "How much effort did you put in?" were used to calculate the effort scores (first effort scores: M = 8.65, SD = 1.34; second effort scores: M = 8.24, SD = 1.77). We did not get any significant interactions in the ANOVA analyses (stress x WMC x two repeated measures on effort). Therefore we could conclude that there were no significant changes in the effort score for either group.

Self-Handicapping Scale

Following Rhodewalt's (1990) procedure, we summed the scores for the items after reversing some items as needed (*M* = 75.97, *SD* = 11.69). The lower the score, the lower the self-handicapping. The self-handicapping scale was only measured once so we conducted 2(stress: treatment vs. control) x 2(WMC: HWM vs. LWM) ANOVA. However, we did not get any significant interactions or main effects. Therefore we could conclude that there were no differences in the self-handicapping scores for either group.

Demographic Questionnaire

We were unable to analyze the effects for gender and ethnicity because of the unequal distribution in our sample (*n* for gender: male = 16, female = 128; *n* for ethnicity: Indian = 1, Asian = 7, Black = 19, Hispanic =22, White = 89, other = 6).

CHAPTER 4
DISCUSSION

The experiment presented here represents an important departure from traditional choking research. Customarily, research exploring choking focuses on the negative influence of stress, typically anxiety (Gimming et al., 2006). Results usually report that only high-powered people choke, (e.g., individuals with high working memory). Often these results are explained by high-powered people's "susceptibility to failure", which implies that performance pressure consumes working memory capacity they rely on to achieve superior performance (Beilock & Carr, 2005). Previous research has not provided a theory or mechanism to explain why LWMs were less affected by stress, or, in some cases actually increased in performance. The purpose of the present experiment was to begin documenting the beneficial as well as detrimental influence of stress by applying cognitive appraisal theory (Lazarus & Folkman, 1984).

In our experiment, we used social comparison to introduce pressure. We presented a high-pressure scenario to participants to test whether performance was affected by this challenge (as choking research indicates), and to explore the reason for performance change. In sum, we proposed that the difference in pressure (no feedback vs. 40th percentile failure) would either motivated or hindered people's performance as a function of their working memory capacities. We expected to find choking in HWMs when given the negative feedback; and LWMs would be motivated by the negative feedback. We also expected that choking in HWMs was due to increased anxiety, increased threat appraisal, decreased challenging appraisal, and increased self-handicapping. In contrast, LWMs' increased performance was due to decreased anxiety, decreased threat, decreased challenging appraisal, and decreased self-handicapping.

Specifically, we hypothesized that: (1) Negative feedback would result in choking in HWMs but not in LWMs. Indeed, we found that HWMs choked; in addition, we found LWMs'

performance increased on the second test. Only HWMs in the treatment condition choked, the control group of HWMs did not show a change in performance. Our results are consistent with Beilock and Carr's (2005) claim that under high-pressure performance situations, HWM participants showed decrements in performance. Our results also mirrored Lai's (2008) finding that while HWM participants did significantly worse after receiving negative feedback, LWM participants' performance actually improved. It was worth noting that in our experiment, LWMs in both stress groups had a trend to increase their performance. One possible reason of LWMs' increasing in scores was a result of statistical regression. That is, scores on the posttest regressed to the population mean. However, we did not find the statistical regression on HWMs because only the treatment group demonstrated choking. Alternatively, LWMs' increasing performance in both conditions might be a result of practice effects. Klein and Fiss (1999) found that scores on the working memory operation span task markedly increased from the first to the second administration of the test, indicating a practice effect. In our study, LWMs might have more room to improve. It is important to note that while LWM participants' performance increased after receiving the feedback, they still underperformed relative to HWM participants.

We also hypothesized: (2) HWMs would demonstrate more state anxiety than LWMs in the treatment condition; (4) HWMs would feel more stressful, put less effort and express high self-handicapping in the second task than in the first task in the treatment condition: We did not find any changes on these variables in our experiment. In Gimmig et al.'s (2006) study, they found that anxiety mediating choking. We obtained no evidence that WMC was jeopardized by pressure via increased anxiety. In Gimmig et al.'s (2006) study, they also found that HWMs had increased self-handicapping under the high pressure. We did not find this effect. Our inability to replicate Gimmig's results could be attributed to the fact that our stress manipulation was

different from theirs. Different stressors could result in different anxiety arousal levels. In Gimming's study, participants were told that the experiment was associated with academic success, which was defined as high-pressure condition. In our study, we used negative performance feedback to induce stress.

We hypothesized: (3) HWM individuals would show more threat appraisal in the second task than in the first task in the treatment condition. The state threat appraisal scores and the self-reported appraisal scores for HWM under the treatment condition increased, while we did not find LWMs increased in threat appraisal. Our review of the previous research suggests that HWMs experienced more pressure than their LWM counterparts (Beilock and Carr, 2005; Gimmig et al., 2006). This was reflected by highly increased threat appraisal in our study. However, we also found HWM control group increased the self-reported threat appraisal scores from the first to the second task but it was a smaller increase compared to the HWM treatment condition (3.38 points comparing to 8.36 points). Perhaps there is a threshold on the amount of threat appraisal necessary to decrease performance.

We hypothesized: (5) In the treatment condition, LWMs would show more challenge appraisal in the second task than in the first task; (6) In the treatment condition, LWMs would feel less stressful, put more effort and express lower self-handicapping in the second task than in the first task: We did not find these results in our experiment. LWMs remained the same from the first to the second task. This result is inconsistent with Lai's (2008) hypotheses that negative feedback might have acted as a motivator for our participants who had lower working memory scores. LWMs did not interpret the negative feedback as a motivator as reflected in our measures of how "challenged" they felt. We are unable to explain LWMs' increasing performance by the change in their mental state (i.e., appraisal, anxiety).

Lastly, we hypothesized: (7) In the no-feedback condition, the control groups would have no significant changes in state anxiety, cognitive appraisal, RSPAN performance, stress, effort in the first and the second task. This was confirmed for all the factors, except that the HWM control group increased the self-reported threat appraisal scores from the first to the second task. But it was a relatively small increase compared to the HWM treatment condition.

The self-reported statement questionnaire (Appendix B), which asked how participants felt after they received the feedback, provided us with open comments. For HWM treatment group, one theme that emerged was the comments that students were scared by the negative feedback as so the test was more difficult than they could handle. This is consistent with our results that HWM treatment group experienced the highest threat appraisal. It is worth noting that some LWMs in both conditions said that they felt the second test to be easier. Since the test order was randomized and the mean scores for the two tests were similar (first RSPAN: 54.60; second RSPAN: 54.92), it is statistically unlikely that one test was inherently more difficult than the other.

Results from this study broadened the literature on individual differences between span types. Previous studies usually reported that only HWMs choked. Gimming (2006) reported that HWMs were more likely to have intrusive thoughts, such as worries, reducing the storage and processing capacity of working memory, as well as attentional resources, resulting in performance decrements. In our study, HWMs in the treatment condition felt pressure to overcome failure thus drawing their attention away from the working memory task. This was reflected by a significant increase in threat appraisal from the first to the second task. Threat appraisals are those in which the perception of danger exceeds the perception of abilities or resources to cope with the stressor, thus decreasing performance (Lazarus & Folkman, 1984).

HWMs appraised the negative feedback as “threatening” and did not have confidence to overcome. External (i.e., feedback) and internal (i.e., feelings and thoughts) distracters could draw attentional resources away from a task (Barrett et al., 2004). HWMs’ performance decrement can be explained through the reallocation of attentional resources from task-relevant information to task-irrelevant worries. On the other hand, we obtained no evidence of choking for LWMs; probably due to the absence of performance pressure to overcome competency threat, since LWMs remained in a neutral state (no changes in anxiety or cognitive appraisal from the first to the second task).

Previous research has suggested that LWMs increase their performance under stressful situations (Lai, 2008), and also demonstrate lower anxiety comparing to HWMs (Gimming et al., 2006). However, we did not get this result because LWMs increased their performance regardless of treatment manipulation. We failed to relate the beneficial influence of stress by applying cognitive appraisal theory because we could not conclude that the increase of performance was due to the stress; and there were no changes in cognitive appraisal from the first to the second task. The increase of performance might be a result of practice effect or statistical regression. The qualitative data from the statement questionnaire (Appendix B) showed some LWMs felt the second test to be easier than the first, suggesting that the increasing performance could be due to practice effects.

While the findings from this study helped to move forward our understanding of working memory, a few limitations and shortcomings exist. For example, the sample size was small (less than 30) for each group. Therefore, some of our results did not reach significance (i.e., LWM control group’s increased performance from first to the second task). Moreover, because we used

test-retest, we were unable to separate the confounding effect of stress, practice effect and statistical regression. Other types of stressors should be considered in future studies as well.

The strength of our present study was to more precisely specify the reason for choking, by aligning previous work with cognitive appraisal theory. The study explores both the negative and (potentially) positive effects resulted from stress. Our findings could change the focus of choking research (i.e., simply anxiety) to a more complex understanding of the appraisal of stress. We argue that increased threat appraisal in high pressure could result in choking, and provides a new research direction in this area.

A great amount of effort has been expended examining the effects of feedback given to students in the education system. The use of negative feedback as stressors in our study helped contribute to this area. Our findings suggest that some students may interpret negative feedback counterproductively and the negative feedback might discourage them from achieving. Specifically, our results point out, assuming that HWM participants are generally high achievers, those students may interpret negative feedback as a threat, which gives rise to their choking in stressful situations. This might be helpful for educators and policy makers to determine how and when to highlight the students' performance in comparison to peers.

APPENDIX A
DEMOGRAPHIC QUESTIONNAIRE

What is your age? _____

What is your gender?

1. Female
2. Male

How would you classify yourself?

1. American Indian or Native Alaskan
2. Asian or Pacific Islander
3. Black (not of Hispanic origin)
4. Hispanic or Latino
5. White
6. Other: _____

APPENDIX B
STATEMENT QUESTIONNAIRE

During the second part of the task, do you perceive your performance change in the second task comparing to the first part? How do you expect your performance in the second part comparing to the first part? Do you change your effort and motivation in the second task?

Do you become lazy in the second task comparing to the first task? (This question is on the reverse side of the questionnaire)

APPENDIX C
SCRIPT

The experimenter says to participants, “We are currently collecting data to create a student normal of University of Florida and we want to know how this test correlates with academic and cognitive ability. The test is related to academic potential. It would be expected to serve as a part of standard test in college exam like SAT in future, please take it seriously. The test is a reliable measure of working memory capacity. Working memory capacity is the ability to hold different pieces of information simultaneously while trying to process one specific piece of information. Your score is based on how accurately you evaluate sentences and the number of letters you can recall in order correctly.”

LIST OF REFERENCES

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.

Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science, 11*, 182–185.

Ashcraft, M. H., & Kirk, E. P. (2001). The relationship among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130,* 224-237.

Baddeley, A. D. (1990). *Human Memory: Theory and Practice*. Hove: Lawrence Associates.

Baddeley, A. D. (2002). Is working memory still working? *European Psychologist, 7,* 85-97.

Barret, L. F., Tugade, M.M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130,* 533-573.

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Educational Psychology: General, 130,* 701-725.

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *American Psychological Society, 16,* 101-105.

Beilock, S. L., Rydel, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Jounral of Experimental Psychology: General, 136*, 256-276.

Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, andCognition, 33,* 983-998.

Bosson, J. K., Haymovitz, E. L., & Pinel, E. C. (2004). When saying and doing diverge: the effects of stereotype threat on self-reported versus non-verbal anxiety. *Journal of Experimental Social Psychology, 40,* 247-255.

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.

Cheryan, S. C., & Bodenhausen, G. V. (2000). When positive stereotypes threaten intellectual performance: The psychological hazards of "model minority" status. *American Psychological Society, 11,* 399-402.

Croizet, J., & 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.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. (2005). Working memory span tasks: A methodological review and users guide. *Psychonomic Bulletin and Review, 12(5),* 769-786.

Corpus, J., Ogle, C., & Love-Geiger, K. (2006). The effects of social-comparison versus mastery praise on children's intrinsic motivation. *Motivation and Emotion, 30(4),* 335-345.

Daneman, M., & Carpenter, P. A. (1980). Individual-differences in working memory and Reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450-466.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5(3),* 69-106.

Engle, R.W., Kane, M. J., & Tuholski, S.W. (1999). Individual differences in working memorycapacity and what they tell us aboutcontrolled attention, general fluid intelligence, and function of theprefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). New York: Cambridge University Press.

Eysenck, M. W. (1992). *Anxiety: The Cognitive Perspective.* Hove: Lawrence Erlbaum Associates.

Folkman, S., Lazarus, R. S., Dunkel-Schetter, C., Delongis, A., & Gruen, R. (1986). Dynamics of a stressful encounter: Cognitive appraisal, coping, and encounter outcomes. *Journal of Personality and Social Psychology, 50,* 992-1003.

Gaab, J., Rohleder, N., Nater, U. M., & Ehlert, U. (2005). Psychological determinants of thecortisol stress response: The role of anticipatory cognitive appraisal. *Psychoneuroendocrinology, 30,* 599-610.

Gimmig, D., Huguet, P., Caverni, J., & Cury, F. (2006). Choking under pressure and working memory capacity: When performance pressure reduces fluid intelligence. *Psychonomic Bulletin & Review, 13,* 1005-1010.

Kelsey, R. M., Blascovich, J., Leitten, C. L., Schneider, T. R., Tomaka, J., & Wiens, S. (2000). Cardiovascular reactivity and adaptation to recurrent psychological stress: The moderating effects of evaluative observation. *Psychophysiology, 37*, 748-756.

Klein, K., & Fiss, W. H. (1999). The reliability and stability of the Turner and Engle working memory task. *Behavior Research Methods, Instruments, & Computers, 31*, 429-432.

Lai., S. (2008). Competency threat: What are the effects on working memory? Unpublished master's thesis, University of Florida, Gainesville.

Lazarus, R. S., & Folkman, S. (1984). *Stress, Appraisal and Coping.* New York: Springer.

Osborne, J. W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology, 26*, 291–310.

Rhodewalt, F. (1990). Self-handicappers: Individual differences in the preference for anticipatory self-protective acts. In R. L. Higgins, C. R. Snyder, & S. Berglas (Eds.), *Self-handicapping: The paradox that isn't* (pp. 69-106). New York: Plenum.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reducesworking memory capacity. *Journal of Personality and Social Psychology, 85,* 440-452.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality & Social Psychology, 85*, 440-452.

Schoofs, D., Preub, D., & Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology, 33,* 643-653.

Skinner, N., & Brewer, N. (2002). The dynamics of threat and challenge appraisals prior tostressful achievement events. *Journal of Personality and Social Psychology, 83,* 678-692.

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.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press.

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., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. Zanna (Ed.). *Advances in experimental social psychology (Vol. 34).* New York: Academic Press.

Thompson, T., & Dinnel, D. L. (2007). Is self-worth protection best regarded as intentional self-handicapping behavior or an outcome of choking under pressure? *Educational Psychology, 27,* 509-531.

Tomaka, J., Blascovich, J., Kelsey, R. M., & Leitten, C. L. (1993). Subjective, physiological, and behavioral effects of threat and challenge appraisal. *Journal of Personality and Social Psychology, 65,* 248-260.

Vick. S. B., Seery, M. D., Blascovich, J., & Weisbuch, M. (2008). The effect of gender stereotype activation on challenge and threat motivational states. *Journal of Experimental Social Psychology, 44,* 624-630.

BIOGRAPHICAL SKETCH

Ye Wang was born and grew up in Shanghai, China. She earned her Bachelor of Arts in Philosophy from Peking University in China in 2007. In fall 2007 she began her graduate study and majored in educational psychology at the University of Florida in Gainesville, Florida.