NOTE TO USERS

This reproduction is the best copy available.
RICE UNIVERSITY

Emotion Perception and Reactions to Tests: Affective Influences on Test Performance

by

Pablo Cruz

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Daniel Beal, Assistant Professor, Chair Psychology

Michelle Hebl, Associate Professor Psychology

Margaret Beier, Assistant Professor Psychology

Andrew Perkins, Assistant Professor Management

HOUSTON, TEXAS

MAY 2010
ABSTRACT

Emotion Perception and Reactions to Tests: Affective Influences on Test Performance

by

Pablo Cruz

The two studies presented here examine the impact of non-cognitive factors on test reactions and test performance. It is well-established that test anxiety can adversely affect test outcomes, particularly in high-stakes testing situations, but what is less well understood are the moderators that can influence that impact. These studies examined the test-takers’ distal appraisals of a testing situation, as well as their ability to process emotional information, and the effects of these factors on proximal test reactions and test performance. It was found that in these studies the positive role of test-taker confidence was more pronounced than the negative influence of test-taker anxiety, and that the ability of emotional perceptiveness had the effect of focusing attention on appropriate signals in the environment to allow the test-taker to have more accurate assessments of their own performance and to discourage the test-taker from using information from other items to influence their reactions to the items they are currently dealing with.
Acknowledgments

I am extremely appreciative of the patience and understanding of my committee members, Mikki Hebl, Margaret Beier, and Andy Perkins, who helped me to finally bring this project to completion. I am especially grateful to my advisor and committee chair, Dan Beal, for his insight and diligence in helping me to push forward and approach each aspect of the research with clear and thoughtful analysis. This paper has been greatly enhanced by their contributions. I would also like to thank my many supportive friends and co-workers, in particular Kelly Goff, Jessica Feinstein, Duc Pham, and Jennifer Rodriguez – their unwavering confidence and encouragement made all the hard work and long hours a lot easier to bear. Finally, I would like to dedicate this work to my mother, who was there with me every step of the way.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Test Anxiety</td>
<td>2</td>
</tr>
<tr>
<td>Emotional Intelligence</td>
<td>8</td>
</tr>
<tr>
<td>Study 1</td>
<td>18</td>
</tr>
<tr>
<td>Overview</td>
<td>21</td>
</tr>
<tr>
<td>Method</td>
<td>22</td>
</tr>
<tr>
<td>Results</td>
<td>24</td>
</tr>
<tr>
<td>Discussion</td>
<td>29</td>
</tr>
<tr>
<td>Study 2</td>
<td>35</td>
</tr>
<tr>
<td>Method</td>
<td>40</td>
</tr>
<tr>
<td>Results</td>
<td>46</td>
</tr>
<tr>
<td>Discussion</td>
<td>55</td>
</tr>
<tr>
<td>General Discussion</td>
<td>57</td>
</tr>
<tr>
<td>References</td>
<td>68</td>
</tr>
<tr>
<td>Appendices</td>
<td>86</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------</td>
</tr>
<tr>
<td>Table 1</td>
<td>Study 1 measurement properties</td>
</tr>
<tr>
<td>Table 2</td>
<td>Self-reported academic backgrounds of study 2 participants</td>
</tr>
<tr>
<td>Table 3</td>
<td>Study 2 reliabilities and intercorrelations between MSCEIT tasks</td>
</tr>
<tr>
<td>Table 4</td>
<td>Study 2 measurement model parcel loadings</td>
</tr>
<tr>
<td>Table 5</td>
<td>Study 2 correlation matrix of latent variables</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1  Study 1 confidence on current item predicted by performance on current item, moderated by emotion perception  p. 30

Figure 2  Study 1 confidence on current item predicted by confidence on prior item, moderated by emotion perception  p. 31

Figure 3  Study 2 structural model with path coefficients ($\beta$)  p. 52

Figure 4  Study 2 confidence predicted by challenge appraisal, moderated by emotion perception  p. 54
List of Appendices

Appendix A  Complete list of ETS measures in study 1  p. 86
Appendix B  Measures used in study 1  p. 87
Appendix C  Study 2 script  p. 89
Introduction

Taking a high-stakes test is an emotional experience. The most obvious emotion related to testing is test anxiety, and a considerable amount of research provides evidence that high-stakes tests are stressful to test-takers, and that this anxiety can hinder performance. The Educational Testing Service (ETS), for example, conducted a survey of over 2,000 students who had recently taken their Graduate Record Examinations (GRE), and found that 42 percent of them reported being nervous while taking the test even if they felt well prepared beforehand, and that correlations between test anxiety and test performance ranged from -.35 to -.40 for each section of the test (Powers, 1986).

Although many people experience negative affect, such as anxiety and frustration, while taking a test, others view the test as an opportunity to demonstrate their competence and therefore may be less likely to experience the same levels of these negative emotions during the test. Adding to the complexity is the fact that people are more or less aware of their own particular emotional experiences (e.g. Mayer & Salovey, 1997; Petrides & Furnham, 2003), and this variability in emotion perception may exacerbate or ameliorate whether negative emotions pose problems for doing one's best on the test.

In addition to these issues, the current studies will examine both proximal and distal factors that link the experience of taking a high-stakes test to one's reactions to the test and, ultimately, to performance on the test. I will begin by discussing the literature on test anxiety and how the goal orientation an individual adopts in an achievement situation, such as taking a high-stakes test, is related to emotions in general and anxiety in particular. Having specified some of the key relations linking tests, test reactions, and performance, I will next consider a potentially important moderating factor: one's
emotional perceptiveness. In particular, I will discuss whether emotion perception will create more cognitive difficulties during test-taking, or whether it will provide a valuable skill that guides emotionally perceptive individuals through the difficult terrain of taking a high-stakes test. Following this discussion, I will present two studies that will shed some needed empirical light on these subjects. Study 1 provides an initial examination of the relevant variables, and study 2 extends and hones our understanding of the relations between emotion perception, test anxiety, and test performance.

**Test Anxiety**

Anxiety describes an aversive emotional and motivational state occurring in threatening circumstances (Eysenck, 1992). Anxiety is of importance to both basic and applied psychologists because it is often associated with adverse effects on the performance of cognitive tasks. State anxiety (the currently experienced level of anxiety) is determined interactively by trait anxiety and by situational stress (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983; Eysenck, 1992). It can be thought of as “a state in which an individual is unable to instigate a clear pattern of behavior to remove or alter the event/object/interpretation that is threatening an existing goal” (Power & Dalgleish, 1997, pp. 206–207).

Test anxiety is conceptualized as a situation-specific form of trait anxiety (Spielberger & Vagg, 1995), based on two factors (Liebert & Morris, 1967): worry, which refers to the cognitive aspect of test anxiety, such as negative expectations regarding the test results, viewing the self as a failure, and perceptions of other potential implications of failing the test; and emotionality, referring to a high arousal level of the autonomic nervous system (e.g., faster pulse and sweating). In terms of performance,
anxiety can be facilitating or debilitating, depending on task difficulty and the
performers' skills and intelligence (Anderson & Sauser, 1995). Meta-analytic studies
have shown negative associations between test anxiety and performance in a variety of
measures and contexts (Zeidner, 2007; 2008). Worry or self-preoccupation is
characterized by concerns over evaluation and failure, and expectations of aversive
consequences (Borkovec, 1994). Worry is activated in stressful situations, and creates
cognitive interference by obstructing the processing and temporary storage capacity of
working memory. Worrisome thoughts consume the limited attentional resources of
working memory, which are therefore less available for parallel task processing.

*Anxiety and Achievement*

Given that individuals typically take high-stakes tests in order to achieve
important goals (e.g., college entrance, graduate or professional school entrance, or to
gain entry into a desired job), a better understanding of emotional reactions such as
anxiety must consider the nature of achievement situations. Achievement situations are
those in which an individual encounters some standard for goal attainment and can expect
to be presented with objective performance feedback to communicate success or failure
(Ames, 1992). Individuals’ goals are thought to exert a broad influence on affect,
cognition, and behavior in achievement settings (Dweck, 1986; Elliot, 1997; Nicholls,
1984). In achievement situations, people generally adopt one of three different goals:
mastery, performance-approach, or performance-avoidance (Elliot & McGregor, 2001).
Which goal an individual adopts is very important to understanding his or her experience
of anxiety during an achievement situation such as taking a test.
Mastery goals are rooted in the desire to improve one's competence during a learning activity. Mastery goals generally cultivate a self-based (or task-based) evaluation of one's competence, and these goals focus attention on developing competence and mastering the task. Achievement in the context of a mastery goal means "making progress." With a performance-approach goal, the individual seeks to demonstrate competence, especially in the presence of an audience. Performance-approach goals generally cultivate a norm-based evaluation of one's competence, and these goals focus attention on the demonstration of ability relative to others (Elliot & Harackiewicz, 1996). Achievement in the context of a performance-approach goal means doing better than others. With a performance-avoidance goal, the individual seeks to demonstrate or prove that he or she is not incompetent, especially in the presence of an audience. Like performance-approach goals, performance-avoidance goals also cultivate a norm-based evaluation of one's competence (Elliot & Harackiewicz, 1996); however, these goals focus attention on the avoidance of a demonstration of low ability relative to that of others. Achievement in the context of a performance-avoidance goal means not doing worse than others.

When individuals pursue either of the first two types of achievement goals, mastery or performance-approach, they experience relatively low anxiety and relatively good test performance (Elliot, McGregor, & Gable, 1999; Barron & Harackiewicz, 2001). Performance-avoidance goals, on the other hand, are the primary cause of worry in achievement settings (Linnenbrink, 2005; Urdan, 2004). That is, the roots of worry are to be found in setting goals for oneself such as, "I just want to avoid making a mistake." Whereas performance-approach and mastery goal-setting orientations toward
achievement situations are negatively related to test anxiety, performance-avoidance goals have been shown to be positively related to test anxiety (Elliot & McGregor, 1999; McGregor & Elliot, 2002).

Test Anxiety and Test Performance

How do emotional reactions help to connect these achievement goals to performance outcomes? There are two mechanisms to consider: the test-taker’s cognitive appraisal of the testing situation, and their tendency to attend to or ignore the emotional information contained in this appraisal. The latter issue will be addressed below in the framework of emotional intelligence theory. Cognitive appraisals determine whether a given situation is regarded as significant for well-being, and if so, whether it is primarily threatening (containing the possibility of harm or loss), or challenging (holding the possibility of mastery or benefit). Blascovich and colleagues (Blascovich & Tomaka, 1996; Blascovich & Mendes, 2000) have developed a biopsychosocial model of challenge and threat that distinguishes challenge reactions to motivated performance situations from threat reactions using patterns of cardiovascular responses. They define motivated performance situations as “goal-relevant and, hence, task engaging for individuals, and that require instrumental cognitive responses” (Blascovich et al., 2004, p. 683-684), and they use test taking as an example of such a situation. According to the biopsychosocial model, challenge appraisals occur when resources (e.g. knowledge, skills, or abilities) are perceived to meet or exceed situational demands (e.g. effort, uncertainty, risk), and threat appraisals occur when perceived demands exceed perceived resources.
How can the emotions resulting from such cognitive appraisals be described? Discrete emotions can be classified on a number of dimensions (e.g., object focus, valence, activation, duration, intensity). As proposed by Pekrun (1992; Pekrun et al., 2002), two such dimensions of particular importance for achievement emotions are object focus and valence. Valence simply refers to whether the situation being appraised is one that is positive or negative in its implications for the individual. Object focus refers to whether the cause of the emotional reaction is activity-related or outcome-related.

Activity-related emotions are emotional reactions induced by the characteristics of a task process itself and include, for example, enjoyment of learning, boredom experienced during learning, and anger about learning assignments or task demands. Outcome-related emotions arise from an individual's perceptions of the broader impacts of an event or situation and can include both retrospective outcome emotions (e.g., pride and shame following success and failure, respectively) and prospective, anticipatory outcome emotions (e.g., hope, anxiety, and hopelessness relating to upcoming success or failure).

Emotions can impact performance by changing brain dopamine levels affecting the consolidation of episodic memory (Ashby, Isen, & Turken, 1999), by directing attentional processes and the use of cognitive resources (Meinhardt & Pekrun, 2003), by inducing and sustaining student interest in learning material (Ainley, Corrigan, & Richardson, 2005; Krapp, 2005), by triggering different modes of information processing and problem solving (Isen, 1993), and by facilitating or impeding self-regulation of learning and performance (Pekrun, Goetz, Titz, & Perry, 2002). Anxiety, in particular, has been associated with a narrowing of attention (Easterbrook, 1959), heightened emphasis on negative information (Riskind, Williams, Gessner, Chrosniak, & Cortina,
2000), interference with working memory (Hayes, MacLeod, & Hammond, 2009), and
generally results in impaired cognitive processing (Beal, Weiss, Barros, & MacDermid,
2005; Eysenck, Derakshan, Santos, & Calvo, 2007).

In one of the first studies of the effects of test anxiety on performance, Mandler
and Sarason (1952) found that those low in test anxiety (as assessed by a self-report
instrument created by the authors) outperformed those high in test anxiety on an
intelligence test. Research into the performance effects of test anxiety has focused
primarily on the role of anxiety as a form of cognitive interference (Sarason, 1984;
Sarason, Sarason, & Pierce, 1990). Wine’s (1971) direction of attention hypothesis, for
example, argues that those high in test anxiety divide their attention between self-relevant
and task-relevant factors under evaluation conditions, whereas the low test-anxious
person concentrates on the task.

The anxious test-taker, therefore, is more frequently focused on self-evaluative,
self-deprecatory thinking (Deffenbacher, 1978; Morris, Davis, & Hutchings, 1981;
Sarason, Sarason, Keefe, Hayes, & Shearin, 1986). This self-focused cognitive activity
occurs during task performance, and it diverts attention away from the task and relevant
performance cues, undermining effective task performance, much in the way that
rumination interferes with cognitive tasks (Lyubomirsky et al., 1999). Self-attention (i.e.
meta-cognitions regarding one’s own abilities and reactions) has been found to have a
negative influence on test performance when coupled with high test anxiety. Being
observed by an audience is one method of heightening this self-attention. Ganzer (1968),
for example, randomly assigned subjects to observed and unobserved conditions and
asked them to learn and later recall lists of nonsense syllables. Observed subjects recalled the lists significantly less efficiently than unobserved subjects.

_Emotiona lIntelligence_

The preceding discussion makes it clear that emotional reactions, and anxiety in particular, play a critical role in determining outcomes in achievement situations such as test-taking. How an individual appraises the testing situation is key to understanding their reactions to it, and individuals differ in the extent to which they attend to and use such emotional information, both interpersonally and intrapersonally (Petrides & Furnham, 2003). However, the mechanisms behind this relation and its boundaries have not been fully explored in past research. An approach that may serve to clarify the role that emotions, especially anxiety, play in test performance is to examine this relation in the context of existing theories that attempt to explain the nature of emotions and their impact on performance. One prominent approach to understanding emotion can be found in the ability-based theory of emotional intelligence.

What is the value of emotional intelligence? A bevy of research over the past decade has sought to answer this question, as well as the more fundamental question, "is there such a thing as emotional intelligence?" To the second question, the research at this point would seem to indicate that the answer is a hesitant "yes." There does appear to be a meaningful construct that can be labeled emotional intelligence that exists conceptually somewhere between personality and cognitive ability, although there are a number of caveats to this conclusion that will be discussed shortly. However, the answer to the first question, which is essentially the question of the predictive validity of emotional intelligence, remains ambiguous. Many claims have been made regarding the real world
outcomes of being emotionally intelligent, but the empirical evidence for these positive effects is limited.

When evaluating the predictive validity of any construct, one must first determine what criteria the construct could be reasonably expected to predict. In the case of emotional intelligence (EI), we must consider what the practical results are of being able to accurately recognize the emotions of others, understand various emotional states and their antecedents, and regulate one’s own emotional responses to events – some of the characteristics included in most definitions of emotional intelligence. It has been claimed that EI predicts successful workplace behavior to a greater extent than intelligence (Cooper & Sawaf, 1997; Goleman, 1998; Hay Group, 2000; Weisinger, 1998). However, a recent review failed to uncover any empirical studies that found EI to be predictive of job success when controlling for personality and intelligence (Zeidner, Matthews, & Roberts, 2004). In fact, there is some evidence that emotional perceptiveness in particular can have negative consequences, causing those who are more emotionally perceptive to be more strongly impacted by stress (Ciarrochi et al., 2002). Such findings regarding emotion perception could point to a potential negative side effect of being emotionally intelligent, specifically that those are more emotionally perceptive may be more prone to experience the negative effects of aversive emotional states, and experience them more strongly, than those who are less emotionally perceptive.

One performance area that would seem in theory to be particularly susceptible to the influence of emotional intelligence is test performance. Much of the literature on emotional intelligence has assumed an overall positive relation between EI and positive behavioral outcomes; however, the actual empirical evidence for these relations is sparse.
Within the context of test performance, we begin to see a potential explanation for these equivocal findings. The skill of being emotionally perceptive, the cornerstone of most conceptualizations of EI, may have both positive and negative consequences. Furthermore, emotional perceptiveness may be more or less relevant to performance outcomes depending on the valence of the emotions being dealt with in a given situation. The studies presented here will attempt to describe the boundary conditions under which emotional perceptiveness may or may not influence reactions to a test and actual test performance, and they will also attempt to elaborate on the psychological mechanisms of these proposed relations.

*What is Emotional Intelligence?*

There is no clear consensus regarding a single definition or model of emotional intelligence. Researchers often distinguish between ability models and mixed models. Mixed models, which are also called trait-based models and incorporate personality, cognitive ability, and other individual difference variables, appear to define EI by exclusion – EI represents a constellation of qualities that are not captured by IQ, including personality traits, perceptions, motivational factors, and other dispositions (Bar-On, 1997; Goleman, 1995). Bar-On described this model of emotional intelligence as "an array of noncognitive capabilities, competencies, and skills that influence one’s ability to succeed in coping with environmental demands and pressures (Bar-On, 1997, p. 14).” In contrast, Mayer and colleagues (Mayer & Salovey, 1997; Mayer et al., 2000) have defined emotional intelligence as a set of abilities dealing with the expression and monitoring of one’s own emotions, the appraisal of others’ emotions, discriminating
between different emotional states, and using emotional information to guide decision making and other cognitive processes.

The ability model

By isolating the ability-based aspects of emotional intelligence, Mayer and Salovey hoped to better understand these characteristics and the degree to which they contributed to individuals’ behavior independent of the contributions from other traits that had already been explored in the context of personality. In an attempt to integrate the disparate bodies of research in social intelligence, the appraisal and communication of emotions, and the use of emotions in problem solving, Mayer and Salovey (1997) proposed that EI is comprised of four abilities, what they refer to as the four-branch model of EI. These four abilities, beginning with the more basic abilities and moving to the more complex, are: (1) the perception of emotions in the self, in other people, and in objects, art, and stories (emotion perception); (2) the ability to generate emotions for use in other mental processes (emotion facilitation); (3) the ability to understand and reason about emotions, transitions between emotions, and relations between emotions (emotion understanding); and (4) the ability to moderate emotions in oneself and in others (emotion management). Each of these abilities has developmental stages within it, and according to the theory those who are higher in emotional intelligence will progress through these stages more quickly than those lower in emotional intelligence (Mayer & Salovey, 1997).

The dimension of emotional intelligence that was addressed in the present research, and which according to Mayer and Salovey’s (1997) model underlies the other three branches in their four-branch model, was emotion perception. Emotion perception
is defined as the ability to perceive and identify emotions in oneself and others, as well as the appreciation of the emotional aspects of works of art, music, and stories (Salovey, Mayer, Caruso, & Lopes, 2003). Mayer and Salovey (1997) proposed four developmental stages of the emotion perception ability: (1) being able to identify one’s own emotions through one’s physical state and thoughts; (2) being able to identify emotions in other people, designs, or artwork based on appearance or behavior; (3) being able to accurately express emotions and identify needs associated with those emotions; and (4) being able to distinguish between accurate and inaccurate expressions of emotion. As will be discussed below, emotion perception has been found to be the most valid and reliable of the four dimensions of the ability model in studies of the various measures of EI (Davies, Stankov, & Roberts, 1998; Ciarrochi, Chan, & Caputi, 2000). In studies using the MSCEIT, the tasks assessing emotion perception have been the only subscales to consistently yield high reliabilities (MacCann, Roberts, Matthews, & Zeidner, 2004; Mayer et al., 2000; Palmer et al., 2005).

*What Outcomes Does Emotional Intelligence Predict?*

Despite the attention researchers have given to the measurement and conceptualization of emotional intelligence, it remains unclear what it can and cannot predict. A considerable amount of research has focused on success in the workplace and interpersonal relations, finding that EI is positively correlated with both the size and quality of social networks (Ciarrochi, Chan, & Bajgar, 2001), optimism (Schutte et al., 1998), prosocial behavior, parental warmth, and positive relations with peers (Mayer et al., 1999; Feyerherm & Rice, 2002). Research also suggests that EI is negatively related to self-destructive behaviors (Brackett & Mayer, 2003; Trinidad & Johnson, 2002),
truancy from school (Petrides, Frederickson, & Furnham, 2004), and depression (Dawda & Hart, 2000; Schutte et al., 1998).

The predictive validity of EI, and its relation to other measures, has largely been a function of the type of EI measure being used in a given study. The two conceptualizations of EI, the mixed and ability models, have led to two categories of assessment – self-report and performance-based measures. Self-report measures sample a diverse range of constructs and therefore are well suited for the mixed model of EI. Ability-based measures of EI follow the ability model of emotional intelligence, and use performance criteria analogous to those of cognitive ability tests, that is, having a person solve a problem, and evaluating their response against criteria for accuracy. The performance-based measures that have been the subject of the greatest amount of research are the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT; Mayer, Salovey, & Caruso, 2000) and its predecessor the Multi-factor Emotional Intelligence Scale (MEIS; Mayer, Caruso, & Salovey, 1999).

Self-report EI and EI assessed through ability-based measures appear to have low or no correlations (Ciarrochi et al., 2002) suggesting that these differing assessment methods are in fact tapping into different constructs. Research indicates that self-report measures of EI have considerable overlap with personality measures and little correlation with cognitive ability measures, and that the reverse is true of ability-based measures of EI (Roberts, Zeidner, & Matthews, 2001; Bar-On, 1997; Newsome, Day, & Catano, 2000; Dawda & Hart, 2000; Brackett & Mayer, 2003; O’Connor & Little, 2003). The degree of overlap between self-reported emotional intelligence and extant personality
measures has in some studies been found to be sufficiently extensive as to call into question the value-added of such measures (Davies, Stankov, & Roberts, 1998).

The relation between ability-based measures of EI and personality traits, on the other hand, tends to be much more modest. For example, the correlations between the MEIS and MSCEIT and the traits of the five-factor model of personality are typically less than .30 (Lopes, Salovey, & Straus, 2003; Roberts, Zeidner, & Matthews, 2001). However, evidence does suggest that ability-based EI does relate to measures of cognitive ability. Studies have found positive relations between EI and SAT scores and other college admission criteria, college grades, and WAIS scores (Lopes et al., 2003; Brackett & Mayer, 2003; Davis & Kraus, 1997). Schulte, Ree, and Carretta (2004) found a correlation of .45 between MSCEIT scores and scores on the Wonderlic intelligence test. This relation appears to be strongest for the emotion understanding branch of the ability model, likely due to the fact this branch is assessed in a manner analogous to a vocabulary test (Lopes et al., 2003; Roberts, Schulze, Zeidner, & Matthews, 2005), but is found in the other three branches of the model as well.

Intercorrelations between the four branches of the ability model as assessed by the MSCEIT range between .25 and .64 (Livingstone & Day, 2005), and several studies have confirmed that the four-factor model fits the data in the MSCEIT well (Livingstone & Day, 2005; Brackett & Mayer, 2003; Day & Carroll, 2004; Mayer, Salovey, & Caruso, 2002).

In sum, the self-report assessments of EI that have been employed in past research suffer from a number of measurement issues due to their overlap with existing trait measures. Ability-based measures of EI are also in their developmental stages, but appear
to demonstrate sufficiently robust psychometric validity, have been used in the literature enough to enable comparisons with other research, and their dissimilarity from personality measures allows for more ease of interpretation when examining outcomes.

EI and behavioral outcomes

Much of the validation research that has been conducted on emotional intelligence has used criteria related to personal adjustment. Theoretically, a person with high emotional intelligence should be able to correctly interpret the moods of others and appropriately convey their own emotions, and therefore have a higher chance of forming close relations and receiving social support. Questionnaire measures of EI, in particular, have been found to be predictive of criteria such as life satisfaction, social support, and low levels of depression and stress (Bar-On, 1997; Ciarrochi, Chan, & Bajgar, 2001; Ciarrochi, Dean, & Anderson, 2002; Saklofske, Austin, & Minski, 2003). A major limitation of this research is its reliance on self-report measures of both the predictor and the criteria. Ability-based measures have also demonstrated modest relations with adjustment criteria. Studies of the MEIS suggest that this index of EI correlates significantly with life satisfaction and self-reported social skills (Mayer et al., 1999). In the case of self-report measures, these relations might be accounted for, at least in part, by the overlap of these self-report measures with established personality traits (Matthews, Roberts, & Zeidner, 2004). Newsome, Day and Catano (2000) found a correlation of -.77 between the EQ-i and trait anxiety, and Dawda and Hart (2000) found correlations approaching .50 between measures of the Big Five personality traits and the EQ-i.

Looking at ability-based EI, Day and Carroll (2004) found that emotion perception (as assessed by a part of the MSCEIT that asks for interpretations of faces and
designs) was the only dimension of EI related to individual performance on a cognitive decision-making task. The ability to recognize nonverbal messages from others, key to the skill of emotional perceptiveness, has been shown to have a positive link to job effectiveness (Rosenthal, Hall, DiMatteo, Rogers, & Archer, 1979). However, there is also evidence that this ability to read emotional signals from others can be maladaptive, and that those who are skilled at perceiving nonverbal messages from others may experience greater difficulty in their social relationships (Blanck, Rosenthal, Snodgrass, DePaulo, & Zuckerman, 1981; Rosenthal & DePaulo, 1980). Reviewing the relation between social outcomes and the ability to infer the thoughts and feelings of others ("empathic accuracy"), Simpson, Ickes, and Blackstone (1995) found positive relations when the issues under consideration were not threatening, but negative relations when the issues were divisive. When discussing potentially harmful issues, it appears that the ability to accurately understand the emotional states of others can have a negative influence on social adjustment.

Closely related to how an individual reacts to threatening issues interpersonally is how they cope with stress in general. An individual’s perceptions of stressors and reactions to those stressors have a large impact on performance. As stress, and the perception of it, is conceptually intertwined with emotional intelligence and emotion perception in particular, we can begin to understand how emotion perception impacts performance outcomes by examining the relation between emotion perception and stress. Salovey et al. (2000) have suggested that EI should be related to an individual’s ability to cope with stressful situations, in that those higher in overall EI are better able to perceive, appraise, and regulate their emotions. However, there is empirical evidence that those
who are skilled at perceiving emotions may, in fact, be more sensitive to the effects of stress. Simpson et al. (1995) demonstrated that inaccuracy in perceptions of thoughts or feelings between the members of dating couples might protect the couple from relationship stress. In addition, studies have shown that the tendency to think about thoughts and feelings is associated with higher anxiety, depression, and paranoia, and with lower self-esteem (Farber, 1989; McCallum & Piper, 2000).

In a similar manner, Lyons and Schneider (2005) found that emotion perception was negatively related to ratings on a speech performance task for males, but positively related to ratings on the same task for females. The speech task called for participants to defend themselves against an allegation of sexual harassment, and the authors proposed that the male participants might have been more defensive than the females while performing this task. Those males who were higher in emotion perception may have been particularly aware of the adverse emotions they were experiencing, hampering their ability to deliver an effective speech.

Lyons and Schneider (2005) also presented subjects with a mental arithmetic task (counting backwards by sevens from a four-digit number as quickly and accurately as possible) and a post hoc analysis confirmed that males reported feeling more threatened by the speech task than the math task. The finding that, for males, emotion perception was not related to performance on the math task but was related to performance on the speech task, would seem to suggest that, to at least some extent, awareness of their stress levels had an impact on their performance.

Lyons and Schneider's (2005) finding of a positive relation between emotion perception and performance of the speech task among females suggests that emotional
perceptiveness is helpful when the task is not particularly threatening or stress-inducing. This finding, along with Day and Carroll's (2004) finding that emotion perception was associated with enhanced task performance, indicates that emotion perception is indeed beneficial in some circumstances, in spite of the negative consequences it can be associated with under stressful conditions.

From this review of the literature relating to emotion perception, it appears that this EI-related ability does have implications for the outcomes related to stress-inducing situations. In particular, it would seem that emotional perceptiveness might have both negative and positive consequences for anxiety-laden events, depending on other contextual factors. Pertinent to the present studies is the potential that emotion perception has for influencing test-taking performance, a specific situation in which anxious reactions have well-documented consequences.

Study 1

How a person reacts to a stressful situation is a major determinant of how well they will perform in said situation. Affective states have been found to influence individuals in a number of situations, including athletic competitions (Beedie, Terry, & Lane, 2000) and academic examinations (Catanzaro, 1996; Lane, Whyte, Terry, & Nevill, 2005; Totterdell & Leach, 2001). One hypothesis put forward to explain these effects is that a person's affective, or emotional, state provides information to the individual about how they can best cope with the stressful situation they have found themselves in (Bless, 2001; Gendolla & Krusken, 2002). Furthermore, an individual's emotional reactions to a situation are a consequence of their cognitive appraisal of that situation. Cognitive appraisal theories have been put forward to explain why a certain event or stimulus
prompts one emotional reaction as opposed to another (Lazarus, 1991; Smith & Ellsworth, 1985). Such theories posit that the emotional reaction to an event is a function of how the event is evaluated in relation to the individual’s goals (Roseman & Smith, 2001).

A study was conducted in order to examine the relations between emotion perception, test anxiety, and test performance more closely. This study proposes that the extent to which emotional reactions will influence performance in a high stress situation is a function not only of the individual’s cognitive appraisal of the situation, but also of their cognizance of the emotional content of that appraisal. Research, as well as the conceits of the ability model of emotional intelligence, has suggested considerable overlap between the ability to perceive emotion in oneself and the ability to perceive others’ emotions (Wong & Law, 2002). The literature on emotional intelligence and other models that describe emotional perceptiveness as an ability have made it clear that there are considerable individual differences in the ability to perceive emotion (Ekman & O’Sullivan, 1991; Matsumoto et al., 2000).

This study uses high-stakes testing as the setting for the stressful situation, and examines test anxiety as the salient emotional reaction. Based on the above discussion, it was anticipated that test anxiety would negatively affect test performance, and that this negative impact of anxiety on performance would be contingent on the test-takers’ emotional perceptiveness. The anxious test-taker is more frequently focused on self-evaluative, self-deprecatory thinking. This self-focused cognitive activity occurs during task performance, and it diverts attention away from the task and relevant performance cues, undermining effective task performance (Deffenbacher, 1978; Morris, Davis, &
Hutchings, 1981; Sarason, Sarason, Keefe, Hayes, & Shearin, 1986). To the extent that emotion perception and cognitive appraisal styles are individual differences, it is to be expected that some people will be more greatly impacted by their emotional reactions to a situation than others. Some people may tend to perceive stressful situations as more threatening than others, and if they also happen to be particularly emotionally perceptive they may have a much more difficult time during those stressful situations.

Given the above described individual differences in emotion perception and cognitive appraisal, it was predicted that those individuals who were dispositionally more anxious and worried about taking tests would tend to perform more poorly on a given ability test. It was also predicted that this relation would be weaker among those test-takers who receive low scores on a test of emotion perception than it would be among those who receive high emotion perception scores.

Hypothesis 1: Dispositionally anxious reactions to tests will be negatively correlated with test performance.

Hypothesis 2: Emotion perception will moderate the relationship between anxious test reactions and test performance such that the negative correlation between anxious test reactions and test performance will be greater for those who score higher in emotion perception.

In addition to the hypotheses regarding the relations between test anxiety, performance, and emotion perception, the role that real-time confidence assessments might have on performance, and how emotion perception may be influencing test-taker confidence was also of interest. Although no hypotheses were made regarding this confidence perception it was anticipated that emotional perceptiveness could affect the
relationship between item-by-item confidence and performance, and that it could also impact confidence across items. To the extent that the RTTQ represents a trait-based tendency to be more or less negatively affected by tests, and the item-by-item confidence ratings obtained in this measure are more situationally-dependent, it was of interest to examine whether there were distinctions between how emotion perception interacted with trait-based test reactions and situation-dependent reactions.

In order to better understand the mechanism underlying how appraisals and emotion perception affect both test reactions and test performance, it will be necessary to explore both the trait-based distal indicators of reactions, and the more situationally-dependent, or proximal, indicators. Here, and in the discussions below, the term distal is used to describe a measure that is relatively removed from a specific testing situation in terms of content, addressing more generalized tendencies and dispositions. Proximal here refers to measures that are more closely linked to the testing situation under analysis in terms of content.

Study 1: Overview

The data for study 1 was drawn from a set of data collected by the Educational Testing Service (ETS) to learn more about student attitudes and values. Participants completed twenty-six self-report measures, including a Vocabulary Test (Ekstrom et al., 1976), and the Reactions to Tests Questionnaire (Sarason, 1984; see Appendix A for complete list of measures included in the ETS study). As part of this study, ETS was also interested in obtaining data on four newly developed measures of emotional intelligence: the Self-Perception Index of Emotions, the Emotion Strategies Index, the test of Affective Quote Completion, and the test of Emotion Perception in Stories.
The current study capitalized on the data available from ETS that was relevant to the issues at hand. The Vocabulary Test included in the battery of measures would serve as a useful ability test, performance on which could be gauged against other measures, and it also included a component of self-perceived confidence. For each item on the vocabulary test, participants were asked to rate how certain they were that they had answered the questions correctly. As will be discussed further below in Exploratory Analyses, it was thought that the conceptual overlap of these certainty ratings with emotional reactions could prove useful as well. Data from the Reactions to Tests Questionnaire was also relevant to the current research. The Emotion Perception in Stories test was employed as the emotional intelligence measure most directly relevant to the concepts under study.

Study 1: Method

Participants

Participants were 431 undergraduate students (270 females) drawn from two-year and four-year institutions located in New Jersey, Maryland, Arizona, and Texas. Participants’ ages ranged from 16 to 57 years \( (M = 22.6, SD = 5.5) \), and they described themselves as Black \( (n = 171, 39.7\%) \), White \( (n = 149, 34.6\%) \), Asian \( (n = 49, 11.4\%) \), Latin American \( (n = 15, 3.5\%) \), Puerto Rican \( (n = 8, 1.9\%) \), American Indian \( (n = 3, 0.7\%) \) or Other \( (n = 20, 4.6\%) \). Most participants learned to speak English first \( (n = 331, 76.8\%) \), some had learned English and another language \( (n = 68, 15.8\%) \) and relatively few first learned a language other than English \( (n = 32, 7.4\%) \).
Measures

*Emotion Perception in Stories.* This was an experimental measure created by ETS and used for the first time as part of this assessment battery. In this assessment of emotion perception, respondents are presented with eight short stories (previously generated by other students from a small, private university in Texas), and are asked to evaluate the extent to which each of seven emotions are expressed in the story (1 = 'not at all,' 5 = 'very strongly'): happiness, sadness, contempt (dislike, disdain), disgust, anger, surprise, and fear. This assessment was scored using the consensus method (Legree, Psotka, Tremble, & Bourne, 2005), such that an individual who provides the same response to an item as half of the sample will receive a score of .50 for that item. Refer to Appendix B for an example of one of the vignettes used in this measure.

*Reactions to Test Questionnaire (RTTQ).* This 40-item questionnaire, created by Sarason (1984) consists of four 10-item subscales: Worry (“Before taking a test, I worry about failure”), Tension (“I feel distressed and uneasy before taking tests”), Test-Irrelevant Thinking (“During tests, I think about recent past events”), and Bodily Symptoms (“My stomach gets upset before tests”). Responses are given on a 6-point scale ranging from “never” to “always.”

*Vocabulary Test II.* This test, created by ETS (Ekstrom et al., 1976), consists of 36 items, for which respondents are asked to select the one word from a set of five that is most closely related in meaning to a target word.

*Confidence.* After each item on the above described vocabulary test, ETS inserted questions asking participants “How confident are you that your answer is correct?” and
presented them with response options ranging from 20% to 100%, in increments of 10 percentage points.

Procedure

Participants were compensated $50.00 for completing the survey. Educational Testing Service and institutional permissions were granted to conduct this study by respective human subjects' committees. Participation was conditional on the individual agreeing to the consent form, which explained the purpose of the study, the amount of compensation, that participation was voluntary and withdrawal at any time was allowed, and the confidentiality of responses.

A proctor was designated at each institution that was responsible for recruiting volunteers, scheduling computer lab time, and proctoring during test sessions. Proctors recruited students by posting signs around campus advertising the study that stipulated participation was offered to current students who had already taken the Scholastic Assessment Test (SAT). Each proctor was compensated $5.00 for every participant he or she recruited. Participants were administered the battery of tests in groups that averaged 20 in size. Tests were administered via a computer, and participants each sat at individual computer workstations. Sessions typically took from 1 to 1.5 hours.

Study 1: Results

Measurement Properties

An exploratory factor analysis (principal components analysis with varimax rotation) determined that the original delineation of four dimensions was adequate to describe the structure of the Reactions to Tests Questionnaire. Exploratory factor analysis of the Emotion Perception in Stories measure (also principal components analysis with
varimax rotation) found some dimensionality due to differences between the stories, but there was one overarching factor with an Eigenvalue of 10.3 accounting for 18.4 percent of the overall variance in responding, with item loadings ranging from .01 to .75. Test-takers got an average of 24.59 questions correct on the 36-item vocabulary test (standard deviation = 5.22). Means, standard deviations, internal consistency coefficients, and zero-order correlations between the measures are reported in Table 1.

**Hypothesis Tests**

Moderated multiple regressions were run to determine if any of the dimensions of the Reactions to Tests Questionnaire were predictive of test performance in accordance with hypothesis 1, and whether an interaction between emotion perception and any of those dimensions was predictive of test performance as stated in the second hypothesis. Bodily symptoms and tension were not found to be related to test outcome ($\beta = -.024$ and $\beta = .019$, respectively), but both worry ($\beta = -.142, p < .01$) and test-irrelevant thinking ($\beta = -.103, p < .05$) were predictive of performance on the vocabulary test. Emotion perception was also strongly predictive of test performance, $\beta = .418, p < .001$, however none of the interactions between emotion perception and any of the four dimensions of test reactions were predictive of performance ($\beta = .030$ for worry, $\beta = -.040$ for test-irrelevant thinking, $\beta = -.005$ for tension, and $\beta = -.011$ for bodily symptoms).

**Exploratory Analyses**

As a measure of general tendencies to react to tests in certain ways the RTTQ is a distal indicator of test reactions, and it is possible that there are more proximal factors influencing test performance and reactions. A proximal indicator of reactions was
Table 1. Study 1 measurement properties

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Alpha</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vocabulary Test</td>
<td>24.59</td>
<td>5.22</td>
<td>.81</td>
<td>.42**</td>
<td>-.17**</td>
<td>.10</td>
<td>-.11*</td>
<td>-.08</td>
<td></td>
</tr>
<tr>
<td>2. Emotion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perception in</td>
<td>22.62</td>
<td>3.54</td>
<td>.90</td>
<td>-.17**</td>
<td>.02</td>
<td>-.12*</td>
<td>-.18**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stories</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Worry</td>
<td>13.75</td>
<td>5.52</td>
<td>.89</td>
<td></td>
<td>.66**</td>
<td>.52**</td>
<td>.58**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Tension</td>
<td>15.26</td>
<td>5.24</td>
<td>.93</td>
<td></td>
<td></td>
<td>.22**</td>
<td>.57**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Test-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irrelevant Thinking</td>
<td>13.32</td>
<td>5.23</td>
<td>.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.41**</td>
<td></td>
</tr>
<tr>
<td>Symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**p < .01, *p < .05
available in this data in the form of the confidence assessments that participants provided
during the vocabulary test. The structure of the vocabulary test, with the confidence
ratings included, made item-by-item analyses possible and allowed a more fine-grained
examination of reactions. In particular, I reasoned that efficacy or confidence on one item
would have implications for one's confidence on a subsequent item. In this sense,
confidence for a given item reflects more immediate appraisals of the test. As such, an
examination of the relations between prior item confidence and confidence on the current
item can provide an index of test-taking appraisals at an item-by-item level of analysis.

Furthermore, consistent with the logic discussed above, an examination of
emotion perception can be included to determine whether the influence of prior
confidence on subsequent confidence is heightened for those who are higher in emotion
perception. This analysis provides another test of the idea that the threatening aspects of
test-taking (i.e., confidence on a given test item) will be perceived more accurately by
those who are higher in emotion perception. The basic model of interest, therefore, was
one in which current item confidence is predicted from prior item confidence, with
emotion perception examined as a moderator interacting with prior confidence. To
disentangle the influence of confidence due to an individual's actual level of performance,
the model also included both current and prior item performance. As a result, the relation
between prior and current item confidence can be interpreted as the extent to which prior
levels of confidence predict subsequent levels of confidence, irrespective of how well the
test-taker actually performed on the items.

The reliability coefficient for the confidence ratings was .93. Average confidence
estimations across all items on the test ranged from .24 (24 percent confident that the
response was correct) to 1.00 (100 percent confident that the response was correct), and the overall average confidence estimation for this sample was .74. Because this model calls for analysis at multiple levels, with nested data in the form of the item-by-item responses for each test-taker, hierarchical linear modeling was used. Level 1 of this model is described by each nested item across each participant, with 14,657 degrees of freedom, and level 2 aggregates the within participant data and deals with the relations between current performance and current confidence, resulting in 417 degrees of freedom. The model regressed confidence level for the current item on performance for the current item, performance for the previous item, and confidence for the previous item. It was found that current item performance predicted confidence on that item, $\beta_{10} = .265$, $t(417) = 49.78$, $p < .001$. Current confidence was also predicted by prior confidence, $\beta_{20} = .090$, $t(14657) = 10.01$, $p < .001$, and by prior performance, $\beta_{30} = -.014$, $t(14657) = 2.72$, $p < .01$ (each test controlled for each of the other variables).

There were also two significant cross-level interactions, in which emotion perception moderated the effect of current performance on current confidence, $\beta_{11} = .283$, $t(417) = 3.07$, $p < .01$, and also moderated the effect of prior confidence on current confidence, $\beta_{21} = -.454$, $t(14657) = -3.04$, $p < .01$. The interaction between emotion perception and current performance appears to indicate that those who are more emotionally perceptive are more accurate in their perceptions of whether or not they responded correctly to an item than those who are less emotionally perceptive. The interaction between emotion perception and prior confidence indicates that the confidence that test-takers reported for an item on the test was less affected by their reported confidence on the prior item when the test-taker was more emotionally
perceptive, whereas the confidence for the prior item was more strongly predictive of confidence on the current item for the less emotionally perceptive test-takers. See Figures 1 and 2 for graphical representations of these interactions.

**Study 1: Discussion**

The worry and test-irrelevant thinking dimensions of the Reactions to Tests Questionnaire appeared to be related to performance on the vocabulary test. This is consistent with existing models of test anxiety that characterize it as a combination of distracting thoughts and negative emotions (Liebert & Morris, 1967). There was a relation between emotion perception and test performance, such that those who were more emotionally perceptive tended to do better on the vocabulary test. However, there were no interactions between test reactions and emotion perception impacting test performance.

Generally speaking, there are two categories of reasons why the predicted interaction did not emerge. One category of reasons involves problems with measures and methods. For example, the RTTQ is a distal measure of test reactions, assessing test anxiety as a trait without taking into account the myriad differences found between specific testing situations and the proximal reactions that those specific characteristics of the environment can induce. Furthermore, the testing situation in which these participants found themselves as part of this study was not presented as a “high-stakes” testing situation, in that the test-takers had nothing to gain or lose due to their performance. As a result, it is possible that the circumstances under which participants took the vocabulary test as part of this study would not have elicited test anxious reactions in accordance with
Figure 1. Study 1 confidence on current item predicted by performance on current item, moderated by emotion perception
Figure 2. Study 1 confidence on current item predicted by confidence on prior item, moderated by emotion perception
the participants’ responses to the RTTQ, as this particular testing situation was unlikely to have resembled the testing situations that have provoked anxious reactions for them in the past. In other words, test anxiety may not have been elicited in this testing situation because the anxiety-provoking environmental cues were not present (e.g., there was no reward or punishment associated with performance on this test, and, as the test was not timed, time pressures that typically occur in high-stakes tests did not occur).

Although the RTTQ, as a distal indicator of trait test anxiety, failed to demonstrate an interaction with emotion perception, it is possible that a more proximal indicator of test anxiety in the domain of vocabulary knowledge, directly relevant to the performance task at hand, or a direct assessment of emotional state going into the test, would have been more susceptible to the influence of emotional perceptiveness. As discussed above, if the trait-based test reactions assessed by the RTTQ were not salient to the test-taker as they took the vocabulary test presented to them in this study, we would not expect to find a perceptible moderating influence of emotional perceptiveness. Real-time, salient test appraisals were available in this study, however, in the form of item-by-item confidence ratings.

Test-takers’ confidence that they had answered a given test item correctly was strongly related to actual performance on that item. In other words, test-takers were generally accurate in their judgments of their item-by-item performance. Furthermore, confidence on an item was predicted by confidence on the prior item, controlling for the influence of performance on that prior item. This implies that irrespective of the test-taker’s performance on a given item, his or her confidence on that item will bleed over to create a sense of confidence on the next item. Put differently, on average, people exhibit a
bias in their confidence for getting an item correct. The source of this bias is prior confidence, and since actual prior performance was statistically controlled in the model its influence has nothing to do with the relation between prior and current confidence.

This finding hints at the possibility of the second category of reasons for not finding the predicted interaction: that the theoretical foundation for the hypothesis is incomplete. Specifically, these follow-up analyses make it clear that feeling confident about one's performance on a test may be as or more important than feeling threatened, anxious, or worried about tests. Certainly it is true that threat appraisals are directly connected to increases in anxiety (Maner, 2009), but it seems worthwhile also to consider the role of challenge appraisals in mitigating anxious reactions to tests.

Furthermore, the role of confidence during the test appears to vary as a function of emotion perception. It was found that the positive correlation between confidence ratings on the prior item and the confidence ratings on the current item was attenuated by emotion perception. Recall that this effect gauged a confidence bias, in that confidence had an influence above and beyond actual prior or current item performance. The pattern therefore suggests that emotionally perceptive individuals are relatively free of this biasing effect of prior item confidence. In addition, those who were emotionally perceptive were also more accurate in their assessments of whether or not they correctly answered each question on the test, in that they provided high confidence ratings for correct responses. It appears then, that people who are higher in emotion perception are better able to gauge their ability on an item-by-item basis. As such, these data suggest that those who are more emotionally perceptive seem able to disregard prior levels of
confidence, place their attention on the item on which they are currently working, and
make a more accurate judgment of their own abilities.

Considering that emotionally perceptive test-takers' confidence ratings on a given
item are more accurate and less likely to be impacted by reactions to prior items than less
emotionally perceptive test-takers, it appears that individuals who are emotionally
perceptive are better able to understand and interpret information in a detailed manner
and make use of the most appropriate information in their environment than those who
are less emotionally perceptive. Emotionally perceptive test-takers seem to be less
swayed by irrelevant information in their assessments of their performance, whereas the
less emotionally perceptive test-takers appear to be more prone to confusion and
considering irrelevant information when assessing their performance. Conceptually this
enhanced capacity for attentional focus does not seem to be inherent in the definition of
emotion perception. It would, however, be consistent with the ability to manage
emotions, or emotion regulation, which has been defined as the ability to initiate,
maintain, and modulate the processes involved in the intensity and expression of
emotions, and which also involves the ability to shift attention either away from or
toward emotionally relevant stimuli (Thompson, 1994).

Although the predicted exacerbating effects of emotional perceptiveness on the
relations between test anxiety and test performance were not found, the role that
confidence appeared to play in the relation between emotion perception and test
performance warrants further inquiry. These findings suggest that appraisals of threat
may not be the only factor to consider in testing situations. Although study 1 focused
initially on the negative, distracting appraisals of tests, immediate levels of confidence
appeared to be an important factor in this situation. Test-takers' confidence with regard to one item appeared to have an influence on their reactions to the next item, and this effect was weakened when the test taker was emotionally perceptive. The next study therefore incorporated not only threatening perceptions of tests, but also the extent to which individuals see tests as a challenging situation in which they can be confident.

These findings also indicate that it may be worthwhile to examine the differential impact of distal and proximal indicators of test reactions in more depth. Study 1 addressed dispositions toward anxious test reactions as predictors of performance in a specific testing situation; however, as the findings related to item-by-item confidence suggest, there appear to be intervening factors between dispositional test reactions and real-time assessments during a testing situation. Study 2 will further explore how and when distal appraisals may differ from proximal emotional reactions.

Study 2

Study 1 made it clear that emotion perception does have an impact on both test reactions and test performance, but the nature of that impact appears to be influenced by a number of factors. In particular, due to the nature of the testing situation in this study it appeared that confidence was the more salient appraisal, rather than threat, and it also appeared that real-time reactions to the testing situation were not affected by distal indications of a test-taker’s reactions. In study 2 these factors were broken out and examined more closely. The finding in study 1 that emotion perception did not appear to influence the impact of emotional reactions on test performance when those emotional reactions were distal, but did have an effect on confidence judgments on an item-by-item basis, suggested that there may be a more complex multiple step process involved in test
reactions, involving both distal and proximal factors. The transactional process model of testing reactions (Sawyer & Hollis-Sawyer, 2005) suggests that both a person’s individual characteristics (such as emotional perceptiveness and cognitive appraisal style) and their processing of information related to a particular situation (such as confidence) are both factors in an individual’s reactions to a testing experience. Whereas study 1 assessed confidence while taking a test, study 2 addressed confidence prior to the test, and sought to further explore the causal chain relating distal dispositions toward anxiety to actual test performance, and how, if at all, emotion perception is involved in that causal chain.

There are both distal and proximal factors involved in an individual’s cognitive appraisal of a given testing situation. In the case of a high-stakes test, the proximal factors are the test-takers’ immediate reactions to the specifics of the testing situation, their comfort level with the specific content of the test items, and their confidence that they are doing well during the test. Distal factors include the individual’s general attitude toward testing and their past experiences with similar tests.

As discussed in the above literature review, an important distal determinant of how an individual will react in a testing situation is their tendency to perceive stressful situations as either threatening or challenging. Threat appraisals suggest potential danger to one’s well-being or self-esteem (Lazarus 1991; Lazarus & Folkman, 1984; Rapee & Heimberg, 1997; Sarason & Sarason, 1990; Smith, 1991) and low confidence in one’s ability to cope with the threat (Bandura, 1997; Lazarus, 1991; Lazarus & Folkman, 1984; Morris, Davis, & Hutchings, 1981; Smith, 1991). Individuals who consistently experience anxiety in stressful encounters in social (Beck, Emery, & Greenberg, 1985;
Rapee & Heimberg, 1997), test (Sarason & Sarason, 1990), and sporting domains (Wilson & Eklund, 1998) anticipate failure and negative evaluation and interpret such negative outcomes as significant threats to self-identity or self-esteem (Schlenker & Leary, 1982).

However, not all stressors in the environment are perceived as threatening. Individuals can perceive certain stressful situations as a challenge, focusing on opportunities for success, social rewards (e.g., recognition and praise), mastery, learning, and personal growth (Lazarus, 1991; Lazarus & Folkman, 1984; Lazarus et al., 1985). A challenge appraisal indicates confidence that the demands of a stressful situation can be overcome (Lazarus et al., 1985; Park & Folkman, 1997). Challenge experiences can be characterized as personally relevant, requiring significant amounts of effort, and having moderate attributions of self-responsibility or personal control (Smith & Ellsworth, 1985).

It was hypothesized, therefore, that cognitive appraisal style, challenge or threat, will have a main effect on domain-specific confidence. Domain-specific confidence is here defined as an individual’s assessment of their ability to perform well on a test in a particular skill area. A high challenge appraisal style will be associated with positive assessments of domain-specific confidence. A high threat appraisal style will be associated with negative assessments of domain-specific confidence. It was originally hypothesized that emotion perception would interact with cognitive appraisal style, such that the relation between appraisal style and confidence would be greater when emotion perception was high. However, when the results of study 1 were reviewed this hypothesis was updated to reflect these new findings. Due to the attenuating effect that emotion
perception appeared to have on confidence assessments in study 1, it was now predicted that there would be a similar effect in study 2.

Hypothesis 1: Dispositional cognitive appraisal style will be related to domain-specific confidence, such that a high challenge appraisal style will lead to high confidence and a high threat appraisal style will lead to low confidence.

Hypothesis 2: The relation between challenge and threat appraisal styles and domain-specific confidence will be attenuated by high emotional perceptiveness, such that appraisal style will have a greater impact on confidence assessments for those who are lower in emotion perception.

It was expected that domain-specific confidence would be directly related to the immediate proximal affective reactions to a testing situation. According to Skinner and Brewer (2002), coping expectancies (i.e. confidence) will determine whether the fears associated with a threat appraisal will be realized and negative affect will ensue. It is therefore predicted that an individual’s confidence will not directly impact their performance, but that the most direct outcome of being confident that one can do well on a particular test will be the positive affect that one experiences during the test, or conversely if one is lacking confidence about a test in a particular domain one will experience negative affect.

Hypothesis 3: Domain-specific confidence will be negatively correlated with negative affect experienced during a testing situation.

And lastly, it was predicted that the immediate affective reactions to a testing situation will have the most direct impact on performance. As discussed earlier, high levels of anxiety can lead to intrusive thoughts that can be difficult to control and can hinder
performance (Glass & Arnkoff, 1994; Sarason & Sarason, 1990; Sarason, Pierce, & Sarason, 1996).

_Hypothesis 4: Negative affective reactions to a testing situation will be negatively correlated with test performance._

The predictions made by this overall causal model are that the affective reactions that a test-taker experiences during an actual test are a function primarily of their confidence regarding the test material, rather than their dispositional attitudes toward test-taking in general.

There were other limitations of study 1 that the second study attempted to compensate for. First of all, the Emotion Perception in Stories measure employed in the first study was a never before used assessment of emotion perception, and therefore although it has face validity given our understanding of emotion perception as an aspect of emotional intelligence, its true validity as an accurate indicator of emotional perceptiveness is questionable. Study 2 made use of the most commonly employed ability-based measure of emotional intelligence, the Mayer, Salovey, and Caruso Emotional Intelligence Test (MSCEIT; Mayer, Salovey, & Caruso, 2000).

It was also important to examine these effects in the context of a testing situation which more closely approximated a high-stakes testing situation. In study 1, participants had no reason to believe that they had anything to lose or gain through their performance on the test. In order to approximate the psychological environment of a high-stakes testing situation, taking the SAT or GRE in order to get accepted into an undergraduate or graduate university program are the obvious examples, it is necessary to create in the mind of the test-taker some sense of urgency, a threat that something or value could be
lost if they perform poorly, and a sense that there is something of value to be gained from a good performance.

The testing situation in the first study had a further limitation in that it was a test of verbal ability, specifically a vocabulary test. The impact of emotional reactions may differ for a test in a different ability domain. Therefore, in study 2 the test that was used contained mathematical problem solving questions.

**Study 2: Method**

**Participants**

Participants were 152 undergraduate students from the University of Houston and from the internship and cooperative education programs at a federal employer in the Houston area.

**Measures**

*Demographics.* Participants were asked to provide their gender (53.4 percent female, 17.6 percent not reported), ethnicity (31.1 percent White, 8.1 percent Black, 22.3 percent Asian, 18.9 percent Hispanic, 19.6 percent not reported), overall GPA, combined SAT score, combined ACT score, high school GPAs in the subjects of English, foreign language, mathematics, sciences, and social science, and their cumulative high school GPA. Those means and standard deviations can be found in Table 2.

*Mayer, Salovey, and Caruso Emotional Intelligence Test (MSCEIT).* The MSCEIT is an ability-based measure of emotional intelligence intended to objectively assess the four-branch model of EI (Mayer, Salovey, & Caruso, 2000). There are 141 questions in eight subtests, with two sets of tasks assessing each of the four branches. The tasks concerning emotion perception ask respondents to identify the emotions expressed
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall GPA</td>
<td>3.16</td>
<td>.39</td>
</tr>
<tr>
<td>Combined SAT</td>
<td>1149.46</td>
<td>165.88</td>
</tr>
<tr>
<td>Combined ACT</td>
<td>25.08</td>
<td>4.20</td>
</tr>
<tr>
<td>HS English GPA</td>
<td>4.61</td>
<td>.60</td>
</tr>
<tr>
<td>HS Foreign Language GPA</td>
<td>4.45</td>
<td>.67</td>
</tr>
<tr>
<td>HS Mathematics GPA</td>
<td>4.36</td>
<td>.76</td>
</tr>
<tr>
<td>HS Science GPA</td>
<td>4.39</td>
<td>.65</td>
</tr>
<tr>
<td>HS Social Science GPA</td>
<td>4.64</td>
<td>.56</td>
</tr>
<tr>
<td>HS Cumulative GPA</td>
<td>10.49</td>
<td>1.64</td>
</tr>
</tbody>
</table>
in photographs of faces, and the feelings evoked by designs and landscapes. The tasks addressing emotion facilitation ask respondents to describe emotions using non-feeling vocabulary, and to identify emotions that would facilitate or interfere with the performance of certain cognitive or behavioral tasks. Understanding emotions is assessed with questions regarding changes in emotions over time, and how emotions blend and interact with each other. Management of emotions is assessed by presenting respondents with scenarios and eliciting the most adaptive ways to regulate one's feelings in those situations, and how to regulate emotions that arise in social situations and in others. Responses were scored by the test publishers using the consensus method to compare individual responses to those of a normative sample.

The Emotion Perception in Stories measure used in study 1 addressed only one of the four branches of the ability model of emotional intelligence. However, its high internal reliability (α = .90) and statistically significant relationship with the vocabulary test used in study 1 (r = .42) make it psychometrically comparable with the MSCEIT, which has been found to have high reliabilities at the overall test level (test-retest reliability was found to be .86 over a three-week interval; Brackett & Mayer, 2003), and has similar correlations with verbal SAT scores (rs ranging from .23 to .39; Brackett, Mayer, & Warner, 2004). The full MSCEIT, including tests of all four branches of the ability model of emotional intelligence, was employed in this study to assess possible moderating effects of the other three facets of EI, particularly emotion management, on emotion perception. As was conjectured in study 1, it is possible that the observed impact that emotion perception has on test reactions may be influenced to some extent by emotion management. Previous research using the MSCEIT has found low to moderate
correlations between the tasks comprising the emotion perception branch and those comprising the emotion management branch (Mayer, Salovey, & Caruso, 2000, for example), indicating that those who are emotionally perceptive also tend to be relatively good at managing their emotions. This correlation may help to explain why those who are high in emotion perception are less susceptible to irrelevant emotional information while performing a task such as taking a high-stakes test.

*Cognitive Appraisal Style.* Trait cognitive appraisal style was assessed by a modified version of the 18-item Cognitive Appraisal Scale (Skinner & Brewer, 2002), composed of Threat Appraisal and Challenge Appraisal subscales. The Threat Appraisal subscale consists of one item from the State-Trait Anxiety Inventory Trait scale (Spielberger et al., 1983) and the nine-item Self-Presentation Concerns Questionnaire (Skinner & Brewer, 1999). The Challenge Appraisal subscale was developed by Skinner and Brewer to assess the anticipation of positive outcomes and the expectation of success, with one item taken from the Optimism-Pessimism Questionnaire (Norem & Cantor, 1986). Items were re-worded to apply specifically to the context of taking a math test ("When taking math tests, I lack self-confidence"). This distal assessment of appraisals was employed here, as opposed to the Reactions to Tests Questionnaire used in the first study, because it considers the possibility of both negative and positive reactions rather than just the negative anxiety-based reactions of the RTTQ, allowing for a more thorough examination of the impact of confidence on test-taking. However, it is to be expected that there will be consistency across the two measures given that the RTTQ is an indicator of cognitive appraisal just as the Cognitive Appraisal Scale is, simply focused on appraisals in a specific situation.
Domain-specific Confidence. Participants were presented with Part 1 of the Mathematics Aptitude Test contained in the Educational Testing Service's Kit of Factor-Referenced Cognitive Tests (Ekstrom et al., 1976) and asked to evaluate their confidence that they could determine the correct answer to each of the 15 items (20% confident, 30% confident, ...100% confident), and also to rate their confidence that they could correctly answer the question in less than one minute. These items were used in order to capture self-efficacy for as relevant a domain as possible.

Mathematical Problem Solving. The mathematical performance measure was Part 2 of the Mathematics Aptitude Test contained in the Educational Testing Service's Kit of Factor-Referenced Cognitive Tests (Ekstrom et al., 1976). Test-takers were asked to indicate the correct answer to 15 word problems in mathematics from a set of four possible answers with a time limit of 12 minutes.

Immediate Affective Test Reactions. Affective reactions to the test of mathematical problem solving were assessed by presenting the test items to participants a second time, and asking them to think back to the test and evaluate their immediate emotional response upon first seeing each item. Respondents were asked to rate on a seven-point scale ranging from 'very much' to 'not at all' the extent to which they experienced the emotions of anxiety, frustration, happiness, and sadness for each of the items that they responded to on the math test. This method of assessing real time test reactions after the event was used in order to minimize the invasiveness of the assessment during the test. Asking the test-taker to periodically indicate their emotional state while they are actually taking the test, although an apparently intuitive means of obtaining this
data, it was feared that this would also be likely to have an unwanted effect on both performance and reactions.

**Procedure**

The first part of the study was administered through an internet link distributed to participants electronically, either through email or through the University of Houston's School of Business online research participation system. In this first part, participants completed the Cognitive Appraisal Scale and the MSCEIT online, as well as providing the requested demographic information. Upon completion of the measures contained in the first part of the study, participants were invited to schedule a time to complete part two in person. At this session, participants first completed the Domain-specific Confidence Measure, took the Mathematics Aptitude Test, and completed the assessment of their Immediate Affective Reactions, in that order.

After completing the Confidence Measure and before moving on to the Mathematics Aptitude Test, participants were provided with information designed to maximize the realism of the experimental context with that of what actually occurs in a high-stakes testing environment. Specifically, they were informed that the test they were about to take was a measure of intelligence used by some companies and schools to select people for internship and scholarship programs, and that their school was interested in using the test for such purposes as well. The participants were told that the results of this portion of the study, the Mathematics Aptitude Test, would be shared with their university for the purpose of awarding these scholarships and internships, and that the researcher would be reviewing their results with them at the end of the session. The participants were informed that the test was timed and that they had 12 minutes to
respond to as many of the 15 problems as they can, and that they were not to use a
calculator but that there was scratch paper and pencils in the room that they could use.
Although scores were not to be used by the university for these purposes, we told
participants this information to simulate several of the anxiety-inducing characteristics of
high-stakes testing situations: there was a personal reward to be gained or lost (Schwarzer
& Jerusalem, 1992), test-takers were under the constraint of a time limit, and their
performance would be critically reviewed by another individual (Zollner & Ben-Chaim,
1988).

The participants were left alone when completing each of the tests. At the end of
the session participants were debriefed as to the purpose of the study, and informed that
the instructions preceding the Mathematics Aptitude Test were invented and intended to
replicate as closely as possible the motivational conditions that a test-taker encounters
when in a high-stakes testing situation. Participants were also given $10 for their time.
Refer to Appendix C for the script used in this procedure as well as the debriefing
protocol.

Study 2: Results

Sample Screening

Three participants were eliminated from the analyses due to missing data, and
visual inspection of rank-ordered Mahalanobis distances identified one additional outlier
who was also removed from the data, leaving 148 participants with usable data.
Measurement Properties

Reliabilities for each of the eight tasks of the MSCEIT were calculated, as were
the intercorrelations between scores on these tasks. It was found that the only two tasks
that had adequate internal consistency (i.e., $\alpha \geq .70$, per Nunnally, 1978) were the two tasks comprising the MSCEIT's assessment of the emotion perception branch of emotional intelligence: face perception ($\alpha = .84$) and picture perception ($\alpha = .95$). See Table 3 for all of the reliabilities and intercorrelations of the MSCEIT tasks.

Structural Equation Modeling (SEM) was used to test the models in this study; however, due to sample size limitations it was decided that item parcels would be created to serve as indicators of the latent variables of interest (Bagozzi & Edwards, 1998). The technique used to create the item parcels was the single-factor method (Landis, Beal, & Tesluk, 2000), in which a single-factor exploratory factor analysis is specified for each construct, and the items with the highest and lowest factor loadings are paired off and placed into one of four parcels.

This method was used for the measures of emotion perception, domain-specific confidence, challenge appraisal, and threat appraisal. The initial measurement model indicated that the two measures of emotion perception used in the MSCEIT (i.e., pictures and faces) did not appear to reflect a single latent variable; therefore those two measures were parcelled separately. Tests of the measurement models indicated that the item parcels strongly reflected the properties of their latent variables. Because test-takers responded to varying numbers of items on the indicators of immediate anxiety and frustration reactions (see description of these measures above), the different items were averaged together. Then, due to their high correlation ($r = .81$), the average anxiety and average frustration scores were averaged together into a single parcel. A confirmatory factor analysis of these measures suggested excellent fit, $\chi^2 (211) = 255.230, p = .020,$
Table 3. Study 2 reliabilities and intercorrelations between MSCEIT tasks

<table>
<thead>
<tr>
<th>Task</th>
<th># of items</th>
<th>Reliability (α)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Emotion Perception</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Faces</td>
<td>20</td>
<td>.842</td>
<td>.419**</td>
<td>.480**</td>
<td>.356**</td>
<td>.218**</td>
<td>.122</td>
<td>.200*</td>
<td>.285**</td>
</tr>
<tr>
<td>2. Pictures</td>
<td>30</td>
<td>.947</td>
<td>.457**</td>
<td>.428**</td>
<td>.328**</td>
<td>.262**</td>
<td>.300**</td>
<td>.380**</td>
<td></td>
</tr>
<tr>
<td><em>Emotion Facilitation</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Facilitation</td>
<td>15</td>
<td>.386</td>
<td>.597**</td>
<td>.488**</td>
<td>.456**</td>
<td>.465**</td>
<td>.475**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Sensations</td>
<td>15</td>
<td>.512</td>
<td>.605**</td>
<td>.570**</td>
<td>.580**</td>
<td>.668**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Emotion Understanding</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Changes</td>
<td>20</td>
<td>.096</td>
<td></td>
<td>.636**</td>
<td>.617**</td>
<td>.557**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Blends</td>
<td>12</td>
<td>.050</td>
<td></td>
<td></td>
<td>.512**</td>
<td>.588**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Emotion Management</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Emotion Management</td>
<td>20</td>
<td>.408</td>
<td></td>
<td></td>
<td></td>
<td>.673**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Social Management</td>
<td>9</td>
<td>.116</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* *= p < .05; ** *= p < .01
Hypothesis Tests

Hypotheses were examined in two steps. First, the underlying appraisal-reaction model specified that distal test appraisals (challenge and threat) affected the more proximal levels of confidence, which in turn generated negative affect during the test, ultimately influencing test scores. At this step, model fit was assessed and direct and indirect effects were examined. In the second step, emotion perception and its interaction with test appraisals were added to the model. This part of hypothesis testing was conducted as a separate step, as estimation of latent interaction terms prohibits testing of indirect effects (Muthén & Muthén, 1998-2007). Testing both sets of hypotheses simultaneously using SEM would therefore have prohibited testing of the hypotheses involving mediation.

Looking first at the causal chain relating cognitive appraisals to test performance (hypotheses 1, 3, and 4), the overall model fit the data well, $\chi^2(87) = 139.355, CFI = .965, RMSEA = .064$ (see Figure 3). Performance on the math test (percent of items correctly answered) was predicted by negative affect ($\beta = -.178, t(147) = -2.148, p < .05$), negative affect was predicted by confidence ($\beta = -.395, t(147) = -4.876, p < .05$), and confidence was predicted by challenge appraisals ($\beta = .329, t(147) = 3.236, p < .05$). Confidence was not, however, predicted by threat appraisals ($\beta = -.124, t(147) = -1.371, p > .05$). Tests of the indirect effects linking appraisals to performance were mostly consistent with the separate paths reported above. Namely, the indirect effect from challenge appraisals to math test performance, going first through confidence and then
Table 4. Study 2 measurement model parcel loadings

<table>
<thead>
<tr>
<th></th>
<th>Est</th>
<th>S.E.</th>
<th>Est/S.E.</th>
<th>Std</th>
<th>StdYX</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Face Perception</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parcel 1</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.162</td>
<td>0.713</td>
</tr>
<tr>
<td>Parcel 2</td>
<td>1.225</td>
<td>0.130</td>
<td>9.422</td>
<td>0.198</td>
<td>0.839</td>
</tr>
<tr>
<td>Parcel 3</td>
<td>1.473</td>
<td>0.157</td>
<td>9.401</td>
<td>0.239</td>
<td>0.836</td>
</tr>
<tr>
<td>Parcel 4</td>
<td>1.220</td>
<td>0.131</td>
<td>9.299</td>
<td>0.198</td>
<td>0.826</td>
</tr>
<tr>
<td><strong>Picture Perception</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parcel 1</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.158</td>
<td>0.893</td>
</tr>
<tr>
<td>Parcel 2</td>
<td>0.924</td>
<td>0.064</td>
<td>14.333</td>
<td>0.146</td>
<td>0.858</td>
</tr>
<tr>
<td>Parcel 3</td>
<td>0.873</td>
<td>0.060</td>
<td>14.549</td>
<td>0.138</td>
<td>0.865</td>
</tr>
<tr>
<td>Parcel 4</td>
<td>0.994</td>
<td>0.072</td>
<td>13.757</td>
<td>0.157</td>
<td>0.840</td>
</tr>
<tr>
<td><strong>Challenge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parcel 1</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.504</td>
<td>0.671</td>
</tr>
<tr>
<td>Parcel 2</td>
<td>0.983</td>
<td>0.136</td>
<td>7.211</td>
<td>0.496</td>
<td>0.713</td>
</tr>
<tr>
<td>Parcel 3</td>
<td>1.259</td>
<td>0.167</td>
<td>7.544</td>
<td>0.635</td>
<td>0.759</td>
</tr>
<tr>
<td>Parcel 4</td>
<td>1.398</td>
<td>0.181</td>
<td>7.712</td>
<td>0.705</td>
<td>0.786</td>
</tr>
<tr>
<td><strong>Threat</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parcel 1</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.775</td>
<td>0.835</td>
</tr>
<tr>
<td>Parcel 2</td>
<td>1.017</td>
<td>0.088</td>
<td>11.549</td>
<td>0.789</td>
<td>0.793</td>
</tr>
<tr>
<td>Parcel 3</td>
<td>1.146</td>
<td>0.076</td>
<td>15.033</td>
<td>0.889</td>
<td>0.952</td>
</tr>
<tr>
<td>Parcel 4</td>
<td>1.022</td>
<td>0.083</td>
<td>12.310</td>
<td>0.792</td>
<td>0.826</td>
</tr>
<tr>
<td><strong>Confidence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parcel 1</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.831</td>
<td>0.906</td>
</tr>
<tr>
<td>Parcel 2</td>
<td>1.048</td>
<td>0.059</td>
<td>17.729</td>
<td>0.870</td>
<td>0.909</td>
</tr>
<tr>
<td>Parcel 3</td>
<td>1.117</td>
<td>0.065</td>
<td>17.061</td>
<td>0.928</td>
<td>0.895</td>
</tr>
<tr>
<td>Parcel 4</td>
<td>1.164</td>
<td>0.062</td>
<td>18.700</td>
<td>0.967</td>
<td>0.927</td>
</tr>
</tbody>
</table>
Table 5. Study 2 correlation matrix of latent variables

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Face Perception</td>
<td>.444*</td>
<td>.147</td>
<td>-.082</td>
<td>-.017</td>
<td>-.116</td>
<td>-.015</td>
<td></td>
</tr>
<tr>
<td>2. Picture Perception</td>
<td>.117</td>
<td>-.147</td>
<td>.039</td>
<td>-.151</td>
<td>.056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Challenge</td>
<td>-.392*</td>
<td>.377*</td>
<td>-.197*</td>
<td>.121</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Threat</td>
<td>-.251*</td>
<td>.190*</td>
<td>-.111</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Confidence</td>
<td>-.391*</td>
<td>.246*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Negative Affect</td>
<td></td>
<td></td>
<td>-.175*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Test Performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05
Figure 3. Study 2 structural model with path coefficients

Challenge Appraisal

-0.39*

Threat Appraisal

0.12

Confidence

-0.33*

Negative Affect

-0.40*

Test Performance

-0.18*

*p < .05
through negative affect, was marginally significant, $t(147) = 1.689, p = .093$. The same indirect effect starting with threat appraisals was not significant, $t(147) = -1.127, p > .05$, due clearly to the non-significant relation between threat appraisals and confidence noted above.

In the second step, examining interaction effects between math test appraisals and emotion perception on confidence, two different models were examined: one using the measure of emotion perception based on face stimuli and one based on picture stimuli (see Measurement Properties section above for details). When examining the latent interaction between emotion perception based on faces and threat appraisals, an interaction was not observed. When examining the same interaction for challenge appraisals, however, a moderating effect was detected, $\beta = -1.381, t(147) = -2.208, p < .05$. Consistent with hypothesis 2 and displayed in Figure 4, the interaction revealed that the positive relation between challenge appraisals and confidence was greater for those who scored lower on the emotion perception measure than for those who scored higher on emotion perception. The emotion perception measure using pictures revealed significant interactions neither for challenge nor for threat appraisals. There is evidence in the literature that may help explain why the face perception and picture perception components of the MSCEIT measure exhibited different relationships with emotion perception. Research in cognitive psychology has found that the visual cues of faces are processed more rapidly and automatically than other pictorial stimuli (Bradley, Mogg, & Millar, 2000). Studies have also found that perception of faces is affected differently by the perceiver's emotional state than perception of non-facial stimuli, with facial perception facilitated by positive emotional states and perception of non-facial stimuli
Figure 4. Study 2 confidence predicted by challenge appraisal, moderated by emotion perception
facilitated by negative emotional states (Bate et al., 2009; D'Argembeau et al., 2003; Mickley & Kensinger, 2008).

Study 2: Discussion

This study proposed a causal chain linking cognitive appraisal style to test performance, such that cognitive appraisals lead to situation-specific confidence assessments, which in turn generate certain affective reactions during a testing experience, which have direct impacts on test performance. It was anticipated that a cognitive appraisal style that is rated high on challenge, in which a test-taker perceives a testing situation as a beneficial opportunity to demonstrate their talent or ability, would cause the test-taker to have a relatively high level of confidence in their ability to perform well on a test. It was also anticipated that a cognitive appraisal style that is rated high on threat, in which a test-taker perceives a testing situation as a negative experience which endangers them or their personal goals in some way, would lead to a relatively low level of confidence on the test-taker's part in their ability to perform well on a test. This relation was partially confirmed for the challenge pattern, but it was not confirmed for the threat appraisal pattern.

This disparity between the challenge and threat appraisal patterns may be a reflection of the testing conditions employed in this study, rather than a true reflection of a differential impact of challenge and threat on test-takers' confidence perceptions. As described above, participants were told that the math test was an established measure of intelligence, that their performance on the math test would be used to place them in consideration for scholarships and internship opportunities, that the test administrator would be reviewing their results with them afterward, and that the test was timed. So
although steps were taken to attempt to simulate the psychological environment of a high stakes testing situation as closely as possible in this study, it is possible that a truly threatening environment was not successfully invoked for the test-takers. To the extent that test-takers accepted the invented context of the testing situation that they were presented with as they began the math test, this scenario may have elicited more challenge than threat since the test-taker stood to gain some reward (scholarships, internships) from good performance, but there was nothing to be lost if they performed poorly, aside from some potential social stigma, which would be limited to the test administrator's relatively impartial perspective.

At the next point in the causal chain, it was expected that the test-takers' confidence in their ability to do well on the math test would be related to their real-time emotional reactions to the test itself. This relation was also confirmed, such that test-takers who indicated that they were confident they would perform well on the test were also lower in their ratings of their own anxiety levels while taking the test. As has been found in previous research, these results also confirmed that lower negative affect (i.e. anxiety and frustration) was associated with better performance on the test.

Although evidence for a threat appraisal pattern affecting negative affect and test performance was not found, these results appear to establish that there is a challenge appraisal pattern influencing performance in high stakes testing situations. Furthermore, evidence was found that emotion perception plays a role in this appraisal. There was an interaction between face perception and the challenge appraisal style impacting confidence ratings, such that for those who scored high on face perception their challenge appraisals seemed to have less influence on their confidence perceptions. In other words,
the more emotionally perceptive test-takers seemed better able to separate their sense of confidence in a particular testing situation from their generalized dispositional attitude toward test-taking. This ability to separate their reactions to a specific event from their overall tendencies appeared to give those test-takers with a lower challenge disposition a boost in their confidence ratings, whereas those test-takers with both a low challenge appraisal orientation and low emotion perception tended to have the lowest assessments of their pre-test confidence.

General Discussion

These studies were intended to shed light on the role that a test-taker’s emotional reactions to a test will have on their performance. Specifically, it was initially hypothesized that an emotionally perceptive test-taker would be particularly disadvantaged during a stressful high-stakes testing situation if they were also prone to experiencing test anxiety. It was surmised that the test-taker’s emotional perceptiveness would cause them to be more vulnerable to the negative effects of test anxiety on their performance, being more aware of the worry and distracting cognitions associated with test anxiety than their more emotionally “ignorant” fellow test-takers. While this supposition is not supported by the current data, these studies did serve to expand our understanding of the mechanisms that relate distal emotional reactions to the “idea” of a test to more real-time reactions to an actual testing situation. Furthermore, these studies demonstrated an appraisal pattern affecting test reactions beyond the commonly assessed pattern involving anxious reactions to tests, namely a beneficial impact of test-taking confidence.
In study 1 it was found that the distal assessments of test reactions, specifically worry, tension, test-irrelevant thinking, and bodily symptoms as assessed by the Reactions to Tests Questionnaire, did not significantly impact performance on a vocabulary test. While emotional perceptiveness also had no relations with these distal indicators of test reactions, it was associated with 1) better performance on the test, recalling past research that has pointed to a correlation between ability-based EI and cognitive ability (Lopes et al., 2003; Brackett & Mayer, 2003; Davis & Kraus, 1997); 2) greater confidence in one’s performance; 3) more accurate judgments of one’s performance; and 4) less association between confidence assessments on one item and confidence assessments on the subsequent item. In contrast to the hypothesized negative effect of emotion perception it appears that being emotionally perceptive, rather than exacerbating test anxiety, allows the test-taker to accurately assess their situation and how they are coping with it, focusing attention on relevant stimuli.

These effects of emotion perception on perceptions of performance were explored further in study 2. The second study looked at both distal and proximal indicators of test reactions, and more explicitly examined the role of confidence reactions in addition to anxiety reactions. While a threat appraisal pattern relating anxiety to performance was not found, a challenge appraisal pattern relating confidence to performance was partially confirmed. In the testing situation presented in this study, at least, it appeared that test-takers perceived very little threat from the test but that there was a perception of an opportunity to demonstrate their ability and perhaps earn a reward. There was nothing to be lost but something to be gained from the test, therefore the test-takers’ challenge appraisal pattern rather than their threat appraisal pattern was triggered. This illustrates
the disconnect that can occur between dispositional appraisal tendencies (toward perceiving situations in general, or testing situations in particular, as challenging or threatening) and the actual emotional reactions elicited by a given situation.

As the results of this study also indicate, the disparity between distal cognitive appraisals and the emotions elicited during an actual event are affected not only by the characteristics of the situation but also by traits of the individual. Namely, it appears that the ability of emotional perceptiveness is a factor in this relation. As was suggested by the findings in study 1, the emotionally perceptive test-takers were less prone to base their assessments of their ability to do well when confronted with actual items on an ability test on their tendency to deal with tests as challenges and positive opportunities. The less emotionally perceptive test-taker, on the other hand, was more likely to judge their confidence with regard to a specific test on how they have reacted to tests in the past.

These findings both elaborate on the mechanisms underlying the relations between distal and proximal reactions to a test, as well as adding further research evidence for the benefits of emotional intelligence. As indicated above, the results of the studies presented here make it clear that it is important to consider more than just dispositional tendencies toward certain cognitive appraisal patterns when assessing how non-cognitive factors will influence both performance on a test and how an individual will react to the test itself. The characteristics of the specific testing situation and of the test-taker also play significant roles. Below I discuss some practical implications of these findings, as well as suggestions for further research, with regard to emotional intelligence, goal orientation, and test taking.
Implications for Emotional Intelligence and Emotion Perception

Future investigations may be able to expand upon these results in a number of ways. First of all, there may be alternative methods of capturing real time test reactions that do not rely on the post-test recollections of the test-taker and are also minimally intrusive that could be employed, such as devices that can monitor physiological reactions without being overly distracting to the participant. Considering the testing situations created in these studies, as discussed above in study 1 there were no constraints placed on the situation to make it resemble a true high-stakes testing environment, and in study 2 it appears that the constraints that were employed succeeded in inducing a sense of challenge in test-takers to some extent, but largely failed to present a legitimate threat. However, aside from the self-reported emotional reactions themselves the only manipulation check that was used was the metric of if and when during the debriefing process the participants indicated that they had suspicions about the testing context.

Future research should also consider alternative methods for presenting the testing scenario in order to more effectively activate anxiety reactions in participants, as well as more direct manipulation checks. Perhaps the most parsimonious solution would be to employ students in an actual testing situation as research subjects. Furthermore, the ability measures used in these studies, a vocabulary test and a mathematical problem solving test, are both relatively narrow in their scope. It would be interesting to examine these effects with tests in a variety of content areas, since it is possible that some ability measures may be more impacted by emotional content than others, and therefore if different content areas or problem types are administered at the same time emotional reactions to the measures as well as the role of emotion perception may differ. For
example, mathematical word problems may inadvertently induce emotional reactions due to their content that would not be found with straightforward formula completion problems.

The assessments of emotional intelligence employed in these studies could also be improved upon. The MSCEIT was used in study 2 to build on the findings from study 1, which used an assessment that only tapped into the emotion perception branch of emotional intelligence. It was surmised that the influence of emotional perceptiveness on test reactions, specifically its apparent ability to help the emotionally perceptive test-taker focus on the item at hand and ignore emotional input from both performance on and reactions to prior items, may have not been due entirely to emotion perception itself but to the influence of other aspects of the ability-based model of emotional intelligence, namely emotion management. Whereas emotion perception would not, in and of itself, seem to be capable of imbuing a test-taker with the attentional focus that these participants seemed to possess, that would be an expected outcome if the test-takers were also good emotion managers. And to the extent that ability-based EI represents a unitary construct composed of four interrelated, and correlated, emotion-related skills (Mayer, Salovey, & Caruso, 2000), an emotionally perceptive test-taker would also tend to be a good emotion manager. By using the MSCEIT in study 2, which includes tasks assessing emotion management, it was hoped that analyses to determine the role of emotion management in test anxiety could be carried out; however, the low reliabilities found for the emotion management tasks in this sample precluded any meaningful examination of this possibility. Future research should consider the inclusion of other assessments of emotion management beyond that employed by the MSCEIT.
Joseph and Newman (2010) have proposed a cascading model to describe how three of the four branches of emotional intelligence impact performance sequentially, drawing on a modal model to explain how emotion affects behavior (Gross & Thompson, 2007). According to this modal model, a stimulus must be first attended to, and then appraised, before a response is elicited. Therefore, Joseph and Newman suggest that the impact of emotions on performance outcomes begins with emotion perception. Those individuals who are more emotionally perceptive have access to a broader range of emotional information about both their own states and their environment, and are able to more accurately appraise their situations. It is at the point of cognitive appraisal that emotion understanding intervenes as a mediator of the relationship between emotion perception and emotion management. The cascading model proposes that the ability to manage emotions is the aspect of EI that is most directly related to performance, in that those who are better emotion managers are able to select and utilize more effective emotion regulation strategies that are less draining on cognitive resources than others (Gross, 1998; Butler et al., 2003). The emotion facilitation branch is not included in this model due to the evidence for its redundancy with aspects of the other branches and the lack of support for the existence of this branch in factor analytic models (Gignac, 2005; Palmer et al., 2005; Rossen, Kranzler, & Algina, 2008). Future research based around this sequential model of emotional intelligence would help to shed more light on the mechanisms behind the relations between emotions and outcomes in stressful situations such as high-stakes tests, and help to incorporate the full range of abilities encompassed by the construct of emotional intelligence.
The current studies employed ability-based measures of emotion perception, but it may be worthwhile to examine trait-based indicators of individual differences in processing emotional information in the context of cognitive appraisals, both to further define the distinctions between trait-based and ability-based EI and to reduce the overlap between the measure of EI and the performance measure that is used. To the extent that ability-based EI is related to cognitive ability (Lopes et al., 2003; Brackett & Mayer, 2003), it is difficult to distinguish the impact of EI on test performance from the impact of general cognitive ability. This relationship between ability-based emotional intelligence and cognitive ability may be responsible for the stronger relationship between performance (responding correctly to an item) and confidence, in other words accuracy in judgments of performance, which was found for emotionally perceptive test-takers in study 1. If an emotionally perceptive test-taker is also a more intelligent test-taker in general this may explain why these participants were more aware of whether they had responded correctly or not, beyond any effect that emotion perception may have had to focus their attention. Looking at the interactive effects that trait-based indicators of EI have on test reactions and test performance, in comparison to the effects of ability-based measures, may shed further light on this problem.

**Implications for Goal-Relevant Appraisals**

As discussed above, there is considerable research demonstrating the links between one's appraisals in achievement situations and specific goal-related behaviors and outcomes. From the perspective of the goal orientation literature, mastery goals are associated with adaptive motivation and affect, which in achievement situations manifests as enhanced persistence, deeper information processing, and more positive
affect than goal orientations that are related to external performance criteria (Elliot, McGregor, & Gable, 1999).

Researchers typically have assumed that achievement goals are relatively stable learner characteristics (Dweck & Leggett, 1988). However, recent research has indicated that although achievement goals generally are stable, they can also fluctuate based on the particular characteristics of the achievement situation (Fryer & Elliot, 2007; Senko & Harackiewicz, 2005). Students can engage in both goal intensification, becoming more or less involved in performance tasks, and goal switching, adopting a mastery or performance goal orientation when moving from one task to the next based on their experiences with the prior task (Senko & Harackiewicz, 2005). While the previous research into shifting goal orientations has looked at changes across tasks, the results of the two studies presented here suggest that a test-taker’s perceptions of a test can change over the course of a single testing experience. In particular, Study 1 confirmed that perceptions may shift from one item to the next, as test-takers encounter items of greater or lesser difficulty and become more or less confident of their responses to each item.

This unstable nature of test appraisals has potential implications for many achievement situations. Studies have indicated that performance on a task is improved when students are primed for mastery goals when presented with the task (by telling the students that through persistence they will be able to master the task) as opposed to performance goals (describing the average performance standards of other students; Thompson & Musket, 2005). Presumably, these primes in goal orientation create different appraisal sets (i.e., challenge vs. threat appraisals) which can directly influence the degree of anxiety experienced during the test (Shell & Husman, 2008). To be
maximally effective, this priming may be needed not only at the beginning of a task calling for more endurance, (i.e., as with standardized tests where each item presents a new performance opportunity), but also intermittently during the task to reinforce a mastery goal orientation and to minimize the emotional impact of drops in confidence that can occur.

Furthermore, additional research examining the relationship between emotional perceptiveness and goal appraisals should be conducted to further explore those links. The results found here would suggest that those who are more emotionally perceptive might also be better equipped to establish and maintain challenge and confidence appraisals in achievement situations. As these studies demonstrate, emotional perceptiveness lends itself to a degree of stability over the course of a task which would enable the individual to maintain appraisal pattern more consistently for the duration of a task.

*Implications for Test-Taking*

Interest in the emotions associated with testing has increased in part because of the current interest in standardized achievement tests. It is important that standardized tests accurately reflect students' knowledge and skills, yet it is clear from the research that test anxiety has a consistent negative impact on performance. As a result there are a growing number of researchers interested in emotion regulation, with a focus on understanding how students regulate their emotions so that we can better help them through the testing process (Gross, 1998; Gross, Richards, & John, 2006, Schutz, Distefano, Benson, & Davis, 2004).
The studies presented here add further evidence to the transactional model of situational appraisals discussed above (Sawyer & Hollis-Sawyer, 2005). It is clear that both individual and contextual characteristics influence the way a test-taker will approach and cope with a given testing situation. If the testing environment is carefully designed and instructions are effectively conveyed, the test can be seen as an achievable goal. As noted above, the adoption of a mastery goal orientation is one way to achieve this mindset which can be induced by the test-taker and facilitated within the testing environment. This involves encouraging the test-taker to do well on the test, while carefully avoiding comparisons to others. Solitary as opposed to group testing is one method of directing the test-taker’s focus on their own performance rather than comparisons with others (Hembree, 1988).

Another individual characteristic than could be further explored with relation to the emotional impact of a test-taking environment is locus of control (internality versus externality), which has been suggested to be an important predictor of test anxiety (Choi, 1998; Volkmer & Feather, 1991). A feeling of decreased personal control can heighten feelings of test anxiety and decrease one's willingness to engage and persist in challenging, feedback-oriented performance tasks (Feather & Volkmer, 1988; Lusk, 1983). It is possible that emotionally perceptive test-takers are also more prone to having an internalized locus of control due to their heightened awareness of their own emotional states and the way that those emotional states influence their environments.

Another research area that was not touched on directly in the present research is the potentially beneficial impact of anxiety on performance. In research examining athletic performance in particular, it has been demonstrated that anxiety can have
performance enhancing effects. In Catastrophe Theory, for instance, if cognitive anxiety is low (the subject perceives herself to be up to the challenge in terms of knowledge and ability) but somatic anxiety is high (the subject experiences nervousness or tension) performance is better than if somatic anxiety is also low (Hardy, 1990). Similarly, in Eysenck's attentional control theory (Eysenck et al., 2007), anxiety is assumed to have both detrimental cognitive consequences yet positive motivational consequences. The implication of this is that the alleviation of test-taker anxiety should be approached with caveats – if the test-taker is too relaxed, they may not be fully engaged in the testing experience. By unpacking the respective consequences of these content-driven and emotion-driven aspects of anxiety, we can better map the point of greatest utility for a healthy level of test anxiety.

In conclusion, as any student who has experienced a high-stakes test can confirm, there are factors other than cognitive ability that will contribute to performance outcomes. As the findings reported in these studies demonstrate, these non-cognitive factors are more complicated than a simple negative correlation between anxiety and performance. Emotional reactions associated with a testing situation can be positive or negative, and qualities in both the test and the test-taker will serve to either mitigate or exacerbate these reactions.
References

Ainley, M., Corrigan, M., & Richardson, N. (2005). Students, tasks and emotions:
Identifying the contribution of emotions to students’ reading of popular culture
and popular science texts. Learning and Instruction, 15, 433-447.

Educational Psychology, 84, 261-271.

C.D. Spielberger & P.R. Vagg (Eds.), Test anxiety: Theory, assessment, and


in organizational research. Organizational Research Methods, 1, 45-87.


Multihealth Systems.

Barron, K. E. & Harackiewicz, J. M. (2001). Achievement goals and optimal motivation:
Testing multiple goal models. Journal of Personality & Social Psychology, 80,
706-722.


emotional intelligence construct. *Personality and Individual Differences, 28*, 539-
561.


relation between stress and mental health. *Personality and Individual Differences,
32*, 197-209.

Cooper, R., & Sawaf, A. (1997). *Executive IQ: Emotional intelligence in leadership and
organizations*. New York: Grosset/Putnam.

D'Argembeau, A., Van der Linden, M., Comblain, C. & Etienne, A. M. (2003). The
effects of happy and angry expressions on identity and expression memory for


NY, US: Guilford Press.

validity of the Bar-On Emotional Quotient Inventory (EQ-i) in university

intelligence to predict individual performance, group performance, and group


Elliot, A.J. (1997). Integrating “classic” and “contemporary” approaches to achievement motivation: A hierarchical model of approach and avoidance achievement


*Personality and Individual Differences, 44,* 1258-1269.


Appendix A

Complete List of ETS Measures in Study 1 Data

- Vocabulary Test II
- Reactions to Tests Questionnaire
- Overclaiming Questionnaire
- Emotion Perception in Stories
- Self-perception Index of Emotions
- Emotion Strategies Index
- Affective Quote Completion
- Abbreviated Time Management Index
- Loci-Alpha Scale
- Study Habits Test
- Reactions to Stress at Work
- Metacognitive Awareness Inventory
- Memory and Reasoning Competence Inventory
- Numeracy
- Student Attitudes on Learning and Intelligence
- Theories of Intelligence – Adult Scale
- Attribution Inventory for Students
- Academic Values Inventory
- Job Area Interest Questionnaire
- Mechanisms of Moral Disengagement
- General Health Questionnaire
- International Personality Item Pool Big Five Measure
- Values Survey Module
- Survey of Dictionary Isms
- Values Survey
- Social Judgment Scale
Appendix B

Measures Used in Study 1

Example Vignette from Emotion Perception in Stories (EPS)

DIRECTIONS:

Below is a series of narratives written by students. After reading each one, your task is to decide what emotion you perceive the person was feeling while writing the passage. A number of emotions are given. Rate each one along the scale provided.

(1) I used to be very angry at my ex-boyfriend's mother. I always felt that she resented me for taking up all of John's time. She was the type of woman who was very domineering and considered her son to be her little pride and joy that she should shelter at all times. As John dated me, he became more independent and did not rely on his mother to make his decisions for him. When John would do something wrong, the blame was always put on my shoulder. Several incidents occurred where she yelled at me. Needless to say, I was very mad. I talked behind her back, bad mouthing her name as well as saying things to her face. Now that I look back on my encounters with her, I now have a better understanding of the situation. Many factors could have been involved that I wasn't aware of. After praying about my feelings towards her, I am finally content. It was wrong for me to judge her or.

What was the author feeling while writing the passage? Rate each emotion on the scale given below.

Happiness
1 2 3 4 5
Not at all Very little Somewhat Strongly Very strongly

Sadness
1 2 3 4 5
Not at all Very little Somewhat Strongly Very strongly

Contempt (dislike, disdain)
1 2 3 4 5
Not at all Very little Somewhat Strongly Very strongly

Disgust
1 2 3 4 5
Not at all Very little Somewhat Strongly Very strongly

Anger
1 2 3 4 5
<table>
<thead>
<tr>
<th></th>
<th>Not at all</th>
<th>Very little</th>
<th>Somewhat</th>
<th>Strongly</th>
<th>Very strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surprise</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not at all</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not at all</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix C

Study 2 Script

*Before the confidence measure...*

This first assessment is going to present you with some mathematical problems and ask you to rate how well you think you would do if asked to respond to the questions it presents you with. You don’t need to actually answer the math problems, just respond to the questions asking you to rate your confidence that you could correctly solve the problem. When you have completed these questions, come get me in the next room, and don’t move on to the next part of the study yet.

*Before the ability test...*

In this test you will be asked to solve some problems in mathematics. This is a test of intelligence that’s routinely used for selection purposes by undergraduate and graduate university programs, including MBA programs, and employers in all job sectors. Your test will be scored by the computer and compared to national averages. Your score from this test will be reported to the university administration to assist in identifying students for certain scholarship and internship opportunities. Test-takers who perform well on this assessment increase their chances of being awarded one of these scholarships or internships. I’ll review the results of your test with you at the end of the experiment.

This is a timed test – you’ll have 12 minutes. There are 15 questions, and you’ll see one per screen. Solve each problem and indicate the best answer. I’ll start the timer once you click to start, and come back in 12 minutes to move you on to the next part of the study.
Before the post-test...

In this assessment, you'll be presented again with the same problems you just saw, but this time you're not being asked to solve the problems but to think back to when you first saw each question on the test and evaluate your immediate reaction to it. For each problem, you'll be asked to rate the extent to which you experienced four emotional reactions. If you skipped over a problem on the test, or didn't get to some problems, skip those problems here as well. When you've finished, leave the survey open on the screen and come join me in the next room.

Study 2 Debriefing Protocol

When the subject returns to the debriefing room, tell them to wait there while you go over to the study room and close out their session. This is when you'll begin buffering their video. Then return to the debriefing room.

The purpose of today's study was to examine factors, other than one's actual ability, that may influence test scores. There are many possible reasons why people perform well or perform poorly on a test, regardless of how much they actually know about the test material. Today's study involved two main factors: how confidently you approach exams and how you experience and perceive your own emotional reactions to exams. Different people react differently to test-taking environments, and experiencing anxiety is one of the most common reactions that interferes with a person's ability to take a test.
Therefore, we needed to get a good idea both of whether you feel threatened by tests generally, as well as whether actually taking a test creates any anxiety.

As you might imagine, measuring anxiety while you are taking a test is not an easy thing to do. Some of the best markers of stress and anxiety are the various facial and other non-verbal movements that people express in stressful situations. So, in order to get an accurate picture of what emotions you were experiencing during the test, we wanted to try to get the best measure possible – your facial and non-verbal expressions during the test. There is also a lot of evidence that people will hold back these expressions if they know they are being observed. So, in order to capture your spontaneous reactions, we videotaped you while you were taking the test using the computer's built-in camera. We should tell you right away that this video will only be seen by the research team (which includes me, two faculty members, and a few other research assistants from a different university, who will help code the tapes). If you would prefer us not to use the video, we will gladly go ahead and delete the file right here and now.

After making sure that the participant understands the purpose of the videotaping, give them the Digital Video Release Form. After the participant has signed this, continue with the debriefing, explaining the deception involving the use of test scores.

Another thing that's often difficult with studies like these is that if we simply asked people to take these tests, they would not be nearly as anxious or motivated as they might be during an actual high-stakes testing situation. [Pause here to allow the Participant to comment or ask a question if desired] Just like with the videotape, if you knew that the
test scores didn't really matter, then you probably would not try as hard, much less get anxious or stressed about taking the test. Does that make sense? [Pause here to answer any questions. Anticipate that some of the participants will at this time guess that the test will not be used for scholarship/internship purposes. If not, continue explanation] The goal in this study was to create an environment that was as close as possible to what people would experience when they actually sit down to take a test like the SAT or the GRE. The main reason that people become anxious or stressed with tests like these is because these tests usually have very important consequences... you are compared to other people and the "high-stakes" often involve entry into desired schools or jobs. So, we created this environment by trying to convince you that the test really was important and that your performance would matter. [Pause one more time to allow the Participant to comment or ask questions. Most participants will have realized the true nature of the study by this point. If there is no evidence that they have, then at this point, tell them...] To increase the importance of the test, we told you that we would be comparing your scores to others' and that your scores could qualify you for scholarships or internship positions. In truth, we are not working in collaboration with a University committee; the scores will only be used for the purposes of the study and will not be reported to anyone else outside of the research team. In fact, no identifying information will even be connected to your data, and we do not really have enough information to provide you with accurate score reports. [At this point, if the Participant expresses any discontent with the deception, the Experimenter will immediately follow up with additional information...] It's important to know that the only reason we did this was to try to understand what happens when people take high-stakes tests. Our hope with this study is
to uncover some of the reasons why these tests measure "test-taking ability" every bit as much as they do actual mathematics knowledge. If we can identify these problems in the test taking environment, then we can recommend new procedures that help make the tests far better at doing what they are supposed to do. Unfortunately, the only way to capture that is to create a sense that the test really does have "high-stakes".

*Finally, to help ensure that the integrity of the study's procedures remains intact, explain the following to each participant:*

Now that you know the full extent of the study, I need to ask a very important favor. As you might imagine, if people knew coming in that the test was just for the study and that their reactions were video-taped, they probably would behave differently than they would otherwise. We are very interested in capturing *natural, spontaneous* reactions to high-stakes testing situations, so we would appreciate it if you would not tell anyone about these parts of the study. If someone asks you what the study was about, you can simply tell them that you completed some questionnaires and took a math test. Can we count on you for this?