THE EFFECTS OF DIFFERENTIAL DISCRIMINATION CUES ON
ATTRIBUTIONS FOR FAILURE:
IMPLICATIONS FOR SUBSEQUENT PERFORMANCE

A thesis presented to
the faculty of
the College of Arts and Sciences of Ohio University

In partial fulfillment
of the requirements for the degree
Master of Science

Anna Berlin
August 2006
This thesis entitled
THE EFFECTS OF DIFFERENTIAL DISCRIMINATION CUES ON
ATTRIBUTIONS FOR FAILURE:
IMPLICATIONS FOR SUBSEQUENT PERFORMANCE

by

ANNA BERLIN

has been approved for
the Department of Psychology
and the College of Arts and Sciences by

Keith D. Markman
Assistant Professor of Psychology

Benjamin M. Ogles
Dean, College of Arts and Sciences
Abstract

BERLIN, ANNA, M. S., August 2006, Psychology

THE EFFECTS OF DIFFERENTIAL DISCRIMINATION CUES ON ATTRIBUTIONS FOR FAILURE: IMPLICATIONS FOR SUBSEQUENT PERFORMANCE (75 pp.)

Director of Thesis: Keith D. Markman

A single study investigated the effects of discrimination on performance. One hundred and three female undergraduates completed an initial anagram test which they believed to be a scholastic aptitude test. They were then given either an overt, ambiguous, or no cue regarding the likelihood that the evaluator discriminated against women. They were later told that a second session of a similar task would be evaluated by the same or a different evaluator. All participants were informed that they had performed poorly on the first task. Results indicated that although the women tended to attribute the failure to external factors (discrimination) rather than internal ones (ability and effort) as discrimination cues increased in strength, only those who received ambiguous discrimination cues did not improve performance from the first to the second task. Implications for the effects of ambiguity on performance are discussed.

Approved:

Keith D. Markman

Assistant Professor of Psychology
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>3</td>
</tr>
<tr>
<td>List of Tables</td>
<td>7</td>
</tr>
<tr>
<td>List of Figures</td>
<td>8</td>
</tr>
<tr>
<td>Overview</td>
<td>9</td>
</tr>
<tr>
<td>Introduction</td>
<td>10</td>
</tr>
<tr>
<td>Stigmatization and Attributional Ambiguity</td>
<td>11</td>
</tr>
<tr>
<td>The Discounting Principle</td>
<td>12</td>
</tr>
<tr>
<td>Situational Ambiguity</td>
<td>16</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>20</td>
</tr>
<tr>
<td>Self-Assessment</td>
<td>21</td>
</tr>
<tr>
<td>Causal Uncertainty</td>
<td>21</td>
</tr>
<tr>
<td>Uncertainty and Stigmatization</td>
<td>22</td>
</tr>
<tr>
<td>Feedback and Subsequent Performance</td>
<td>23</td>
</tr>
<tr>
<td>Feedback Intervention Theory</td>
<td>23</td>
</tr>
<tr>
<td>Achievement Motivation</td>
<td>25</td>
</tr>
<tr>
<td>Implications of Attributing Failure to Discrimination on Subsequent Performance</td>
<td>25</td>
</tr>
<tr>
<td>Goals of the Present Research</td>
<td>28</td>
</tr>
<tr>
<td>Study Overview</td>
<td>28</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>30</td>
</tr>
<tr>
<td>Preliminary Hypotheses</td>
<td>30</td>
</tr>
<tr>
<td>Preliminary Hypothesis 1</td>
<td>30</td>
</tr>
<tr>
<td>Preliminary Hypothesis 2</td>
<td>30</td>
</tr>
<tr>
<td>Major Experimental Hypotheses</td>
<td>30</td>
</tr>
<tr>
<td>Hypothesis 1</td>
<td>31</td>
</tr>
<tr>
<td>Hypothesis 2</td>
<td>31</td>
</tr>
<tr>
<td>Individual Differences</td>
<td>33</td>
</tr>
<tr>
<td>Implicit Theories of Intelligence</td>
<td>33</td>
</tr>
<tr>
<td>Causal uncertainty</td>
<td>33</td>
</tr>
<tr>
<td>Achievement motivation</td>
<td>34</td>
</tr>
<tr>
<td>Method</td>
<td>34</td>
</tr>
<tr>
<td>Participants</td>
<td>34</td>
</tr>
<tr>
<td>Materials</td>
<td>34</td>
</tr>
<tr>
<td>Design and Procedure</td>
<td>35</td>
</tr>
<tr>
<td>Discrimination cues</td>
<td>37</td>
</tr>
<tr>
<td>Second Anagram Session</td>
<td>37</td>
</tr>
<tr>
<td>Attributional assessment</td>
<td>38</td>
</tr>
<tr>
<td>Results</td>
<td>39</td>
</tr>
<tr>
<td>Preliminary Hypotheses</td>
<td>39</td>
</tr>
<tr>
<td>Preliminary Hypothesis 1</td>
<td>39</td>
</tr>
</tbody>
</table>
Preliminary Hypothesis 2 ........................................................................................................ 41
Major Experimental Hypotheses .......................................................................................... 41
Performance ....................................................................................................................... 41
Persistence .......................................................................................................................... 45
Discussion .......................................................................................................................... 46
Comparison to previous studies ........................................................................................ 46
Performance ....................................................................................................................... 48
Future directions .................................................................................................................. 51
Conclusion ........................................................................................................................... 53
References .......................................................................................................................... 54
Appendices .......................................................................................................................... 62
Appendix A: Pretest Measures .......................................................................................... 62
Appendix B: In-Session Computer-Based Questionnaires .................................................. 66
Appendix C: Anagrams ......................................................................................................... 68
Appendix D: Failure Feedback ............................................................................................ 69
Appendix E: Attributions for Performance Outcome .......................................................... 70
Appendix F: Session Protocol ............................................................................................ 72
List of Tables

- Table 1. Means and $sds$ of Attributions as a Function of Cue and Evaluator..................... 40
- Table 2. Means and $sds$ of Persistence and Performance Differences............................. 43
List of Figures

Page

Figure 1. Mean Performance Difference for High and Low Desire for Control ............ 45
Overview

Research on how individuals react to discrimination has recently begun to focus on accurately identifying when discrimination has occurred. Much attention has been paid to the counterintuitive finding that accurately attributing a negative outcome to discrimination may have a buffering effect on self-esteem. Specifically, attributing a failure to an external cause such as discrimination may allow the individual to avoid the temporary decrease in self-esteem that is generally associated with attributing a failure to stable, internal factors.

Because of the possible benefits to psychological well-being, research has often focused on factors that lead to accurately identifying discrimination. Little attention, however, has been paid to the potential long term consequences of such attributions. When individuals receive feedback about their performance it gives them ideas for possible changes needed to improve future performance. Research in the area of feedback also suggests that effort and performance tend to increase when an individual fails on a personally relevant task. Attributing a failure to discrimination may protect self-esteem in the short term. However, it may also render the feedback uninformative and consequently circumvent the typical increase in performance. Even if self-esteem is protected in the short run, a lack of improvement on future performance may lead to a long-term deficit in efficacy-based self-esteem.

In order to investigate these possibilities, the current proposal will give brief reviews of the attribution and stigmatization literatures, followed by a more detailed review of the literature in the domain of attributions to discrimination. A brief overview of feedback intervention theory will also be provided. Finally, there will be a presentation
of a single study which investigates the role of attributions to discrimination on subsequent performance.

Introduction

What happened the last time you experienced a failure? How did you feel after thinking about why the failure occurred? What did you do about it? Most people would agree that they have experienced a failure at one point or another during their life. One common reaction is to attempt to identify the cause of the failure. Internal attributions identify personal shortcomings and behaviors that may have led to the failure, whereas external attributions identify sources outside of the person (Heider, 1958). External attributions include such statements as, “The test was too difficult,” “There was a distraction,” or “The equipment malfunctioned.” Attributions of all varieties are made as a means of making sense of one’s social world (Kelly, 1971).

Failure attributions have a significant impact on subsequent emotions. When internal attributions are made, negative feelings such as guilt, regret or feelings of incompetence often arise (Weiner, 1985). In turn, negative emotions may motivate the individual to make adjustments to subsequent behavior. Identifying lack of effort as the cause of failure may lead one to try harder next time. External attributions, on the other hand, do not tend to elicit the same degree of negative feelings about the failure because such attributions tend to deflect blame from the self (Major, Kaiser & McCoy, 2003). Essentially, making an external attribution is akin to saying “It’s not my fault.” Thus, although changes may still be made to improve subsequent performance, a lack of negative emotion may result in less motivation to do so.
In addition, internal attributions can identify specific, controllable actions that one may have an opportunity to change in the future, whereas external attributions may not always identify potentially controllable actions. For example, a student who attributes a failure on a test to insufficient time spent studying knows what to do in order to improve next time, but a student who attributes failure to the fact that the test was written poorly may feel that there is little that can be done to improve future outcomes in that class.

Individuals may identify a number of potential causes for any given outcome. However, for some members of society, it may be particularly difficult to ascertain the extent to which their outcomes are due to internal or external factors; members of stigmatized minorities may frequently encounter situations in which they are unclear about the degree to which their stigmatized status influences the nature of social interactions.

**Stigmatization and Attributional Ambiguity**

Stigmatization, according to Crocker, Major, and Steele (1998) “…is considered to be primarily a situational threat, the predicament of being in a situation where one’s stigma could influence how one is treated and judged… stigmatized individuals possess (or are believed to possess) some attribute, or characteristic, that conveys a social identity that is devalued in a particular social context” (p. 504). As such, stigmatized individuals are frequently put into positions where they must use situational cues to determine the extent to which their stigma is affecting their interactions with others.

Stigma is, essentially, a social construct. Archer (1985) noted that the accepted rules about which characteristics place an individual in a stigmatized category are defined by society or the government, but not by nature. Although nature may assign specific traits, it is society that devalues, or stigmatizes, those traits. Members of a society are
aware of the content of stereotypes held by that society regardless of their endorsement of them (Devine, 1989). Chronic awareness of stigmatized status and stereotype content may leave individuals in a state of ambiguity regarding the extent to which their stigma is either positively or negatively affecting their outcomes.

The difficulties that face the stigmatized are as varied as the types of stigmatization within a society. Goffman (1963) divided stigma into three major categories: tribal stigma, such as racial and ethnic characteristics; “abominations of the body” such as mental or physical disabilities and obesity; and “blemishes of character,” which include stigmas about substance abusers, criminals and homosexuals. Stigmatized status varies in terms of visibility and perceived controllability (for a review of more distinct dimensions see Jones, Farina, Markus, Miller, & Scott, 1984). People with highly visible stigmas may not be able to conceal them and, relative to people with less visible stigmatized traits, may more often question the extent to which they are receiving differential treatment because of their stigma.

The Discounting Principle

Kelly (1971) posited that, “the role of a given cause in producing a given effect is discounted if other plausible causes are also present” (p. 8). This phenomenon is referred to as the discounting principle. When an individual can identify many potential explanations for another’s behavior, it may be difficult to infer the other’s actual motivation. Initial investigations of the discounting principle examined how people made attributions for another’s behavior (e.g., Thibaut & Riecken, 1955; Jones, Davis, and Gergen, 1961). However, the same concept can be applied to other causal attributions, such as explanations for one’s own performance outcomes. Identifying the cause of an
outcome plays an important role in how the events that surround the outcome are perceived. This may be especially true in the context of discrimination. Discrimination is the behavior that results from the prejudicial attitudes, and stereotypes associated with stigmatized traits (Crocker et al. 1989). Overt expressions of prejudice and discrimination are not socially acceptable. As a result, discrimination often occurs in subtle ways. Consequently, stigmatized individuals can rarely be certain whether or not discrimination was the primary cause of a given negative outcome. Nonetheless, the ability to determine the extent to which discrimination has affected an outcome has important psychological consequences.

Accurate identification of discrimination is a critical step in the attribution process. Research has shown that accurately attributing failure to discrimination enables individuals to buffer themselves against decreases in self-esteem (Crocker & Major, 1989; see also Crocker, Voelkl, Testa, & Major, 1991; Ruggiero & Taylor, 1995, 1997; Major, Kaiser, & McCoy, 2003, Major, Quinton, & Schmader, 2003). Those who do not make attributions to discrimination are more likely to suffer decreases in self-esteem following a failure. Pinpointing discrimination as the primary cause of negative feedback provides the attributor with a reasonable alternative to assigning fault to internal factors. Whereas, individuals who do not attribute negative outcomes to discrimination are more likely to identify internal causes and thereby suffer temporary decreases in self-esteem (Crocker & Major, 1989).

Application of the discounting principle to stigmatization. One of the first studies to apply the discounting principle to a discrimination context was conducted by Crocker et al. (1991), who studied the effects of attributions to discrimination on subsequent
ratings of affect and self-esteem. In two studies, the authors exposed both women (Study 1) and black students (Study 2) to ambiguous cues regarding the likelihood that discrimination had played a role in their positive or negative outcomes. The women in study one learned that their (fictitious) male partner endorsed either liberal or extremely conservative views of women’s role in society. The women then wrote persuasive arguments their male partner was to evaluate for clarity, quality and persuasiveness. Their partner’s feedback on the essay was either clearly positive or clearly negative. After receiving the feedback, participants completed mood, self-esteem, and attribution measures. Following positive feedback, the evaluator’s attitudes toward women had no effect on attributions to discrimination. However, following negative feedback participants were more likely to attribute their results to the evaluator’s attitudes towards women if he had previously expressed conservative attitudes about women’s role in society. Moreover, women who believed that their evaluator held traditional opinions toward women also exhibited less depressed affect than did those women who believed that their negative evaluation came from a more liberal male evaluator. Similar results were obtained in Study 2 by placing African American students in a room where the blinds were either open (race known) or closed (race unknown) to the room where the while evaluator was seated. Although not all effects were statistically significant, the combined results of studies one involving women, and a follow up study involving African Americans, suggest that self-esteem decreased less following failure when participants believed the evaluator was biased against them than when they believed that bias had played no part.
The Crocker et al. (1991) work established a general methodological paradigm for studies examining attributions to discrimination. Studies conducted employing this general format have examined moderating variables such as knowledge that the evaluator is a member of a group that is prototypically prejudiced (Morera, Dupont, Leyens, & Désert, 2004; Inman, 2001), receipt of verbal cues about the likelihood that the evaluator will discriminate (Ruggiero & Taylor, 1995; Major et al, 2003), and exposure to prejudiced comments written by the evaluator (Crocker et al, 1991, Study 2; Sechrist, Swim, & Stangor, 2004). The discrimination cues generally range in strength from subtle to overt, and often include a control group who receives no mention of discrimination. With few exceptions (e.g., Crocker et al.; 1991, Inman, 2001; Sechrist et al., 2004), most studies present only negative feedback given by an unknown evaluator, and focus on the perception of the individual experiencing discrimination rather than observations of discrimination experienced by similar others.

Generally, these studies have found that those who make attributions to discrimination subsequently exhibit little change in state self-esteem after receiving negative feedback. Participants who are given overt discrimination cues generally attribute their failure to discrimination and are therefore the least likely to exhibit decreases in self-esteem. Conversely, those who are not provided cues regarding the likelihood of discrimination tend to make primarily internal attributions and generally exhibit decreases in self-esteem. Subtle or ambiguous discrimination cues have been shown to elicit differential responses dependent upon personality and situational factors.
Situational Ambiguity

One of the key contextual elements in the discounting model is situational ambiguity. When their stigmatized status is concealed, participants rarely make attributions to discrimination. Conversely, when their stigmatized status is not concealed and cues are overt, participants tend to attribute outcomes to discrimination. However, when cues are ambiguous, participants respond differentially to the possibility that they may have been the target of discrimination. In everyday life, individuals rarely encounter completely unambiguous feedback. In fact, Ruggiero and Taylor (1995) found that although most females would agree that women in the United States are often discriminated against, most would also assert that they have not personally experienced discrimination on the basis of their gender. There are numerous explanations for why people who are aware of their membership in stigmatized groups may choose not to believe they have personally experienced discrimination (see Crosby, 1984). A number of recent studies (e.g., Ruggiero & Taylor, 1997; Major et al., 2003; Sechrist et al., 2004; Berlin & Markman, 2006) have focused on both situational and personality variables that impact the likelihood that an individual will make an attribution to discrimination.

Control. Drawing external attributions is in some ways akin to believing that one has little control over an outcome. In fact, Ruggiero and Taylor (1997) demonstrated that individuals who made attributions to discrimination subsequently indicated lower perceptions of domain-specific control than those who did not. Perceived performance control was measured by questions such as, “How good are you at tests such as these?” and “How much would you improve with practice?.” Responses indicating expectations of improvement were considered to reflect perceptions of control over performance and
outcomes. Attributions to discrimination were negatively associated with perceived control over the outcome. The authors concluded that participants were motivated not to make attributions to discrimination because, by doing so, they were acknowledging a lack of control over their personal outcomes. Although self-esteem was higher in participants who made attributions to discrimination, the authors argue that the trade-off may have come at too high a cost in perceived control.

While some research suggests that people minimize attributions to discrimination in order to maintain control (Ruggiero & Taylor, 1997), other research indicates that certain types of people may actually make attributions to discrimination in order to regain control. Sechrist et al., (2004) pointed out that making attributions to discrimination often has social consequences. Those who claim to have been discriminated against may feel they will be disliked by others as a result. Consequently, such attributions come at a higher cost under public reporting conditions than when made privately. Sechrist et al. hypothesized that individuals higher in desire for control would be more likely to make attributions to discrimination¹ under public reporting conditions than would those low in desire for control, “because the benefit of regaining personal control would counteract the cost of being disliked ” (p. 112). Consistent with this reasoning, the authors found that in public reporting conditions, participants who scored high on Burger and Cooper’s (1979) Desire for Control (DFC) Scale were more likely to make attributions to discrimination than were their low DFC counterparts. Overall then, the level of desire for control moderates one’s tendency to draw attributions to discrimination given particular situational cues.

¹ Attributions to discrimination were measured by five questions on a 7-point Likert type scale. Participants indicated their agreement about how much their feedback was due to: ability, effort, response quality on the part of the participant, and bias and discrimination on the part of the evaluator.
The relationship between control and perceptions of discrimination, at first glance seems somewhat ambiguous. In one instance, individuals feel less control over a specific outcome which may lead them to reduce attributions to discrimination. In another instance, individuals who are higher in desire for control are more likely to draw attributions to discrimination in public reporting conditions. Part of the discrepancy between the findings of Ruggiero and Taylor and those of Sechrist et al., lies in the construct definition of control. Ruggiero and Taylor were primarily concerned with domain-specific control and asked questions regarding the control individuals felt over one particular outcome. Sechrist et al. conceptualized control in more global terms and referred to a more general, dispositional tendency to seek control in everyday interactions.

*Mood.* One of the more subtle cues that people attend to is their own mood. Sechrist, Swim, and Mark (2003) found that when provided with no alternative explanation for their mood state, participants in positive moods were less likely to make attributions to discrimination than were those in negative moods. When people were aware of external explanations for their mood, however, there were no differences in the likelihood of attributions to discrimination. The authors suggested that negative moods enhance self-protection needs. Moreover, encounters with discrimination are likely to be affectively charged, and neutral moods may be less likely than valenced moods to inform the individual that there is cause to suspect unfair treatment.

*Gender Identification.* Another factor that affects the likelihood that a person will draw attributions to discrimination is the extent to which they identify with the target group. Major et al. (2003) found that women who scored high on a gender identification
scale were more likely to make attributions to discrimination when ambiguous discrimination cues were present than were those who scored lower on the gender identification scale. Participants in their study completed a gender identification scale adapted from Luhtanen and Crocker’s (1992) collective self-esteem scale. In control conditions where there was no mention of discrimination, no differences were found between high and low gender identification participants with regard to attributions to discrimination. Moreover, in the overt cue condition, all participants, regardless of strength of gender identification, tended to make attributions to discrimination. When discrimination cues were ambiguous, however, high gender identification participants made more attributions to discrimination than did participants low in gender identification. Regardless of gender identification, participants in the overt cue condition maintained self-esteem relative to those in the ambiguous and no mention groups.

In general, those individuals who strongly identify themselves as women may be more likely to believe that other individuals also consider gender to be a useful variable to take under consideration when evaluating performance. It is therefore plausible for such individuals to expect that others will treat them differently on the basis of their gender.

*Summary of Attributional Ambiguity Literature.* The preceding paragraphs provide a detailed account of specific studies that have investigated factors that contribute to the likelihood of drawing attributions to discrimination. The overall findings of such research are that given strong enough cues, most individuals will make attributions to discrimination, and doing so often results in the protection of self-esteem. These benefits, however, might come at the cost of perceptions of control. In general, individuals who
strongly identify with their in-group, or who are in negative moods may be more likely to attribute an ambiguous situation to discrimination.

**Uncertainty**

It is with good reason that accurate identification of discrimination has been one of the primary focuses of recent research. The very nature of stigma leads to high levels of ambiguity in social interactions. Overt displays of prejudice and discrimination are considered to be socially unacceptable. For this reason, most of the discrimination that members of stigmatized groups face occurs subtly. Although individuals may suspect that discrimination has occurred, it is often difficult to be certain. Attributional ambiguity is an inherent part of life for members of stigmatized minorities. As such, these individuals encounter a degree of uncertainty regarding the extent to which their outcomes are affected by their stigma. The resolution of uncertainty is a much-studied topic in social psychology. Social comparison theory (Festinger, 1954) posits that individuals compare themselves to others as a means of gaining information about themselves. Social comparisons are more likely to occur when an individual is high rather than low in uncertainty (Goethals & Darley, 1977). Uncertainty motivates individuals to seek information via causal attributions (Heider 1958; Kelley 1967, 1973) and also leads individuals to seek diagnostic information (Trope & Ben-Yair, 1982). The following sections provide a brief overview of the ways in which individuals react to and seek to reduce uncertainty.
Self-Assessment

According to Trope’s (1979, 1980, 1982) self-assessment model, it is important to accurately assess of one’s own abilities because it renders future outcomes more predictable and controllable. Achievement settings that provide critical and objective feedback about ability level, or that are diagnostic of ability, are important because they allow an individual to glean “ability-relevant information.” When given ambiguous performance feedback in one domain and unambiguous feedback in another, individuals tend to choose to repeat the domain in which feedback was ambiguous. This is usually the case whenever the individual believes that further trials will reduce uncertainty regarding ability level. There is also an inverse relationship between task diagnosticity and persistence; on tasks where sufficient information about ability level is provided after a few trials, participants persist less. Less diagnostic task feedback elicits longer task persistence because under such conditions more time is required in order to ascertain skill level.

Causal Uncertainty

Uncertainty is often a product of an individual’s situation. According to Weary and colleagues (Weary & Edwards 1994, 1996; Weary & Jacobson, 1997) some individuals are more prone to feelings of uncertainty than others. These chronic individual differences comprise the construct of causal uncertainty. Causal uncertainty beliefs are defined as “generalized self-constructs about one’s uncertain or inadequate understanding or detection of causal relations to the social world” (Weary & Edwards, 1996, p. 159). Given the proper circumstances, individuals who are high in causal
uncertainty are more likely to seek diagnostic information than are low causal uncertainty individuals.

*Uncertainty and Stigmatization*

The discrimination context is inherently ambiguous enough that many possible attributions can often be made for any single outcome. Stigma carries with it the property of chronic uncertainty regarding the degree to which any outcome is caused by “typical” internal and external factors as opposed to stigmatization. Moreover, because many stigmatizing characteristics are permanent, there is a great risk associated with consistently making attributions to stigma (Crocker & Major, 1989). When one attributes negative outcomes to controllable factors, one has an idea about how to influence future outcomes. However, when attributions are made to discrimination, one is relatively powerless to make changes that will prevent such negative outcomes in the future. Consequently, consistent attributions to discrimination may lead to feelings of helplessness and despair. Given the potential negative consequences of discrimination, recent research has begun to focus on the counterintuitive finding that attributions to discrimination can lead to self-esteem maintenance. Yet, it is also important to consider the effects of drawing discrimination attributions on subsequent behavior. If the same attribution that protects self-esteem also inhibits the individual from focusing on controllable actions that might ameliorate future performance, the long-term consequences of discrimination attributions may not be worth the short-term benefits.

When an individual discounts negative feedback, the immediate effects may be self-esteem protective, but the long-term effects are likely to lead to a lack of improvement relative to those who did not discount the negative feedback. In order to
assess such differences, it is first important to understand how people react in general to valenced feedback. Previous research has repeatedly demonstrated that following failure those who make attributions to discrimination are less likely to exhibit decreases in self-esteem than those who do not make such attributions. Yet, little is known about the differential effects of attributions on subsequent performance and behavior.

*Feedback and Subsequent Performance*

Before considering how discounting feedback affects subsequent performance, it is first necessary to understand the role of feedback under normal circumstances. The following section will take a brief detour from the literature on stigma and attributions to discrimination in order to provide a basic review of the literature on feedback and performance. The previous section on self-assessment highlighted the way in which individuals use feedback to gain diagnostic information about their ability level, as well as how performance is affected by a lack of diagnosticity. The next section, on feedback intervention theory, will focus specifically on how individuals tend to respond to negative feedback. It further explains research on self-assessment by describing how individuals use feedback to evaluate themselves relative to ideal performance standards and by suggesting processes by which individuals reach for or reject those standards. A final section will discuss personality differences in achievement motivation before returning to the domain of stigma and attributions to discrimination.

*Feedback Intervention Theory*

Kluger and DeNisi (1996) proposed a preliminary feedback intervention theory (FIT) that explains the likely effects of feedback under normal circumstances. One of the key principles of FIT is that people compare their feedback to where they would ideally
like to be (a standard), and make adjustments accordingly (see also Carver & Scheier, 1982). When feedback indicates that actual performance is below the standard (failure feedback), FIT holds that there are four basic strategies that are typically employed. Feedback-standard gap eliminating strategies include: 1) increasing effort (generally occurs when there is a high perceived likelihood of achieving the standard); 2) abandoning the standard (generally occurs with extreme or repeated negative feedback, as is the case with learned helplessness paradigms); 3) changing the standard (e.g., changing from a performance goal to a learning goal [see Senko & Harackiewicz, 2005; Dweck, 1999]); and 4) rejecting the feedback. Although much research needs to be done to determine when each of the four strategies is likely to be employed, FIT does offer some rudimentary predictions. When feedback is positive, indicating that performance is above the standard, subsequent effort depends upon the perceived opportunity to attain other self-related goals. Specifically, if there is a perceived opportunity to reach higher self-related goals, the likely consequence is a raise in standards that will, in turn, engender increased effort and performance. However, when no perceived opportunity exists, effort is likely to be reduced. When feedback is negative, on the other hand, the most likely response is to increase effort. As long as the individual believes that increased effort will serve to decrease the performance-standard gap, effort is likely to be increased until the gap is eliminated, at which point effort will plateau. In general, increased effort leads to improved performance and is therefore likely to be the default reaction unless the feedback is extremely negative. If, however, increased effort does not seem to reduce the feedback-standard gap, attention is likely to be shifted elsewhere. If success seems possible, attention is likely to be directed toward learning. Alternatively, if success seems
unattainable, attention will then be shifted to the self. Attention to the self may or may not produce improved performance, depending on the relevance of the task to other self-related goals and the amount of attention required.

Achievement Motivation

In addition to cognitive processes and self-relevant goals, personality differences may also influence the way an individual responds to failure feedback. People high in achievement motivation tend to value competence, expect success, and seek challenges (Atkinson, 1957, McClelland, 1961), whereas those low in achievement motivation tend to expect failure and avoid meaningful challenges. Those high in achievement motivation are likely to respond to a failure with increased effort, while people low in achievement motivation are less likely to do so (Kukla, 1972; Kuhl, 1978; Revelle & Michaels, 1976).

Implications of Attributing Failure to Discrimination on Subsequent Performance

The nature of the discrimination context may play a crucial role in shaping how an individual responds to failure feedback. Discrimination may be perceived in a variety of situations, sometimes they may be circumscribed, isolated events, whereas other times, such as in a job context, they may reflect ongoing discrimination.

Up to this point, research on discounting has applied to discrimination contexts that contain fairly circumscribed, “one-time” events. Discounting is arguably advantageous in such situations because it allows the individual to disregard the event and its outcome entirely. However, in occupational settings, or other contexts in which discrimination persists across time, discounting may not be so advantageous. Disregarding such feedback is likely to prevent an individual from implementing the necessary improvement procedures outlined by FIT. An individual who has attributed a
negative evaluation to discrimination is unlikely to believe they had control over the outcome or that subsequent changes in behavior are likely to affect future outcomes (Ruggiero & Taylor, 1997). Hampered by doubts that increased effort will lead to improved performance, the individual may be more likely to reject the feedback and/or abandon their performance goals. Attributing failure to discrimination may be more detrimental when the individual believes the discrimination is likely to persist such as in academic and occupational settings. This may often the case in applied settings such as academic or occupational contexts. However, as has been the case for most laboratory research, when individuals believe the discrimination was an isolated event and is unlikely to occur again, they are more likely to benefit from the self-esteem protective benefits of making attributions to discrimination (Crocker and Major, 1989).

When feedback is accompanied by cues that undermine its validity the nature of the feedback itself is affected. It is unlikely that the response to such feedback would be the same as it would in cases where the individual had no reason to question the validity of the feedback. Belief that discrimination plays a role in an outcome or evaluation raises the specter of doubt and uncertainty. In the case of negative feedback, there is little reason to believe that performance should improve with increased effort if the same biased evaluator also will be evaluating the subsequent performance. However, providing individuals with information suggesting that subsequent performances will not be subject to bias (i.e., by informing participants that they will be evaluated by a different, unbiased judge) gives individuals an opportunity to reduce their uncertainty regarding the extent to which bias influenced their previous evaluation (e.g., D’Agostino & Pittman, 1982; Trope & Bassok, 1982; Sorrentino & Short, 1986; Weary & Edwards, 1994). In addition,
such an opportunity may also inspire some individuals to attempt to prove to themselves (and possibly others) that their capabilities are higher than that for which the biased evaluator gave them credit (Lazarus & Folkman, 1985).

Making external attributions for negative feedback may lead to a denial of personal responsibility. In the case of discrimination, such responses can be beneficial because a person who believes that discrimination is to blame for a negative outcome need not feel bad about it, thereby facilitating self-esteem maintenance. Nonetheless, such attributions may also lend themselves to beliefs about one’s inability to have changed the outcome in the past (cf. Thompson, 1981, Ruggiero & Taylor, 1997), or the inability to change future outcomes. An individual who believes that an outcome is not contingent upon his or her efforts or abilities has little incentive to make improvements (Abramson, Seligman, & Teasdale, 1987). This is particularly problematic in situations where discrimination was not the actual cause of the failure but, rather, the outcome actually was contingent upon one’s efforts and abilities. Individuals who make no attempt to improve are less likely to experience subsequent success than those who do make the attempt. Furthermore, a secondary consequence of lack of improvement is that one is unlikely to enhance one’s beliefs about their abilities, thereby precluding any increases in self-esteem.

Ironically, the very same thought processes that buffer individuals from proximal decreases in self-esteem may ultimately prevent distal increases in self-esteem. Relative to internal attributions, external attributions for failure may lead to performance deficits on subsequent tasks. In turn, performance decrements may eventually lead to lower self-esteem relative to those who made internal attributions.
Goals of the Present Research

The purpose of the present research is to examine how and when drawing attributions to discrimination following failure feedback affects subsequent performance. Based on prior research (e.g., Major et al., 2003), overt discrimination cues are more likely to elicit attributions to discrimination following failure than are ambiguous or absent discrimination cues. In turn, attributions to discrimination should exert differential effects on subsequent performance depending on whether individuals believe that they will be re-evaluated by a biased judge as opposed to a new, more objective judge. By default, when individuals receive failure feedback, there is a general tendency to increase effort on subsequent tasks (e.g., Kluger & DeNisi, 1996). However, there is little reason to increase effort on a future task when an individual believes that the same biased evaluator who negatively influenced previous incomes will also be evaluating future outcomes. Performance increases may still occur following a previous discriminatory evaluation, on the other hand, if the individual believes that a new and presumably less biased rater will evaluate future outcomes because an opportunity has been presented to demonstrate that the initial set of negative feedback did not properly characterize their true abilities. In addition, it provides an opportunity to reduce uncertainty about the underlying causes of the initial failure.

Study Overview

Female participants will complete a set of anagrams that purportedly measure their verbal ability and which are scheduled to be included in actual college entrance exams. In addition to receiving points for each anagram they solve correctly, participants will be told that their overall score will also depend upon subjective measures of their
performance such as originality, creativity, variety, and perceived difficulty of the identified words. Moreover, all participants will be informed that a male graduate student will be completing the initial evaluation (cf. Major et al., 2003). Following completion of this set of anagrams, a female confederate (ostensibly, a participant in the study) will either provide an overt discrimination cue (i.e., suggesting that she knows the evaluator and that he is “definitely” biased against women), an ambiguous discrimination cue (i.e., suggesting that the male student “might” be biased against women), or will make a control statement with no mention of evaluator bias. Additionally, in the overt condition, participants will hear the experimenter mention to the male confederate that he received the highest evaluation on the initial task.

Immediately prior to the receipt of feedback, half of the participants will then be told that the same evaluator (i.e., who had been described as potentially biased in the overt and ambiguous cue conditions) will be evaluating them again on a second set of anagrams (no perceived opportunity condition). The other half will be told that a different evaluator will be evaluating them on a second set of anagrams (perceived opportunity condition).

All participants will then receive feedback indicating that they performed poorly on the initial anagram task. Next, they will respond to several questions regarding the likelihood that they had been discriminated against and the degree to which their previous score was due to various internal and external factors.

Participants will then complete a second set of anagrams. Persistence (i.e., time spent completing the second set of anagrams) and (objective) performance will be the key dependent variables.
Hypotheses

Preliminary Hypotheses

The present research examines how and when attributions to discrimination for a negative outcome affect subsequent performance. Before describing the major hypotheses regarding subsequent performance, a set of preliminary hypotheses are presented that focus on replicating the results of prior research that has examined the impact of differential cues to discrimination on failure attributions (e.g., Major et al., 2003).

Preliminary Hypothesis 1

Participants will discount their feedback more (i.e., display a greater tendency to draw external attributions to discrimination relative to internal attributions for their negative performance) as discrimination cues increase in strength. Thus, participants in the overt condition will discount more than will those in the ambiguous condition, and those in the ambiguous condition will discount more than will those in the no mention condition.

Preliminary Hypothesis 2

Discounting will buffer against decreases in self-esteem. Because participants in the ambiguous and especially the overt discrimination cue conditions are more likely to discount than are no mention participants, these two groups should also be more likely to maintain self-esteem relative to participants in the no mention condition.

Major Experimental Hypotheses

In general, failure will lead to enhanced performance on the subsequent task. This prediction is supported by Kluger and DeNisi’s (1996) FIT theory and has also been
replicated by Berlin and Markman (2005) within a discrimination context. This overall effect, however, should be moderated by several factors.

Hypothesis 1

For participants in the same evaluator (i.e., no future opportunity) condition, performance change will decrease as discrimination cues increase in strength, from no mention, to ambiguous, to overt. Because discounting is expected to increase as cues increase in strength, there should also be a negative relationship between discounting and performance change in the same evaluator condition.

This hypothesis is based on the notion that individuals who believe their outcomes are determined by internal factors are more likely to believe they have the ability to affect changes (Abramson, Seligman & Teasedale, 1978; Testa & Major, 1990), whereas those who believe their outcomes are determined primarily by external factors are less likely to believe that increased effort will lead to improved outcomes.

Additionally, although negative feedback generally results in enhanced performance, cues about the evaluator’s bias undermine the utility of the feedback. Individuals who have reason to believe that their score is inaccurate are unlikely to make the same use of feedback because it is no longer informative with regard to how they performed in relation to their performance standard. Such biased feedback loses its diagnostic value and no longer informs individuals about the necessity of behavioral changes.

Hypothesis 2

For participants in the different evaluator (i.e., future opportunity) condition, performance will increase as discrimination cues increase in strength. Consequently,
there should be a positive relationship between discounting and performance change in the different evaluator condition.

Exposure to the discrimination cue is likely to lead participants to attribute their initial failure to external sources. For individuals in the same evaluator condition, there is little reason to believe that their feedback on the second performance will be any different. However, participants who are aware that subsequent trials are to be evaluated by a different (and presumably unbiased) evaluator are more likely to believe they that have the ability to affect future outcomes, and to demonstrate that their first score was not reflective of their abilities. Support for this prediction comes from Kluger & DeNisi’s (1996) FIT in that participants in this condition have the opportunity to attain higher self-goals such as affirming that the biased evaluator was unjustified in his scoring, and disconfirming the gender-based stereotypes that such bias brings to mind.

Uncertainty regulation is another mechanism by which individuals who attributed the initial feedback to discrimination might exhibit performance enhancement relative to those who believed that their feedback was veridical. Individuals who receive ambiguous discrimination cues, and to a lesser extent overt cues, are faced with a certain degree of uncertainty regarding the extent to which their negative feedback is an indicator of actual performance or purely a byproduct of the evaluator’s gender bias. However, only those individuals in the different evaluator condition have the opportunity to resolve this uncertainty. It is this opportunity that is likely to lead individuals in the discrimination cue conditions to exhibit enhanced performance above and beyond that generally associated with the receipt of legitimate negative feedback. Moreover, this is particularly likely to be the case for those who receive ambiguous discrimination cues, because
individuals who receive overt cues can more easily discount the initial feedback and therefore experience less ambiguity.

*Individual Differences*

Six individual difference factors will be considered. These factors are not central to the major experimental hypotheses, but will be examined nonetheless because they are potential moderators of the predicted effects. Three of the six individual difference factors have been previously employed in studies examining attributions to discrimination. These include Gender Identity (Luhtanen & Crocker, 1992), Desire for Control (Burger & Cooper, 1979), and Global Self-Esteem (Rosenberg, 1965). Three additional scales will also be included that may potentially moderate post-feedback reactions: the Implicit Person Theory scale (Dweck, Chiu & Hong, 1995), the Causal Uncertainty scale (Weary & Edwards, 1994), and the Achievement Motivation scale (McClelland, 1961).

*Implicit Theories of Intelligence*

Individuals who believe that intelligence is malleable and can be increased, rather than finite and inflexible, often try harder after failure (Dweck, 1999). Thus, the use of this scale may be of use with regard to interpreting changes in performance.

*Causal uncertainty*

Causal uncertainty indicates the extent to which individuals feel that they understand the causes of outcomes and events, and it is hypothesized that this variable is likely to have an effect on how participants react to ambiguous cues. Thus, individuals high in causal uncertainty may be more likely to feel uncertain that discrimination contributed to negative feedback than would individuals low in causal uncertainty. Thus,
participants who score high in causal uncertainty may try harder on the subsequent task in
the face of ambiguous discrimination cues than those who score low in causal uncertainty
because the former are more motivated to understand whether the feedback they received
was veridical, as opposed to being the result of discrimination.

Achievement motivation

In general, participants high in achievement motivation should be more likely to
respond to a failure with increased effort than participants low in achievement
motivation.

Method

Participants

Data was collected on 103 female Ohio University students in exchange for
partial credit in a general psychology course. Information from 13 participants was
omitted; 11 participants failed to successfully complete the study (9 of whom failed to
follow instructions such as using all of the letters for the anagrams, two were unable to
finish in the allotted time) one participant was omitted for mentioning to the confederates
that she had not read the questions before responding, and a final participant was omitted
because she responded to all questions using only the extreme endpoints on each scale.
All data analyses were conducted on the remaining 90 participants.

Materials

Prior to their participation in the experiment, participants completed six individual
difference measures during a computerized pretest session (see Appendix A). These
measures include the Rosenberg (1965) self-esteem scale, the Desire for Control Scale
(Burger & Cooper, 1979), the collective identity subscale of Luhtanen and Crocker’s
(1992) Collective Self-Esteem Scale (as adapted by Major et al., 2003), the Achievement Motivation questionnaire (McClelland, 1961), the Causal Uncertainty Scale (Weary & Edwards, 1996), and the Implicit Person Theory Scale (Dweck, Chiu & Hong, 1995).

During the session, participants also completed the social and performance subscales of Heatherton and Polivy’s State Self-Esteem Scale (1991; see Appendix B). This scale was designed to measure momentary levels of self-esteem, and reveals changes in state self-esteem over short periods of time. It was administered at the beginning of the session (before the initial anagram task) and after receiving performance feedback (but before the second anagram task). A 28-item mood measure identical to the one developed by Major, Kaiser and McCoy (2003; originally the Multiple Affective Adjective Check list, Zuckerman & Lubin, 1956; see Appendix B) was also administered immediately following the state self esteem scale both times.

**Design and Procedure**

The study employed a 3 (Discrimination Cue: no cue, ambiguous cue, overt cues) by 2 (Future Opportunity: same evaluator, different evaluator,) design.

After arriving at the lab and signing informed consent forms, participants were told that they would be helping to standardize a test of verbal ability. They were told that the test an important predictor of success in college and was scheduled to be included in upcoming college entrance exams. A female experimenter explained that the session that day would involve solving anagrams, and would be graded not only by the quantity of solutions found, but also by the quality and creativity of their responses. For this reason their tests would be evaluated by a rater who was described as a male graduate student named Luke Young. Next, the experimenter explained that in order to make the testing
situation as much like real life as possible, it will involve the potential for reward. Because good test performance can lead to financial scholarships, the highest scoring student in each of the two anagram sessions would be entered into a drawing for the opportunity to win $100 in several weeks after the study was complete.

Before they could begin the actual test, participants were told that they must respond to brief demographic questions, including gender and first name, and then complete the state self-esteem and mood measures. Participants next completed the first of two sets of 10-item anagrams (adapted from Shah, Higgins and Friedman, 1998). Each anagram contained 5-7 letters that could be arranged to reveal multiple solutions (see Appendix C). These anagram sets have been previously used in psychological research and are known to be fairly equivalent in terms of average difficulty and completion times.

After completing the initial anagram session, the computer notified participants that their answers had been sent to the evaluator, and that they should remain in their seats until the experimenter informed them that it was time to move on. Two confederates, one male and one female, always completed their anagrams about one to two minutes after the last participant has finished. After completion of the first session, the experimenter asked all participants to move to a table in the center of the room where a filler task was completed. The filler task was described as a visualization task and required participants to list as many buildings on campus as possible. On the back of the sheet there were instructions to list as many countries of the world as possible, which they could do if they were unable to come up with any more buildings. The filler task lasted seven minutes.
Discrimination cues

After providing instructions for the filler task, the experimenter left the room, ostensibly to make sure that the rater had received all of their responses. While she was gone, a female confederate made one of three statements to the group at large. In the no mention condition, she said, “You know, the psychology department must do a lot of these studies. I was in one just like it last week,” in the ambiguous cue condition she said, “You know, I have friends who were in this study and they told me that the guy doing the evaluating totally grades guys and girls differently,” and in the overt cue condition she said, “You know, I know the guy who is doing the evaluating, and he’s totally prejudiced. There’s no way he’ll say a girl got the best score” (cf., Major et al., 2003).

Second Anagram Session

After seven minutes the experimenter returned and informed participants that the evaluator had finished scoring their first session responses. She reminded participants of the anagram rules and grading criteria. At this point the experimenter told participants either that their next set of responses would be evaluated by the same evaluator, or by a different evaluator. Next, she asked them to return to their computer terminals and enter their first name, “so that the rater can send the scores to the right computer.” This served to reinforce the idea that the rater had access to their demographic information including, most importantly, their gender. At this point in the overt condition, the experimenter told the male confederate that he received the highest score and that he would receive a lottery ticket at the end of the experiment. Although this was stated directly to him, it was audible for all participants and occurred while they were waiting for their feedback. In all cue conditions, participants read a statement on their score sheet stating that they were
not the recipient of the lottery ticket. Following a pause while the scores were being
“sent,” the participants received negative feedback indicating that their ratings fell in the
34th percentile of all Ohio University students who had previously completed the test (see
Appendix D). The score sheet also provided an interpretation of their specific percentile
rank. Additionally, the score sheet included the phrase, “Responses were mostly lacking
creativity. Few of the more difficult answers were found.”

Attributional assessment

After receiving feedback, participants responded to questions designed to assess
the extent to which their feedback was due to internal versus external factors, as well as
more specific attributions to discrimination (see Appendix E). The questions were similar
to those used in prior research (e.g., Morera et al., 2004; Sechrist et al., 2004; Major et
al., 2003), including, “My score is a reflection of my ability,” and, “I feel that my score
was due to a prejudiced evaluator.” Likewise, participants responded to questions
assessing the extent to which they felt they had control over the outcome they just
received, as well as control over subsequent outcomes (cf., Ruggiero & Taylor, 1997).

After completing the attributional assessment, participants completed a second
round of state self-esteem and mood questionnaires, identical to those completed before
the first set of anagrams. Following completion of the second set of anagrams,
participants completed a funnel debriefing in which they first described the purpose of
the study, and then commented on the believability of the feedback. All participants were
then informed of the confederates’ role in the study and were told that their feedback was
not reflective of their actual performance or abilities. Finally, they were informed that all
participants would be entered into the lottery with equal chance of winning the $100.
Results

Preliminary Hypotheses

Preliminary Hypothesis 1

Discounting. A weighted mean of attributions to ability and effort was computed to form an overall internal attributions variable. Attributions to discrimination were also averaged. Both of these averages were then standardized separately. The standardized internal attributions were subtracted from the standardized attributions to discrimination to form a discounting variable (cf. Major et. al, 2003), where positive numbers indicate relatively higher external than internal attributions. Attributions to internal factors and to discrimination were negatively correlated, $-.32, p < .01$.

A 2(Evaluator: same, different) x 3 (Discrimination Cue: no mention, ambiguous, overt) between subjects analysis of variance (ANOVA) revealed a significant main effect of discrimination cue on discounting, $F(2,84) = 3.64, p < .05$. As predicted, discounting increased as discrimination cues increased in strength from no mention ($M = -.61$), to ambiguous ($M = -.15$) to overt ($M = .40$). Planned comparisons revealed that participants in the overt discrimination condition discounted significantly more than did participants who received no mention about the likelihood of discrimination, $t(87) = 2.76, p < .05$. Although the hypothesized planned comparisons failed to reach significance from no mention to ambiguous, $t(87) = 1.14, p > .25$ or from ambiguous to overt, $t(87) = 1.40, p > .15$), means were in the expected direction.

For exploratory purposes, discounting was then analyzed in two separate one-way ANOVAs for each evaluator condition. There was a significant main effect of cue on discounting for the same-evaluator condition $F(2,42) = 4.16, p < .025$. The means
followed the same trend, increasing from no mention, to ambiguous to overt (see Table 1 for means). For participants in the different-evaluator condition there was no significant main effect for discounting ($p > .7$). These exploratory analyses suggest that the main effect of cue in the previous 2 x 3 ANOVA was driven by participants in the same evaluator condition.

Table 1.

<p>| Means and Standard Deviations of Attributions as a Function of Cue and Evaluator |</p>
<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Ability</th>
<th>Effort</th>
<th>Discrimination</th>
<th>Discounting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same evaluator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No mention</td>
<td>16</td>
<td>3.3 1.1</td>
<td>4.2 1.3</td>
<td>1.9 0.7</td>
<td>-.72 1.1</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>11</td>
<td>2.9 0.7</td>
<td>3.9 1.3</td>
<td>2.3 0.7</td>
<td>-.10 1.0</td>
</tr>
<tr>
<td>Overt</td>
<td>18</td>
<td>2.5 1.1</td>
<td>4.0 1.6</td>
<td>3.5 1.4</td>
<td>.82 1.7</td>
</tr>
<tr>
<td>Different evaluator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No mention</td>
<td>16</td>
<td>3.4 1.1</td>
<td>4.2 1.4</td>
<td>2.5 0.9</td>
<td>-.49 1.4</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>13</td>
<td>3.2 1.6</td>
<td>4.1 1.4</td>
<td>2.8 0.9</td>
<td>-.20 1.6</td>
</tr>
<tr>
<td>Overt</td>
<td>16</td>
<td>3.4 1.1</td>
<td>4.7 1.0</td>
<td>3.3 1.4</td>
<td>-.07 1.7</td>
</tr>
</tbody>
</table>

Note. Means for Ability, Effort and Discrimination are reported on a scale from 1 (strongly disagree) to 7 (strongly agree). The score of 4 was represented by the value label Neutral.

In part because the null findings in the different evaluator condition were unanticipated, and in part because of the integral role that discounting played in the formulation of all hypotheses, many of the subsequent analyses were first conducted to test the hypothesized effects, and were followed by separate, exploratory analyses for each evaluator condition and reported based on the adjusted alpha(.025) level of significance. Although this manner of conducting separate analyses is not usually called for in standard statistical procedures, the analyses are reported because of their additional informative value.
Preliminary Hypothesis 2

State self-esteem. Change in state self-esteem (SSE) was computed by subtracting the total performance and social SSE scores at Time 1 from the total of the same scores at Time 2. SSE was submitted to a 2 (Evaluator) x 3 (Discrimination Cue) between-subjects ANOVA. There was no significant interaction, nor were there main effects for either manipulation. Analyses conducted separately on each evaluator condition also yielded no significant results\(^2\). Planned comparisons were not conducted. Although these results are not supportive of Preliminary Hypothesis 2, SSE did not play a central role in the subsequent hypotheses.

Major Experimental Hypotheses

Due to the large number of dependent variables being considered, a 2 (evaluator) x 3 (discrimination cue) multivariate analysis of variance was conducted on change in self-esteem, discounting, performance difference, and persistence difference. There was a significant cue by evaluator interaction, Wilks’ Lambda = .823, \(F(8,162) = 2.07, p < .05\).

Following the recommendations of Rencher (1995), subsequent univariate tests will be reported at the alpha(.05) level, unless otherwise stated.

Performance

A performance change variable was computed by subtracting the total number of anagrams correct on the second anagram set from the number of anagrams correct on the first so that positive values indicate an increase in the total number of correct solutions

\(^2\) Although the results of state self-esteem change were nonsignificant, other measures suggest that there may have been some effects to psychological well-being. There was a marginally significant negative relationship between attributions to discrimination and increase in depression, \(r = -.18, p = .08\). This was particularly true in the same evaluator condition where the correlation was marginally significant, \(r = -.30, p = .04\) at the adjusted alpha(.025) level. The relationship was nonexistent in the same evaluator condition, \(p > .50\). These results suggest that in the same evaluator condition, even though state self-esteem did not change, participants who did not attribute their failure to discrimination did, in fact, suffer negative consequences to psychological well-being.
identified. Performance was submitted to a 2 (evaluator) X 3 (discrimination cue) between subjects analysis of variance (ANOVA) \(^3\). The test of intercept was significant, \(F(1,84) = 9.96\), \(p < .01\), indicating that, as expected, there was a general tendency for performance to improve at Time 2. There was a significant main effect for cue, \(F(2,84) = 3.89\), \(p < .05\), such that participants in the overt condition exhibited the greatest increase \((M = 2.00)\), followed by participants in the no mention condition \((M = 1.78)\), while participants in the ambiguous discrimination condition performed worse \((M = -0.29)\) than on the initial session. Planned comparisons revealed that participants in the ambiguous condition exhibited less performance improvement than participants in both the no mention, \(F(1, 87) = 2.24, p < .05\), and the overt, \(F(1, 87) = 2.51, p < .05\), conditions. Contrary to the expectation that performance would increase least in the overt condition, participants in that condition did not differ significantly from the no mention condition in terms of performance change. To determine whether improvement occurred for each discrimination cue, the mean for each cue condition was compared to zero via three separate t-tests with the alpha(0.017) level adjusted by the Bonferroni method, revealing that the means for both no mention, \(t(87) = 8.63, p < .01\), and overt cues \(t(87) = 11.57, p < .01\) were greater than zero, while the ambiguous cue mean, \(t(87) = 0.17, p > .5\) was not different from zero.

**Exploratory analyses on performance.** Although neither the main effect for evaluator or the cue by evaluator interaction were significant, separate analyses were run for participants in the same and different evaluator conditions as exploratory analyses.

\(^3\) Performance difference was also analyzed in an ANCOVA with Time 1 performance entered as a covariate, and Time 2 performance as the dependent variable. Results of the ANCOVA and the reported ANOVA on difference scores yielded very similar results. Throughout the results section performance and persistence differences are reported rather than the results of ANCOVAs because of the ease with which difference scores can be interpreted in terms of performance increases and decreases.
For participants in the same evaluator condition, the one-way ANOVA yielded a significant main effect for cue on performance, $F(2,42) = 5.25', p < .01$, as in the previous analysis, means for the no mention ($M=3.0$) and overt ($M=2.8$) conditions, were higher than the ambiguous condition ($M=-1.1$). However, for participants in the different evaluator condition, contrary to the hypothesis that performance would improve as discrimination cues increased in strength, the one-way ANOVA dropped to nonsignificance, $p > .7$. The results indicate that the findings for the 2x3 ANOVA were driven by the performance of participants in the same evaluator condition.

Table 2.

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Persistence</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Same Evaluator</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Mention</td>
<td>16</td>
<td>-26.5</td>
<td>145.9</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>11</td>
<td>-50.2</td>
<td>204.3</td>
</tr>
<tr>
<td>Overt</td>
<td>18</td>
<td>43.1</td>
<td>123.0</td>
</tr>
<tr>
<td>Different Evaluator</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Mention</td>
<td>16</td>
<td>76.0</td>
<td>127.5</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>13</td>
<td>-1.4</td>
<td>98.8</td>
</tr>
<tr>
<td>Overt</td>
<td>16</td>
<td>10.6</td>
<td>146.0</td>
</tr>
</tbody>
</table>

*Note.* Persistence means are reported in terms of the difference in the total number of seconds spent from Time 1 to Time 2, and performance reports the total change in correct solutions identified across all 10 anagrams in each set.

**Performance and desire for control.** In order to examine the effects of Desire for Control (DFC) on performance, a median split was conducted on the DFC scale. The median desire for control score was 4.55 on a scale from 1 (low desire for control) to 7 (high desire for control) averaged across all 20 questions. Participants who scored higher than 4.55 on the DFC scale were classified as high DFC ($n = 43$) and those who scored at 4.55 or lower comprised the low DFC ($n = 47$) group. A 2 (DFC) by 2 (evaluator) by 3
(cue) ANOVA yielded a marginally significant Cue by Feedback by DFC interaction, 
\( F(2,78) = 2.25, p = .11 \) (see Figure 1). Separate analyses of the same and different 
evaluator conditions suggest that this interaction was, again, driven by the same evaluator 
condition. A 2 (DFC) by 3 (cue) ANOVA conducted on participants in the same 
evaluator condition revealed a marginally significant cue by DFC interaction, 
\( F(2,39) = 3.73, p = .03 \). While the performance of low DFC participants did not differ from no 
mention (\( M = 2 \)) to ambiguous (\( M = .83 \)) or overt (\( M = 1.29 \)), all t’s < 1, the performance 
of high DFC participants did vary as a function of cue. Planned comparisons (collapsed 
across evaluator) revealed that high DFC participants in the ambiguous condition (\( M = -3.4 \)) performed worse than high DFC participants in the no mention (\( M = 3.78 \), \( t(24) = 3.78, p < .01 \)), and overt (\( M = 3.82 \), \( t(24) = 4.14, p < .01 \)) discrimination cue conditions. 
Each of the twelve means in the cue by evaluator by DFC analysis were submitted to t-
tests against zero, to determine for which groups performance changed significantly from 
the initial task. Results indicated that only the means in the same evaluator, high DFC 
group differed from zero when a Bonferroni adjustment was applied to the alpha level. 
Participants in the same evaluator no mention, (\( M = 3.78 \) \( t(78) = 12.33, p < .01 \)), and overt, (\( M = 3.82 \) \( t(78) = 15.40, p < .01 \)), both showed significant increases while high 
DFC participants in the same evaluator ambiguous condition exhibited a decrease in 
performance, (\( M = -3.40 \) \( t(78) = 5.58, p < .01 \)); all p’s reflect the Bonferroni adjustment. 
None of the means in the different evaluator condition differed significantly from zero, 
nor did the means of any of the low DFC groups. The means were in the same direction 
as previous analyses of performance differences and suggest that the effect of
discrimination cue on performance was driven by individuals who were high in desire for control. There were no significant effects of DFC in the different evaluator condition.

Figure 1.

Mean performance difference for low and high DFC as a function of cue and evaluator.

Persistence

The total amount of time spent on the initial anagram set was subtracted from time spent on the second set to form a measure of persistence difference. Performance and persistence differences were positively correlated, $r = 0.45, p < .001$. Persistence was submitted to a 2 (evaluator) by 3 (discrimination cue) ANOVA. There were no significant main effects for either cue or evaluator and no cue by evaluator interaction.

For exploratory purposes, persistence differences were subjected to separate one-way ANOVAs for each evaluator condition. Performance and persistence were positively correlated for both same, $r = 0.54, p < .001$, and different, $r = 0.39, p < .01$, evaluator conditions. There were no effects of cue on persistence for either evaluation condition.
**Discounting and performance.** A correlation was conducted to investigate the relationship between performance change and discounting. There was a marginally significant positive relationship between discounting and performance difference, $r = 0.194, p < 0.07$. Contrary to expectation, the association between discounting and performance was moderately positive. Separate correlations were computed for the same, $r = 0.224, p < 0.15$, and different, $r = 0.130, p < 0.40$, evaluator conditions, and both yielded positive but nonsignificant results.

**Discussion**

The results of this study indicate that discrimination cues did, in fact, have an impact on performance changes during two trials of an identical task, but only when that performance was to be evaluated by the same person. In general, performance decreased when participants were given ambiguous information about the likelihood that they would be discriminated against, but increased when they were given either clear cues that discrimination would occur or no mention about the possibility of discrimination. When participants believed that the second session would be evaluated by a different, and presumably unbiased, evaluator the effects of discrimination cues were attenuated. Although the pattern of performance change was not as predicted, possible explanations for these findings will be discussed. It is first necessary to consider how the results compare with previous research on attributions to discrimination.

**Comparison to previous studies**

**Discounting.** Similar to previous research, participants discounted feedback more as discrimination cues increased in strength (e.g., Major et al, 2003). Unexpectedly, this was only the case when participants believed they would be evaluated by the same rater.
in the subsequent session. Participants who believed they would be evaluated by a
different rater did not discount their feedback on the initial task.

One possible explanation for this finding is that participants who were aware that
they would be evaluated by a different person might have felt threatened by the situation.
Receiving approximately the same score on the second evaluation could have the effect
of validating that the initial feedback was accurate. If they were to have claimed that the
initial negative feedback was due to a biased evaluator, and then received a similar
negative feedback on the second round, they may have feared they would be perceived as
having ‘cried wolf’, so to speak, on the first task. If women are generally reluctant to
claim personal discrimination under ambiguous circumstances (e.g., Ruggiero & Taylor
1995, Kaiser & Miller, 2001), then they may be particularly unlikely to do so when there
is a possibility that they may be proven wrong. Perhaps it is psychologically less painful
to make internal attributions for a failure than to risk making falsifiable claims of
discrimination.

Yet another explanation for why participants in the different evaluator condition
did not discount their feedback may be related to the experimental design. Specifically, it
is possible that participants believed that because their overall performance would be
evaluated by different people, any discrepancies in the scoring process would be
resolved. Similarly, they may have believed that because there were two evaluators, the
test, as a whole, was not likely to reflect biases, or at the very least was set up to detect
evaluator biases.
Performance

As expected, there was a general tendency for performance to increase from Time 1 to Time 2. These findings are consistent not only with practice effects, but also Feedback Intervention Theory (Kluger & DeNisi, 1996) which posits that effort and performance both increase following negative feedback on personally relevant domains where performance standards cannot easily be abandoned. Performance change means were greater than zero in all conditions except the same evaluator – ambiguous condition. In this condition performance change was significantly lower than the performance increase exhibited by the other cue conditions for the same evaluator.

Discrepancies between hypotheses and obtained results. The original hypothesis was that in the same evaluator condition, performance would decrease as discrimination cues increased in strength. Two separate justifications lent support to this expectation. First, individuals who believed that their outcome was due to a biased evaluator rather than to their own efforts and abilities might be unlikely to believe that increased effort would lead to an improved outcome on the subsequent task because the same evaluation biases are likely to recur. Second, attributing feedback to an external source, such as evaluator bias, renders the feedback nondiagnostic of the actual level of performance, and is unlikely to result in the same performance increases as would diagnostic feedback. Thus, although all individuals receive the same negative feedback, only those who receive no mention of discrimination cues are likely to believe that feedback is diagnostic of their actual performance. Consequently, only those individuals were expected to increase effort and experience performance improvement. However, the findings were not in the expected direction. In fact, performance in the no mention and overt
discrimination cue condition did not differ significantly. Instead, performance decreases were only exhibited in the ambiguous condition.

It is unclear why in the same evaluator condition performance change dropped drastically in the ambiguous cue condition but remained high in the overt condition. The most plausible explanation is that it was situational ambiguity itself that led to the decrease in performance. Research in the area of fairness judgments indicates that individuals are particularly sensitive to fairness when they are in uncertain or ambiguous situations (van den Bos & Lind, 2002). Individuals who were uncertain about the degree to which their performance was due to discrimination rather than their performance may have been uncertain about how to interpret their information. This ambiguity may have prevented them from concluding that increased effort would lead to performance improvements. Consequently, they may have been unmotivated to increase effort or performance. Participants in the overt condition, on the other hand, may also have been deprived of diagnosticity from feedback but would have the motivation to increase effort in order to demonstrate to themselves, the evaluator, or to both, that the first evaluation was due to bias and that they were better than their score indicated. This would explain not only how performance increased in that condition, but also accounts for the positive (although only marginally significant) correlation between discounting and performance. That is to say, the more certain individuals were that their feedback was due to external factors, the more their motivation to improve (or to disprove their initial feedback).

The current study provides insufficient information to directly support the explanation that ambiguity caused the performance differences. Some of the results even suggest reasons why ambiguity may not be the best explanation. One reason why
ambiguity might not be the best explanation is that participants in the overt condition reported discrimination means that fell in the “somewhat disagree” range of the scale. An individual who “somewhat disagrees” with an overt cue may not be very certain about whether or not discrimination occurred. Because individuals in the overt condition were probably not certain that their feedback was due to discrimination, it is unclear whether a lack of certainty caused the performance decrements in the ambiguous condition. Consequently, it is unclear why ambiguity would have had a more dramatic effect on individuals in the ambiguous cue condition than overt cue condition when in fact, both groups may have experienced a high degree of ambiguity. In other words, although the condition was labeled “overt discrimination,” participants in that condition may still have felt some ambiguity about whether or not discrimination had actually occurred.

The performance drop in same evaluator – ambiguous condition seemed to be exacerbated by participants who were high in desire for control. This finding is consistent with the idea that it was ambiguity itself that led to the decrease in performance. Individuals who are high in desire of control might have been particularly affected by uncertainty about the degree to which their outcomes were in their own control, or were due to an external source. Although the results involving DFC were only marginally significant, this is most likely due to power problems arising from very low numbers in that condition and would need more thorough investigation in future research.

Discounting and performance. No support was provided for the hypothesis that performance changes would decrease as individuals discounted their feedback. Although discrimination cue affected both discounting and performance difference, the relationship between discounting and performance change was marginally significant and positive. A
quick glance at the means provides the explanation for this finding. Discounting increased as discrimination cues increased in strength. Performance change, on the other hand, was positive for those in the no mention condition (low discounting) as well as for those in the overt condition (high discounting), while performance change was negative for participants in the ambiguous (mid-range discounting) condition. It was expected that, especially in the same evaluator condition, participants in the no mention condition would perform better than those in the ambiguous cue condition. What was not expected was that participants in the overt condition performed as well as those in the no mention condition. It is their high performance increase that led to the positive correlation between discounting and performance. It does not appear that discounting, per se, was responsible for the drop in performance exhibited by those in the ambiguous discrimination condition.

Future directions

Direct measurement of ambiguity. The current results give little indication of how much ambiguity was perceived as a result of discrimination cues. It is possible that both participants in the ambiguous and overt conditions were fairly uncertain about exactly how much of their outcome was due to discrimination. Future studies might make use of confidence ratings following each attribution question in order to gain a better understanding of the degree of perceived ambiguity and its relationship to performance change.

Strategies for coping with negative feedback. Multiple strategies may be employed in reaction to negative feedback. Feedback Intervention Theory (Kluger & DeNisi, 1996) suggests four basic strategies, only one of which involves increasing
effort. The other three strategies are: abandoning the standard, changing the standard, and rejecting the feedback. The current study attempted to avoid these other three strategies by employing a domain (tests of intelligence and verbal ability) that is difficult to devalue or change. College freshmen are likely to believe that intelligence matters, that outcome matters more than effort on standardized tests of intelligence, and that standardized tests of intelligence are at least somewhat accurate. However, it is possible that they did not believe that anagrams were a valid test of verbal ability. In fact, several participants mentioned that very issue during the debriefing. In cases in which individuals reject performance goals by abandoning or changing the standard or by simply rejecting the feedback, perceived control over future outcomes might not lead to increased effort or performance. Although the current study sought to prevent the use of such alternative strategies, future studies should seek to measure the extent to which these alternatives may have occurred anyway.

Comparison of nonstigmatized groups. It is wise, when studying the effects of stigma, to understand how the severity of a stigmatized identity interacts with the perception of discrimination. Acknowledgement of discrimination may lead to more negative consequences to psychological well-being for highly stigmatized groups than for relatively advantaged ones (Branscombe, 1998; Branscombe, Schmitt & Harvey, 1999; Schmitt, Branscombe, Kobrynowicz & Owen, 2002). Partial explanation for this comes from the fact that attributions to discrimination may have a larger internalized component for members of disadvantaged groups (Schmitt & Branscombe 2001, 2002). The current study sought to avoid such internalization by using verbal ability, a domain in which women are generally believed to perform at least as well as men. However, it may be the
case that women are generally more threatened than men by the possibility of
discrimination in academic contexts. Future research should investigate these issues by
comparing not only stigmatized and nonstigmatized groups, but also stigmatized and
nonstigmatized domains.

Conclusion

Research on successful coping strategies and reactions to discrimination has
recently begun to focus on identifying factors that lead to the accurate identification of
discrimination. Accurately attributing negative outcomes discrimination protects self-
esteem in certain situations. Although it is important to understand situational and
personality factors that may lead to these beneficial effects, it is also important to
understand the potential long term consequences. When discrimination occurs in
performance contexts, it is particularly important to understand the long term
consequences of both attributing negative outcomes to discrimination as well as to
internal factors. The current study demonstrates that there may be detrimental effects on
performance in ambiguous discrimination contexts. In situations with ambiguous
discrimination cues, it is possible that although attributing a failure to discrimination may
protect an individual from temporary decreases in self-esteem, subsequent performance
deficits may have detrimental effects on self-esteem in the long term.
References


Major, B., Quinton, W. J., & McCoy S. K., (2002). Antecedents and consequences of attributions to discrimination: Theoretical and empirical advances. *In M. P. Zanna*


Appendices

Appendix A: Pretest Measures

Achievement Motivation

1. People who work so hard they make the rest of us look bad really bother me.*
2. I like taking risks.
3. I ask others for advice even if I think I know the answer, because it is always better to be safe than sorry.*
4. I like it when people say in front of others that I am doing a good job.
5. I dislike entering a room full of strangers and trying to connect.*
6. If you met me, you would say that I get on well with other people.
7. I would be more successful, but others try to disrupt my plans.*
8. In a tight situation, I like it when I am in charge and the blame or praise will come to me.
9. I enjoy spending much of my time alone.*
10. I like being independent.
11. In fantasies about my career, often I am in a group and they are praising me.
12. Life would be better if more people stuck to their business so I could stick to mine.
13. I am basically a competitive person, and I compete just for the sake of competing.
14. People do not like to admit it, but success in life has less to do with hard work and more to do with luck and being in the right place at the right time.*
15. Having good friends is important, but I could quickly make new ones if I had to.
16. When it is possible to avoid conflict, I do so.*
17. Most people who know me say I am ambitious.
18. If jobs and money are on the line, it is a good idea to let someone else be in charge, in case things go sour.*
19. If I knew others disapproved of my actions, it would cause me to rethink my plans.*
20. Being part of a team at work is less important than doing good work on your own.
21. I regularly list my goals where I can see them during the day.
22. Most evenings, I kick back and relax rather than prepare for the next day's tasks.*

Note. Responses were reported in true-false format with a score of one indicating true, and a score of zero indicating false. Items marked with an asterisks (*) indicate reverse scoring. The scale was developed by McClelland (1961).
(Global) Self-Esteem Scale

1.) On the whole I am satisfied with myself.
2.) At times I think I am no good at all.*
3.) I am able to do things as well as most other people.
4.) I do not feel I have much to be proud of.*
5.) I certainly feel useless at times.*
6.) I feel I am a person of worth, at least on an equal playing field with others.
7.) I wish I could have more respect for myself*
8.) All in all, I am inclined to feel that I am a failure.*
9.) I take a positive attitude toward myself.
10.) I am worried about looking foolish.*

Note. Responses were reported on a 5 point scale where higher numbers indicate more agreement. Items marked with an asterisks (*) indicate reverse scoring. The scale was developed by Rosenberg (1965).

Causal Uncertainty Scale

1. I do not know what it takes to get along well with others.
2. When I receive good grades, I usually do not understand why I did so well.
3. I do not understand what causes most of the problems that I have with others.
4. When I see something good happen to others, I often do not know why it happened.
5. When I receive poor grades, I usually do not understand why I did so poorly.
6. When someone I know receives a poor grade, I often cannot determine if he or she could have done anything to prevent it.
7. I do not understand what causes most of the good things that happen to me.
8. When things go right, I generally do not understand what to do to keep them that way.
9. When bad things happen, I generally do not know why.
10. When there is more than one possible reason for a person’s action it is difficult to determine which one is the actual reason.
11. I often feel like I don’t have enough information to come to a conclusion about why things happen to other people.
12. When I see something happen to others, I often do not know why it happened.
13. I often feel like I do not have enough information to come to a conclusion about why things happen to me.
14. When I think about why someone does something, there are usually so many possible reasons for it that I cannot determine which one was the cause.

Note. Responses were reported on a 7-point scale where higher numbers indicated increasing agreement. The scale was developed by Weary & Edwards (1994).
Desirability of Control scale

1.) I prefer a job where I have a lot of control over what I do and when I do it.
2.) I enjoy being able to influence the actions of others.
3.) I enjoy making my own decisions.
4.) I enjoy having control over my own destiny.
5.) I consider myself to be generally more capable of handling situations than others are.
6.) I’d rather run my own business and make my own mistakes than listen to someone else’s orders.
7.) Others usually know what is best for me.
8.) I wish I could push many of life’s daily decisions off on someone else.*
9.) There are many situations in which I would prefer only one choice rather than having to make a decision.*
10.) I like to wait and see if someone else is going to solve a problem so that I don’t have to be bothered by it.*
11.) When driving, I try to avoid putting myself in a situation where I could be hurt by someone else’s mistake.
12.) I like to get a good idea of what a job is all about before I begin.
13.) When I see a problem I prefer to do something about it rather than sit by and let it continue.
14.) I am careful to check everything on an automobile before I leave for a long trip.
15.) I try to avoid situations where someone else tells me what to do.
16.) I prefer to avoid situations where someone else has to tell me what it is I should be doing.
17.) I enjoy political participation because I want to have as much say in running government as possible.
18.) I would prefer to be a leader rather than a follower.
19.) I would rather someone else took over the leadership role when I’m involved.
20.) When it comes to orders, I would rather give them than receive them.

Note. Responses were reported on a 7-point scale where higher numbers indicated increasing agreement. Items marked with an asterisk (*) indicate reverse scoring. The scale was developed by Burger and Cooper (1979).
Gender Identification Scale

1. Overall, being a woman has very little to do with how I feel about myself.*
2. Being a woman is an important reflection of who I am.
3. Being a woman is unimportant to my sense of what kind of person I am. *
4. In general, being a woman is an important part of my self-image.

Note. Responses were reported on a 5 point scale where higher numbers indicate increasing agreement. Items marked with an asterisk (*) indicate reverse scoring. The Gender Identification Scale was adapted by Major Quinton & Schmader (2003) from Luhtanen and Crocker’s (1992) Collective Self-Esteem Scale.

Implicit Person Theory

1. You have a certain amount of intelligence and you really cannot do much to change it.
2. Your intelligence is something about you that you cannot change very much.
3. You can learn new things, but you cannot really change your basic intelligence.
4. The kind of person someone is, is something very basic about them and it cannot be changed very much.
5. People can do things differently, but the important parts of who they are cannot really be changed.
6. Everyone is a certain kind of person and there is not much that can be done to really change that.

Note. Responses were reported on a 7-point scale where lower numbers indicate malleable theories of intelligence and personality, and higher numbers indicate fixed theories. The scale was developed by Dweck, Chiu & Hong (1995).
Appendix B: In-Session Computer-Based Questionnaires

State Self-Esteem Scale

This questionnaire is designed to measure what you are thinking at this moment. There is, of course, no right answer for any statement. The best answer is what you feel is true of yourself at this moment. Be sure to answer all of the items, even if you are not certain of the best answer. Again, answer these questions as they are true for you RIGHT NOW.

1. I am confident about my abilities.
2. I am worried about whether I am regarded as a failure.
3. I feel frustrated or rattled about my performance.
4. I feel that I am having trouble understanding things that I read.
5. I feel self-conscious.
6. I feel as smart as others.
7. I feel displeased with myself.
8. I am worried about what other people think about me.
9. I feel confident that I understand things.
10. I feel inferior to others at this moment.
11. I feel concerned about the impression I am making.
12. I feel that I have less scholastic ability right now than others.
13. I feel like I’m not doing well.

Note. Original scale was created by Heatherton & Polivy (1991). Items marked with an asterisks (*) indicate reverse scoring. The subscales are for Performance and Social self-esteem. The appearance subscale has been omitted from the original format.
MAACL – Mood Index

<table>
<thead>
<tr>
<th>Affect</th>
<th>Hostility</th>
<th>Anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worthless</td>
<td>Angry</td>
<td>Fearful</td>
</tr>
<tr>
<td>Proud *</td>
<td>Cooperative*</td>
<td>Worried</td>
</tr>
<tr>
<td>Embarrassed</td>
<td>Cruel</td>
<td>Calm*</td>
</tr>
<tr>
<td>Like a failure</td>
<td>Agreeable*</td>
<td>Secure*</td>
</tr>
<tr>
<td>Disappointed in</td>
<td>Mad</td>
<td></td>
</tr>
<tr>
<td>myself*</td>
<td>Scornful</td>
<td></td>
</tr>
<tr>
<td>Pleased with myself*</td>
<td>Irritable</td>
<td></td>
</tr>
<tr>
<td>Humiliated</td>
<td>Hostile</td>
<td></td>
</tr>
<tr>
<td>Ashamed</td>
<td></td>
<td>Depression</td>
</tr>
<tr>
<td>Inferior to others</td>
<td></td>
<td>Discouraged</td>
</tr>
<tr>
<td>Sad</td>
<td></td>
<td>Fine*</td>
</tr>
<tr>
<td>Depressed</td>
<td></td>
<td>Active*</td>
</tr>
<tr>
<td>Mortified</td>
<td></td>
<td>Blue</td>
</tr>
</tbody>
</table>

Note. Responses were provided on a scale from 1 (not at all) to 7 (extremely). Items marked with an asterisk are reverse coded. All items were presented in a random order.

The scale was originally part of Zuckerman & Lubin’s (1965) Multiple Affective Adjective Check List. It appears as it was adapted by Major Kaiser & McCoy (2003).

Demographic Information

Gender  
First Name  
Date of Birth  
Year in School  
Major of Study (If necessary write “undecided”).  
Oak ID  
I generally like solving anagrams.  
I am better at solving anagrams than the average person.  
Solving anagrams is easier for me than for the average person.

Note. The demographic information served to increase the plausibility that the rater would be aware of gender and therefore able to discriminate on its basis. The questions regarding anagram performance have proven to be useful covariates in previous studies (Berlin & Markman, unpublished data set.)
Appendix C: Anagrams

<table>
<thead>
<tr>
<th>Session 1</th>
<th>Session 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>EACHP</td>
<td>NELMO</td>
</tr>
<tr>
<td>ALSET</td>
<td>ANETLM</td>
</tr>
<tr>
<td>IKCTS</td>
<td>ILESMD</td>
</tr>
<tr>
<td>MSDOE</td>
<td>OLSPO</td>
</tr>
<tr>
<td>RCTIDE</td>
<td>LEESTC</td>
</tr>
<tr>
<td>RFERE</td>
<td>NIEDM</td>
</tr>
<tr>
<td>CETNRA</td>
<td>HRBOT</td>
</tr>
<tr>
<td>ERSET</td>
<td>IDFEL</td>
</tr>
<tr>
<td>EPSRE</td>
<td>RTNIPS</td>
</tr>
<tr>
<td>LPEESA</td>
<td>VEERL</td>
</tr>
</tbody>
</table>

Note. Anagrams were presented in a random order within each session.
Appendix D: Failure Feedback

Scoring Index

Please take a moment to review your score. Looking at all the information will help you to understand your score. Your score was 22.

<table>
<thead>
<tr>
<th>Your Score</th>
<th>22</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Score</td>
<td>39</td>
</tr>
<tr>
<td>Your Percentile*</td>
<td>34</td>
</tr>
</tbody>
</table>

Rater’s Note: Responses were mostly lacking creativity. Few of the more difficult answers were found.

*Percentile reports the percent of obtained scores at or below your score. The 34th percentile means that up to 65% of obtained scores were higher than 22 and 34% were equal to or lower than 18.

Rater: Luke Young
Appendix E: Attributions for Performance Outcome

Instructions: The purpose of the following questions is to understand how you viewed the initial anagram session. Your answers may be used to refine the way this experiment as run, as well as determine the optimal way to run experiments such as these in general.

Please be as candid as possible, and be assured that your responses will be kept confidential. These will NOT affect your scores in any way nor will your responses be available to the evaluator.

Attributions to Discrimination
- I feel the score I received was due to an unfair evaluation. (Question 11, from Major et al. 2003)
- I feel that my score was due to a prejudiced evaluator. (Question 12, from Major et al. 2003)
- I believe the quality of my responses deserved a better grade than what I received. (Question 14)
- I have reason to believe that gender discrimination played a role in the evaluation of my performance. (Question 26)

Internal Attributions
- Ability
  - The score I received was due to my abilities. (Question 1)
  - My score accurately describes my verbal ability. (Question 2)
  - My score is a reflection of my ability. (Question 19)
- Effort
  - My score was a reflection of my effort. (Question 20)
  - I would have done better if I had concentrated harder. (Question 21)

Perceived Control
- Past
  - I had complete control over the score I received on the last task. (Question 10)
  - I would improve on the previous task with practice (given the exact same task and the same evaluator). (Question 15; From Ruggiero & Taylor, 1997)
  - I know what I could have done to get a better score on the task I just completed. (Question 16)
  - I have no idea what I could have done to receive a better score on the last task. (Question 17, reverse scored)
  - There is nothing I could have done to improve my score on the previous task. (Question 24, reverse scored)
- Future
  - I know what to do to improve on the next session. (Question 7)
  - I have complete control over how well I do on the next task. (Question 18)
  - I have what it takes to improve on the next task if I so choose. (Question 22)
  - It is up to me, and me alone, to make improvements on the next set of problems. (Question 23)
  - I have control over my score on the next task. (Question 25)
Questions not included in reported results
  As a reflection of my verbal ability, my score accurately predicts my success in college. (Question 3)
  After receiving my score, I am worried about success in college. (Question 4)
  I enjoyed working on the anagrams in the first session. (Question 5)
  I generally enjoy working on anagrams. (Question 6)
  I expected to do well. (Question 8)
  I expected to do poorly. (Question 9)
  I believe the quality of my responses prevented me from getting a better grade. (Question 13)

Note. Responses were made on a 7 point likert-type scale with higher numbers indicating agreement. The numbers were presented in the order indicated in parenthesis following the question.
Appendix F: Session Protocol

Participants will read and sign the informed consent form.

Cover Story and Initial Instructions

The following instructions will appear on the computer screen and will also be read aloud by the experimenter. The experimenter will encourage all participants to read along.

“Thank you for your participation in today’s research session. Your responses will be used to develop the standardization of educational assessment tests. Today we are interested in several factors including the best order to arrange the sections of the test, as well as the relationship between personality, mood and testing. In order to study mood, you will be asked to complete questionnaires immediately preceding or following each subsection of testing. The questionnaires are meant to be answered as you are feeling at that particular moment.

“Today’s session will consist of two 10-item anagram sessions. Anagrams are sets of scrambled letters that you rearrange to form words. You must use all of the available letters in order to receive credit. All of the anagrams contain 5 to 7 letters and most have multiple solutions. For some there no solutions in which case you should simply write the word NONE and then move on to the next problem.

“Performance on tests such as these is being used increasingly in tests of verbal ability, and is strongly linked to success in college. Similar versions of the test you are about to take are scheduled to be included in college entrance exams within the next two years. Results of such tests are often used to determine scholarship distribution. In order to simulate this reality, a fund of $100 dollars is available to participants who do well. The participant from each timeslot who receives the highest score on the each session will be entered into a drawing for a chance to win the money.

“In order for this to occur, we must determine your scores before you leave today.

“For this reason, you will take the first session immediately. When you finish the first session, your responses will be sent to Luke Young, a graduate student who will evaluate them. He will give you a score, comment on your performance, and let you know if you will be entered into the drawing.

“Answers to this session will be scored in two ways: quantitative and qualitative. Your quantitative score will be judged based on the number of correct solutions you found compared to the number of possible solutions. Response quality will be determined by a trained judge and will depend upon creativity, perceived difficulty, and variety in your responses.

“Because we are interested determining the optimal order in which to present such tests, you will be presented with shorter sections of the test at two different times. You will receive feedback from the first session before you move on to the second session.

“You may now click CONTINUE to begin.”
Computer Questionnaire Order – Session 1

Demographic information
State Self-Esteem
MAACL (mood)
Anagram Session 1

Group Activity Filler Task

The experimenter will wait until all participants have finished. She will then instruct all participants to pull their chairs over to a table in the center of the room in order to work together on the next activity. She will then make the following statements:

“At this time your responses are being sent electronically to the rater. He will examine your responses and provide you with a score. While he is working on that, you will participate in a group activity. The purpose of this activity is to have you engage in visualization. You are to work together to create a list of as many of the buildings on campus as you can name. If you run out of buildings, you may flip the sheet over and work on one of the lists on the back. This activity is not scored, but rather its purpose is to cause you to visualize the locations as you list them. You have 7 to 10 minutes. Remember to work together.

“I am going to go upstairs and make sure that he received all of your scores. I will come back in about 10 minutes when your scores are ready.”

The experimenter will then leave the room for 7 minutes.

Discrimination Cue

During the filler task, the female confederate will make one of three statements:

Control. “You know, they must do a ton of these studies. I was in one just like this last week.”

Ambiguous. “You know, I have friends who were in this study last week and they told me that the guy doing the evaluating totally grades guys and girls differently.” (This is taken verbatim from Major et al, 2003 p. 224)

Overt. “You know, I know the guy who is doing the evaluating, and he’s totally prejudiced. There’s no way he’ll say a girl got the best score.”
After 7 minutes the experimenter will return and make the following statement.

“After 7 minutes the experimenter will return and make the following statement.

“The results from the first session are in. In a minute you can go to your computer and enter your first name so that he knows which computer to send the scores to. After you receive your scores you will be asked several questions about the first session. These questions are simply an assessment of your perception of the first task. They are used to evaluate the procedure and do not effect your score. Because many of the questions refer to the fairness of the evaluator, your specific responses will not be available to him. After those questions you will respond to more questions about mood and then take the second session. Both sessions will be evaluated by the same criteria. You still must use all of the letters, and list as many responses as you can identify.

**Same versus Different Rater Manipulation**

*Same evaluator condition.* “Your responses will be sent to the same evaluator. While he is reviewing your answers, you will respond to a few open ended questions about the session. You will receive your feedback before you leave.”

*Different evaluator condition.* “The only difference is that your responses for this session will be sent to a different evaluator. While he is reviewing your answers, you will respond to a few open ended questions about the session. You will receive your feedback before you leave.”

For both conditions the experimenter then reminded participants that the participant who received the highest score on the next session would be entered into the drawing. In the overt conditions, after the participants returned to their computers, the experimenter walked over to the male participant and informed him that he had received the highest score on the first session, and that he would need to fill out a form before leaving the lab so that he can be entered into the drawing. This was said loudly enough for all participants to hear. (This manipulation is also similar to the overt cue employed by Major et al., 2003)
Computer Questionnaire Order – Session 2
Feedback
Attributions Questionnaire
State Self-Esteem
MAACL (mood)
Anagram Session 2
Debriefing

Written Debriefing.
Participants were asked to type short responses to the following questions:

1. Please put into your own words what the experiment is about. State the purpose. Also, please state if you think it might involve any elements other than those explicitly stated.
2. Please tell us what you think of your feedback from the first session. Based on your performance, how accurate do you believe it was?

Verbal Debriefing
After all participants finished typing responses to the written debriefing questions, the experimenter asked the following questions.

1. “Would one of you volunteer to describe what the study is about? Just put into your own words what it’s about, why we had you do this, what are we trying to get out of it. Also, if you feel it involves any other elements than what was explicitly stated in the instructions, please include that too.”
2. “Could you please let me know what you thought about your feedback? Specifically, I’d like you to give me a sense of how accurate it was, especially based on your responses.” She will call on each of the participants to respond to the question.

General debriefing
After all participants responded to the verbal debriefing, the experimenter revealed the deception and the role of the confederates, debrief participants on the true nature of the study, thanked them for their time, and informed them that all participants would be entered into the drawing with equal likelihood of winning regardless of actual performance.