

The Pennsylvania State University

The Graduate School

College of the Liberal Arts

DELEGITIMIZATION AND WOMEN'S PERCEIVED EMOTIONALITY: "DON'T BE SO
EMOTIONAL!"

A Thesis in

Psychology and Women's Studies

by

Leah R. Warner

© 2007 Leah R. Warner

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2007

The thesis of Leah R. Warner was reviewed and approved* by the following

Stephanie A. Shields
Professor of Psychology and Women's Studies
Thesis Advisor
Chair of Committee

Theresa K. Vescio
Associate Professor of Psychology and Women's Studies

Karen Gasper
Associate Professor of Psychology

Nancy Tuana
Du Pont/Class of 1949 Professor of Philosophy and Women's Studies
Director, Rock Ethics Institute

Melvin M. Mark
Professor of Psychology
Interim Head of the Department of Psychology

*Signatures are on file in the Graduate School.

ABSTRACT

When a woman engages in an argument a common reaction is, “she shouldn’t have been so emotional.” In a series of two studies I look at the consequences that this reaction has on other individuals’ perceptions of women and on women who are themselves targets of being called emotional. In Study 1 I test the hypothesis that calling a person emotional is more delegitimizing than identifying that person’s argument as wrong, and that women will be more delegitized than men when this occurs. Delegitimization is defined as invalidating a claim in the eyes of an actual or implied reference group (e.g., Zelditch, 2000). The inability to properly control emotions, which is believed to invalidate a claim, is one of the most salient stereotypes of women in the West, and thus calling a woman emotional will be more believable than calling a man emotional. Participants (N=127) evaluated vignettes of two characters in which one (the observer) either calls the other (the target) emotional or disagrees with the target. Results revealed that when the observer disagreed with the target, male and female targets’ arguments were perceived similarly. However, when observers called targets emotional, male targets’ arguments were seen as more legitimate than female targets’. In Study 2 I focused specifically on the consequences that being called emotional has on women targets, and tested the hypothesis that, due to concerns about being delegitized, women would respond to being called emotional by being more preoccupied with their emotions than men. Results are discussed in terms of the negative consequences associated with preoccupation, such as emotion suppression.

TABLE OF CONTENTS

List of Tables	v
List of Figures	vi
Acknowledgements	vii
Chapter 1. INTRODUCTION.	1
Defining Legitimacy.	2
Emotion and the Negotiation of Legitimacy.	4
Gender, Emotion, and Legitimacy.	8
A Caveat About Generalizing These Predictions to All Women	12
Research Overview	12
Chapter 2. STUDY 1	14
Method	16
Results.	20
Discussion	26
Chapter 3. STUDY 2	31
Method	34
Results.	41
Discussion	49
Chapter 4. GENERAL DISCUSSION.	54
When Does Emotion Legitimize and When Does It Delegitimize?	55
Consequences of Delegitimization for Women.	59
Emotion, Legitimacy, and Structures of Gender Inequality	63
Does Emotion’s Relationship with Legitimacy Reach beyond Gender?	64
Conclusion	66
Endnotes	68
References	69
Appendix A: Study 1 Legitimacy Scales	78
Appendix B: Factor Analyses for Study 1 Emotion Scales	79
Appendix C: Correlations Between All Variables in Study 1	84
Appendix D: Study 2 Social Issues Questionnaire	87
Appendix E: Study 2 Handwriting Analyses	89
Appendix F: Correlations between All Variables in Study 2	90
Appendix G: Comparing Participants and Raters in Study 2	91

LIST OF TABLES

Table 1: Factor Loadings for the Measure of Target Emotional States When the Target is called Emotional	19
Table 2: Mean Perceived Legitimacy of Target’s Arguments as a Function of Participant Gender, Observer Gender, and Observer’s Evaluation	22
Table 3: Means and Standard Deviations of Potential Mediating Variables for the Effects of Observer Evaluation and Target Gender on Legitimacy of Targets’ Arguments	24
Table 4: Mean Attitude Strength as a Function of Target Gender and Feedback	42
Table 5: Mean Raters’ Assessments of Thoughts about Emotion and Wordiness as a Function of Target Gender and Feedback	44
Table 6: Mean Participants’ Assessments of Thoughts about Emotion and Wordiness as a Function of Target Gender and Feedback	47
Table 7: Mean Thought Intensity as a Function of Target Gender and Rater Type	49
Table 8: Factor Loadings for the Measure of Target Emotional States When the Observer Disagrees with the Target.	80
Table 9: Factor Loadings for the Measure of Observer Emotional States When the Target is called Emotional.	81
Table 10: Factor Loadings for the Measure of Observer Emotional States When the Observer Disagrees with the Target.	83
Table 11: Number of Participants per Essay Type by Gender and Feedback Type	84
Table 12: Simple Main Effects Comparing Participants and Independent Raters for the Feedback X Target Thought Focus X Rater Type Interaction	88
Table 13: Means and T-Tests for Handwriting Pretesting	89
Table 14: Correlations between All Variables in Study 2	90
Table 15: Simple Main Effects Comparing Participants and Independent Raters for the Feedback X Target Thought Focus X Rater Type Interaction	91

LIST OF FIGURES

Figure 1. Mean evaluation of the perceived legitimacy of the target's arguments as a function of observer's evaluation and target gender	21
Figure 2. Legitimacy of observer calling the target emotional mediates the relationship between target gender and legitimacy of the target's arguments	25
Figure 3. Raters' assessments of mean thoughts about emotion and wordiness as a function of target gender and feedback	45
Figure 4. Mean participants' assessments of thoughts about emotion and wordiness as a function of target gender and feedback	47

ACKNOWLEDGMENTS

This Thesis was funded in part by the Florence L. Geis Dissertation Award, awarded by the American Psychological Association Division 35, The Psychology of Women. Special thanks goes to Stephanie Shields, head of my dissertation committee, for her role in providing feedback and mentorship through out the development of this thesis. I also thank my doctoral committee, Karen Gasper, Terri Vescio, and Nancy Tuana for their valuable comments on the thesis. I also thank the undergraduate research assistants involved in collecting data: Halley Anolik, Karen Berman, Stacey Cohen, Danielle Conrad, Lindsay Filoleta, Moira Gallagher, Jennifer McBride, Lindsey Mitchell, Christy Pflieger, Rebecca Schuberth, Taryn Stoken, Betsy Van Noy, and Joy Walters.

CHAPTER 1: Introduction

In *The Cultural Politics of Emotion*, Sara Ahmed describes an incident in which a Middle Eastern professor in Canada spoke out against Canada's response to "the war on terror," and the media and government reacted by labeling her as so emotionally out of control that she was considered a terrorist threat herself:

What is significant about the attacks on Thobani is how personalized they became and how much they focused on delegitimizing her very right to speak. Her speech is described as the 'ranting and raving of a 'nutty' professor', as 'wingnut ravings,' as 'mean mouthed,' as 'sly and sick,' as well as 'hate-filled discourse'. Such attacks on Thobani's speech worked to exclude it from the register of legitimate speech by constructing her as motivated by a purely negative passion (2004, pg. 169).

In this event, the professor threatened her government by criticizing its decisions about an important and controversial issue. The response to this threat was strong and was aimed directly at undermining the legitimacy of her threat. Specifically, her comment was reduced to mere "emotionality," rather than as a result of careful or reasoned thought.

When a woman makes a challenging statement in an argument, a common reaction is, "she shouldn't have gotten so emotional." By attributing her arguments to "emotionality," an assumption is made that she is being unreasonable or inappropriate. Indeed, the inability to properly control emotions is one of the most salient and consistent stereotypes of women in the West (e.g., Fischer & Manstead, 2000). However, appropriately regulating one's emotions is a hallmark of skill in social interactions (Butler et al., 2003), a competence that men are believed to better achieve than women (Timmers, Fischer, & Manstead, 2003). Characteristics of emotion, at once, both seem to act in service of reason and against it. Gender, it seems, may tip the scale in one direction or the other.

In this thesis first I develop the argument that, because of emotion's paradoxical status of both serving reason and subverting it, biased perceptions of emotionality can function to

delegitimize individuals, and in particular, women. I examine the general hypothesis that calling an individual “emotional” functions to delegitimize him or her, especially if that person is a woman. I begin by defining legitimacy within a psychological context. Next, I highlight emotion’s relation to reason and legitimacy, and then why perceived emotionality functions to delegitimize women in particular. I then report two studies and conclude with a discussion of how findings relate to systems of structural inequality between men and women.

Defining Legitimacy

At the psychological level, legitimizing entails making a claim or a person seem valid in the eyes of an actual or implied reference group, while delegitimizing makes such a claim or person seem invalid (Berger, Ridgeway, Fisek, & Norman, 1998; Zelditch, 2000). A claim is seen as invalid, for example, if considered illogically or unjustly based. Research in procedural justice demonstrates that people respond positively to claims and procedures that appear to be rational and fair, because such qualities suggest that a claim or procedure has been said or implemented for a true reason and for the greater good (Elsbach, 2000). There are many aspects of a person or situation that contribute to whether a claim appears to be rational or fair, such as their position or status (Ridgeway, 2000) or the complexity of their arguments (Kelman, 2000). I focus, in particular, on how one’s emotional response is used as a source of information about rationality and fairness of claims.

Both Berger et al. and Zelditch (2000) define legitimization and delegitimization as operating within the context of a relevant group of people. They argue that legitimacy is determined by all members of the group – that both the advantaged and disadvantaged accept what is deemed “right” or “true.” For example, if a representative of upper management determines that an entry-level employee was being too sensitive and was overreacting when she

made a claim that she was sexually harassed, other employees may perceive that the claim is invalid. As a result, they will not treat the alleged harasser with caution nor will they try to punish him. Berger et al. (1998) additionally assert that if two people are struggling over legitimacy, legitimacy is contingent on the reactions of others who validate (or invalidate) the claim or action for the group and create the appearance that the claim is appropriate (or inappropriate). Thus, the “rightness” of the claim does not merely rest with the two individuals having the conversation. Rather, the legitimacy of the claim is contingent on whether it is considered “right” outside of the context of the conversation as well, including by those who have no investment in whether the claim is right or wrong.

This definition of legitimacy provides a framework for understanding delegitimization within the context of one person calling another emotional. When calling someone emotional, delegitimization is a process involving three contributors. The first contributor is the one who calls the other emotional, which I label the *observer*. The second is the one who is being called emotional, which I label the *target*. The third is the uninvolved observer or reference group, which I call the *third party observer*. If calling someone emotional is indeed delegitimizing, the third party observer should then view the target’s arguments as less legitimate when the observer calls him or her emotional. In addition, if the target is motivated to maintain his or her legitimacy, then he or she will react to being called emotional by trying to regain his or her legitimacy. It is the goal of the present research to evaluate the reactions that both the third party observer and the target have to the observer’s comment about the target’s emotionality.

With this understanding of legitimacy, one can take a closer look at why emotion can affect whether a claim is perceived to possess the rationality and validity necessary to be considered legitimate.

Emotion and the Negotiation of Legitimacy

There is a long tradition in Western cultures of construing emotions as being in opposition to reason (Averill, 1980; Ben-Ze'ev, 2000). This tradition assumes that emotions do not involve rational cognition. The historical term for emotions, passions (from the Greek word *pathe*), are considered changes that a person suffers, rather than an action that a person does (Averill, 1990). According to this view, emotions result from uncontrollable instinct and thus are animalistic, a property at odds with human thought (Lutz, 1990). This type of thinking about emotion is associated with common U.S. ways of describing the occurrence of emotion like “she was overcome with emotion.”

Previous work in psychology and other disciplines has suggested that this traditional belief about emotions is currently the dominant popular belief about emotions in the West (see Averill, 1990). However, recent work challenges this conventional dichotomization of emotion and reason (e.g., Hutson-Comeaux & Kelly, 2002; Shields, 2002; Warner & Shields, *in press*). This view emphasizes the critical role of emotion in confirming deep, authentic, justifiable feeling, or “humanity” (Morgan & Averill, 1992). Emotion can be briefly defined as a phenomenon containing behavioral, physiological, subjective, and cognitive components that together indicate one’s evaluation of the personal significance of a person or object in the environment (Oakley & Jenkins, 1996). Emotions, by communicating one’s reaction to a stimulus, indicate that a person is being consistent with his or her values and convey a person’s true intentions (Averill, 1990), which can either be reasonable and trustworthy or irrational and dangerous. As such, emotions are not inherently irrational, but they have the *potential* to be perceived as irrational if not appropriately regulated (Shields, 2005). In a discussion about the social construction of mental disorders, James Averill (1990) argues that individuals tend to

determine an emotional response is “disordered” when the emotion is dictated by ulterior or unrecognized motives, such as instinctual or reactionary responses. In contrast, individuals tend to determine an emotional response as “normal” when the emotion results from accurately assessing the situation and choosing the appropriate response to manage that situation.

My analysis of the role that emotion plays in such negotiations of legitimacy derives from this premise that lay theories of emotion view emotion as potentially dangerous to reason yet essential to humanity. Lutz (1990) argues that emotion’s association with the uncontrolled and the natural leads to deleterious consequences. I agree with her conclusion, but not from the standpoint that emotion is only associated with irrationality. Emotions may delegitimize when they threaten that the instinctual (and therefore irrational) nature of emotion takes over. At the same time, emotions may legitimize when they signal that something important is happening (Averill, 1990). The capacity to feel deeply is often seen as disconnected from or even beyond reason, but it is this deep feeling that indicates that one has opinions and cares about issues of importance. The individual without emotion is viewed as wooden or robotic, lacking values or opinions; the individual with too much, lacks self-control and clear thinking. This ability to negotiate one’s humanness is the basis for contesting emotion and thus what makes it powerful in asserting status and delegitimizing others who might challenge that status.

Research in psychology demonstrates how emotion’s paradoxical status can positively affect some individuals and negatively affect others. For example, a series of studies by Leyens and colleagues illustrates this point in demonstrating that prejudice toward an outgroup is reflected in beliefs about that group’s emotions (Leyens et al., 2000; Paladino, Leyens, Rodriguez, Rodriguez, Gaunt, & Demoulin, 2002). Leyens et al. examined attributions of *primary* and *secondary* emotions made by ingroup and outgroup members toward themselves

and toward the other group. They identified *primary* emotions as those commonly believed to be universal across both animals and humans and to appear early in life. *Primary* emotions include joy, sadness, anger, fear, disgust, and surprise. *Secondary* emotions, such as admiration, pride, nostalgia, remorse, are those, Paladino et al. note, that are social emotions commonly considered exclusively human. Across several ingroup/outgroup combinations and employing several different methods of obtaining emotion attributions, they find a consistent pattern. While there tend to be differences in the number of primary emotions attributed to each group, both groups selected more secondary emotions for their ingroup than for the outgroup. By assigning more secondary emotions to themselves the ingroup implies that they are somehow more developed, more human than the outgroup. Here, then, emotion's paradoxical status is divided among types of emotions: primary emotions are primitive, whereas secondary emotions are more connected to human rationality. The Leyens group's research underscores the power of emotion beliefs even in defining what it means to be human. In other words, emotion beliefs help to determine who deserves to be seen as valid, believable, or reasonable and who deserves to be seen as invalid, suspicious, or illogical.

As an initial step in examining the role of emotion in negotiating legitimacy, I first explore the role that perceived emotionality has on processes of delegitimization (as opposed to legitimization). Specifically, in the context of a conversation, I predict that by calling a target emotional an observer will delegitimize the target's arguments in the eyes of the third-party observer. In this context, "emotional" will be interpreted to mean that the individual is feeling and/or expressing emotion without appropriate reason. The term "emotional," at least in many mainstream American contexts, is typically associated with uncontrolled actions and irrationality. For example, Shields and Crowley (1996) found that when a target's responses

were described as “emotional” they were judged as less controlled and less appropriate than when those responses were described by specific emotion terms (e.g., angry). Because being emotional suggests irrationality, the third-party observer has “evidence” that the target’s arguments are not valid.

I predict that, in addition to the third-party observer’s reaction to the observer’s comment, the target will also react to being called emotional in that they will try to protect or regain legitimacy. I make the assumption that individuals are, in general, *motivated* to maintain their legitimacy, and thus individuals will respond to delegitimization by enacting behaviors to maintain or reinstate legitimacy. I base this assumption on the idea that in our everyday lives we seek to attain legitimacy in the eyes of others as well as ourselves (Jost & Major, 2001). In their review of the psychology of legitimacy, Jost and Major (2001) extensively cite the importance of legitimacy in everyday interactions. They argue that legitimacy is crucial to impression management and self-worth, in part because people value rationality, integrity and other characteristics associated with legitimacy (e.g., Bierhoff, Cohen, & Greenberg, 1986, as cited in Jost & Major, 2001). Thus, if being called emotional is related to delegitimization, individuals may be particularly motivated to alter that perception because delegitimization has strong negative social consequences, such as lowering status or power (Elsbach, 2000). In order to determine the appropriate response to the accusation, an individual needs to first examine what he or she is currently emoting. In particular, an individual may examine his or her emotional responses as a way to determine whether the accusation is warranted or not, and also to monitor behavior so as not to confirm the accusation. Thus, an initial response to being called emotional may include an increase in one’s preoccupation with one’s own emotional responses. I predict

that individuals will become more preoccupied with their emotional responses when being called emotional.

In sum, I predict that calling a target emotional will have several negative consequences. First, when an observer calls a target emotional, the target will be viewed as less legitimate in the eyes of the third-party observer than a target in a control condition where the observer simply disagrees with the target. Second, as compared to non-emotion feedback, calling a target emotional will lead to an increase in the target's preoccupation with his or her emotions. Notably, however, while I believe that calling any target emotional will affect all targets, as I describe in the next section there is reason to believe that consequences will be more negative for women than for men.

Gender, Emotion, and Legitimacy

Differing beliefs about men's and women's emotions may dictate whether perceived emotion leads to legitimization or delegitimization. Women are stereotypically associated with letting emotions take over and control thought or speech (e.g., Rosencrantz, Vogel, Bee, Broverman, & Broverman, 1968); men's emotions, on the other hand, tend to be stereotypically associated with competence (Timmers, Fischer, & Manstead, 2003). In this section I will discuss when gendered beliefs play a role in emotion perception and also evidence that, when beliefs do inform perception, stereotypes about women's emotions may lead to more negative consequences for legitimacy than stereotypes about men's emotions.

The ambiguity and complexity that occurs in everyday emotion provides ample opportunity for gendered beliefs to inform perception. Emotion in extreme situations provides a clear indication as to whether or not a person is expressing emotion in a way that signifies rationality and humanity. The idea of emotional appropriateness, for example, is the basis for

widely used concepts of “display rules,” the often tacit social rules directing when, how much, and which emotions should be expressed to others (Ekman, 1993). Such rules generally refer to clearly prescribed norms (e.g., be happy when given a gift; show sadness when hearing sad news). However, in many day to day situations, more than one interpretation is possible. Emotion is an aspect of human nature that is in a unique position. It is ongoing behavior that is visible, and should therefore have potential for “objective” measurement. As research on eyewitness testimony has shown, however, emotion’s visibility gives the observer a false sense of confidence in his or her assessment (Memon, Hope, & Bull, 2003). At the same time that emotion is visible, it is also fleeting and/or complex in everyday interactions. People have the capacity to express multiple emotions at once, to mask emotions, and to exaggerate other emotions. Thus, while there is consensus regarding clear departures from a norm (Hutson-Comeaux & Kelly, 2002), less extreme every day expressions of emotion allow more room for multiple interpretations (Gross et al., 2000).

Ambiguity both makes emotion perception subject to bias and its fleeting nature makes it difficult for assertions regarding its quality, quantity, and appropriateness to be countered with objective evidence (Shields, 2005). When there is ambiguity or complexity people have to look at diverse sources of information to guide interpretation of the expressive behavior. The situation is one source (e.g., funeral or birthday party); another are features of the target, such as stereotypes about the target’s propensity for emotion.

Stereotypes about men’s and women’s emotions seem to reflect an assumption that men are able to uphold emotion competence standards better than women. For example, Timmers, Fischer, and Manstead (2003) found that participants believed men are better than women in situations where sensitivity is required for competence (such as in nursing). Indeed, Timmers and

colleagues conclude, “for men, emotions still seem to be more associated with ability, with good social and emotional skills, whereas for women emotions remain linked to stereotypical femininity, that is, to their vulnerability, and thus to their loss of control and power” (p. 58). In other words, when men express emotions, it is believed that they do so out of competence whereas when women express emotions, it is believed they do so out of vulnerability.

Research on gendered perceptions of emotions supports the notion that men are believed to emote out of rationality more than women are. For example, Shields and Crowley (1996) generated scenarios depicting either a male (Brian) or female (Karen) responding to an emotionally-provoking situation. They found in people’s open ended responses to how they imagined the scenario that they “normalized” Brian’s response; for example, “I just imagined any average reaction (i.e., my own) if I found out that my car was stolen. I just imagined that he probably worked pretty hard for his car, and that he had taken care of it, so of course it would be upsetting.” On the other hand, they imagined Karen as letting emotions take over: “[I imagined] Karen at a parking lot crying hysterically because her car had been stolen. She lacked control and was too emotional for that particular situation.” Further evidence of the relative advantage men have over women is present in the literature on human tears. While flowing tears by anyone are considered a loss of control (Vingerhoets, Cornelius, VanHeck, & Brecht, 2000), men who cry in modest amounts are more positively evaluated (e.g., more likeable) than women who cry (Labott, Martin, Eason, & Berkey, 1990; Warner & Shields, *in press*). In part, the gender specificity of tear evaluation stems from a belief that the situation must have been serious for a man to cry (Labott et al., 1990).

If, as I argue, emotion beliefs help to determine who deserves to be seen as valid, believable, or reasonable and who deserves to be seen as invalid, suspicious, or illogical, and that

this quality of emotion helps to determine legitimacy, men and women may face different consequences. When faced with being called emotional, stereotypes of men's emotions as controlled and competent may temper this accusation for men, while stereotypes of women's emotion as uncontrolled or unreasonable will increase the third party observer's perception that women's arguments are not valid. As such, I predict that, while both men and women will be delegitimized when they are labeled as emotional, women will be more delegitimized compared to men.

In addition to facing negative consequences in terms of the observer's perceptions, I predict that women will be more likely than men to face negative consequences in terms of their own reactions as targets to being called inappropriately emotional. In other words, women may be more greatly affected as targets of delegitimization when emotions are the subject matter. Women's experience of being called emotional may entail more than general concerns about being seen as socially inappropriate. Women additionally may expect that an observer endorses the stereotype of women's emotionality, and as preparation for compensating for this expectation they may be even more strongly preoccupied with their emotions.

Because perceived competence is a necessary component of perceived success in the workplace (Yoder, Schleicher, & McDonald, 1998) and other interpersonal domains, women may be concerned at how to overcome such negative expectations. Because men do not have past history with discrimination based on their emotions, they may not be as preoccupied. As I discuss in Chapter 4, a future series of studies will address this phenomenon's relation to negative consequences for women.

A Caveat About Generalizing These Predictions to All Women

While it is possible that the above predictions extend to all women, it is important to caution against making this assumption. Evidence from studies that have addressed stereotype content more generally indicates that, across ethnic and national groups, women are believed to be more emotional than men (Shields, 2002). The degree to which this is so, however, may depend on the race of the target. Very little research concerns perceptions of women of color's emotions, but the study I did find suggests that stereotypes depend on the gender and race of the target. Landrine (1985) examined stereotypes held by White women and found differences in emotion stereotypes of Black women and White women. Black women were rated as significantly more "dirty," "hostile," and "superstitious" than White women, while White women were rated as significantly more "dependent," "emotional," and "passive." Landrine's study specifically suggests that negatively valenced words and words that signal uncontrolled anger are associated with Black women, while emotion more generally is associated with White women (at least from a White women's perspective). Such results put into question the degree to which such an effect can be generalized to non-White individuals. Thus, one must proceed with caution when discussing stereotypes about "women's" emotion. Thus, for the present studies I will only focus on the general category of "man" and "woman," knowing that any effects due to stereotyping may only extend to White, heterosexual individuals. In Chapter 4 I will discuss future research to test whether effects vary as a function of subordinate race and sexual orientation.

Research Overview

For Study 1, I will test the hypothesis that calling a target emotional is more delegitimizing than disagreeing with the target's arguments, and that women will be more

delegitimized than men when this occurs. Participants will read a vignette of two people interacting and then have participants rate the legitimacy of the target's arguments when he or she is called emotional or not. The purpose of this study is simply to establish that being called emotional actually results in delegitimization and also to show that it is more severe for women than for men.

For Study 2, I will test the hypothesis that labeling a target's response as "emotional," compared to nonemotion criticism, leads the target to be more preoccupied with emotion. I predict that this effect will be greater for female targets than male targets. This second study will provide grounds for future projects on how identifying a person's state or behavior as inappropriately emotional functions as a delegitimization tool and how that person resists such delegitimization.

CHAPTER 2: Study 1

In Study 1 I hypothesize that when the observer calls the target emotional, the target's arguments will be seen as less legitimate by a third party observer than when the observer disagrees with him or her, and that female targets relative to male targets will be especially delegitimized when called emotional. To test this hypothesis, participants, who will serve as the third party observers, will read vignettes that depict two individuals engaging in an argumentative conversation. Participants will read two vignettes. In one vignette, one individual in the vignette (the observer) evaluates the other individual (the target) as emotional. In the other vignette, the observer disagrees with what the target is saying. Thus, observer evaluation will be a within-subjects variable. In these vignettes I vary the genders of the target by changing the names of the characters. Target gender will be varied across participants, making target gender a between-subjects variable. After each vignette participants will rate the degree to which they perceive the target's arguments to be legitimate.

I chose disagreeing with the target as the control group for the following reasons. When choosing a control condition, one can do one of several different things: 1) pick a condition that is equivalent to the treatment condition in every way except for the treatment itself; 2) pick a condition that is as neutral as possible. For this particular study, I chose a neutral control condition. Characters were arguing in the vignette, and the vignette ended with one disagreeing with the other; hence, the key phrase simply extended the argument they were already having.

I included several exploratory subject variables into my design: observer gender and participant gender. I included observer gender because stereotypes of women indicate that they are more irrational than men (Haslam, 2006), suggesting that female observers' evaluations may not be seen as legitimate as male observers' evaluations. Thus, I wanted to see if results held

across male and female observers. I also included participant gender because own-gender bias might appear. I do not expect any effects due to participant gender, however, for the majority of studies on emotion stereotyping find that men and women tend to equally endorse stereotypes of men and women's emotionality (e.g., Robinson, Johnson, & Shields, 1998).

Hence, including the exploratory subject variables, the design for this first study will be 2 (gender of target: man or woman) X 2 (gender of observer: man or woman) X 2 (gender of participant: man or woman) X 2 (observer's evaluation: emotional and disagree with statement) mixed-factor design with repeated measures on the last factor. The dependent variable is the perceived legitimacy of the target's arguments.

I also included several potential mediators to explore why labeling the target emotional might affect perceptions of legitimacy. First, I included a scale measuring the perceived legitimacy of the *observer's* arguments. I reasoned that if the act of calling emotional affects the target's legitimacy, the individual who does the labeling may also be affected. If the act of calling emotional is seen as reasonable given the particular circumstances, the observer may be seen as legitimate in calling the target emotional. However, if the act of calling emotional is not seen as reasonable, the observer may be seen as less legitimate. Second, I included scales measuring both targets' and observers' perceived emotionality. Variations in the degree to which targets are perceived to be emotional in the vignette may affect the degree to which targets are delegitimized when they are called emotional. Specifically, if participants do not believe that a target is being emotional, he or she will be less likely to be convinced when the target is being accused of that action. Given that men are believed to be less emotional than women (e.g., Shields 2002), participants may be less likely to be convinced when a man is accused of being emotional. I also included the measure of the observers' perceived emotionality to see if the

observers' own perceived emotionality was related to their ability to delegitimize another via calling emotional.

Method

Participants. 127 undergraduate students (66 men and 61 women) were recruited from the Penn State psychology subject pool for either course or extra credit. 109 participants identified as European American, 6 as African American, 5 as Asian American; 5 as Latino/a American, 1 as Native American, and 1 as mixed racial origin.

Procedure. For this study I used a vignette-type paradigm in which participants read two vignettes about two individuals of equal power in a dispute over a work-related project. In the vignettes the observer and target each state their point of view on the issue at stake (one statement each). Then the observer either calls the target emotional or disagrees with the target. Participants received both the emotional and disagree versions of the vignette and I counterbalanced which version they received first. Participants were asked to think about the situations as if it were really happening and to provide feedback as if they were outside observers to this situation. I varied the gender of observer and target by varying the names used (e.g., Carla, Lisa, John, and Matt).

The topic of the vignettes were chosen based on several key factors: 1) neither character should clearly be more 'right' than the other; 2) the characters should be equally invested in the situation; 3) the topic should be relatively understandable to college students; and 4) the situation should be about work-related discussions, because being too emotional is generally seen as inappropriate in work contexts (e.g., Grandey, 2000), and therefore being called "emotional" will be seen as delegitimizing. The situations were set up such that each character presents an

opposing argument and then one of the characters either says, “You’re being emotional about this” or “I just don’t agree with you.” The two vignette situations are as follows:

Lisa and Carla are working on a group project together worth 40% of their grade. The major component of this project is a joint paper. The syllabus clearly states that late papers will be penalized 15% and the professor reminded everyone of that fact earlier in the week. An hour before their paper is due, Lisa brings her disk to a campus computer lab to print out their completed paper. A box appears on the computer screen saying that there was a fatal error in the disk. Carla does not have a copy of the paper. There might not be enough time for either of them to get home, search their computers for another copy of the paper, and make the deadline.

Lisa says, “I can’t believe this is happening! You told me you would bring a back-up copy.”

Carla says, “Well, if you hadn’t decided to wait until the last minute to print, then we wouldn’t be in this situation in the first place!”

Lisa then says, [either “You’re being emotional about this.” or “I just don’t agree with you”]

Kim and Amy are student officers for a martial arts club on campus. They planned an intercollegiate competition where they had to raise money to pay for judges, materials, and space. Now that the competition was over, they realize that they had spent \$350 over their budget, which means that they have to pay \$175 each out of their own pockets.

Kim says, “I cannot afford this. You handled the money and should have warned us that we were running over budget.”

Amy says, “There was just no way that I could have known. Those costs happened at the last minute.”

Kim then says, [either “You’re being emotional about this.” or “I just don’t agree with you”]

Vignettes were also counterbalanced such that I varied whether (1) the first or second character was the observer or the target; (2) participants received the vignette with the emotional or disagree condition first; and, (3) which character I asked participants to rate first.

After reading the vignette, participants were asked to evaluate the perceived legitimacy of the target’s arguments. This dependent variable consists of ten items that measure perceived legitimacy in terms of how valid the arguments seemed to be, for example, “[the target’s] opinion was rational” ($\alpha = 0.91$ for targets who were called emotional; $\alpha = 0.92$ for when the

observer disagreed with the target). All items were rated on a 1 (“strongly disagree”) to 7 (“strongly agree”) Likert-type scale. The same scale was also used to measure the exploratory variable of how legitimate the observer’s evaluation was perceived to be ($\alpha = 0.91$ for when the observer called the target emotional; $\alpha = 0.89$ for when the observer disagreed with the target). (For full versions of the scales, see Appendix A.)

Next, participants rated the target and the observer (both in the emotional and disagree conditions) on how much they perceived him or her to be feeling a variety of different emotions. Using a 1 (“not at all”) to 7 (“very much”) Likert-type scale, participants rated 11 different emotional states (annoyed, joyful, mad, happy, nervous, angry, scared, anxious, frustrated, afraid, and emotional). Each of the four scales (target in the emotional condition; target in the disagree condition; observer in the emotional condition; observer in the disagree condition) were subjected to principal axis factoring with varimax rotation.

Factor analysis for the target in the emotional condition yielded three factors with eigenvalues over 1.0. A scree test suggested a three factor solution which conformed to the three predicted dimensions, together accounting for 67% of the variance. The first factor, Fear/Anxiety (eigenvalue = 4.53), is comprised of 4 items (nervous, scared, anxious, and afraid, $\alpha = .88$). The second factor, Anger (eigenvalue = 2.13), is comprised of 4 items (annoyed, mad, frustrated, and angry, $\alpha = .82$). Happiness (eigenvalue = 1.50), is comprised of 2 items (happy and joyful, $\alpha = .89$). An item was included in a subscale if loaded higher than .30 on one factor, but not others. The “emotional” item was left out of the three factors because it loaded both on the Fear/Anxiety and Anger factors. Thus, it was tested separately from the other factors. Items for each subscale and factor loadings are listed in Table 1. Because the factor analyses were

similar for the other three scales (target in the disagree condition and observer in the emotional and disagree conditions), analysis of those scales can be found in Appendix B.

Table 1
Factor Loadings for the Measure of Target Emotional States When the Target is called Emotional

	Fear/ Anxiety	Anger	Happiness
Annoyed		0.73	
Joyful			0.95
Mad		0.84	
Happy			0.81
Nervous	0.82		
Angry		0.88	
Scared	0.86		
Anxious	0.56		
Frustrated		0.68	
Afraid	0.91		
Emotional	0.49	0.34	

NOTE: loadings below 0.30 are not shown.

Because factor analyses of my emotion measure suggested four components (Fear/Anxiety, Anger/Frustration, Happiness, and Emotional), I treated them as four separate variables in my analyses. These same components appeared for the other three scales (target in the disagree condition, observer in the disagree condition, and observer in the emotional condition) and thus I treated the emotion scale four separate variables for these scales as well.

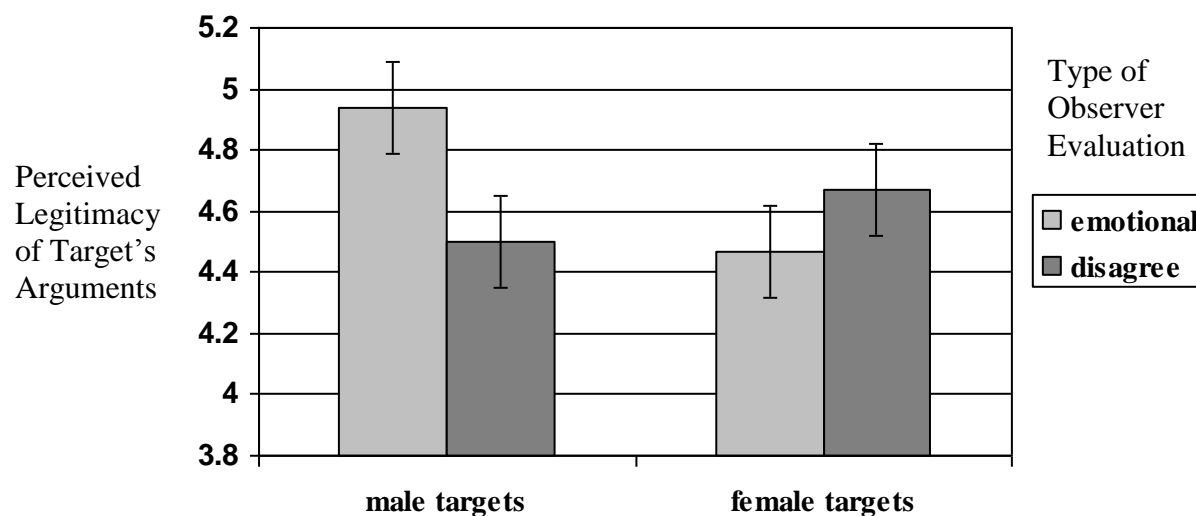
Results

The approach to my analyses are as follows. First, I report the test of my hypothesis. Second, I report the results of my exploratory variables specifically as they pertain to explaining the pattern of results obtained in tests of my hypothesis. Third, I report mediational analyses to test whether the exploratory variables mediate my effects. For a table of the correlations between all scales in this study, please see Appendix C.

Test of Hypothesis: Perceived Legitimacy of the Target's Statements. To test the hypothesis that when targets are called emotional their arguments will be delegitimized, especially arguments of female targets, I conducted a 2 (target gender: man or woman) X 2 (observer gender: man or woman) X 2 (participant gender: man or woman) X 2 (observer's evaluation: emotional and disagree with statement) mixed-factor ANOVA with repeated measures on the last factor. The dependent variable is the perceived legitimacy of the target's arguments. In partial support of the hypotheses, I found an observer's evaluation by target gender interaction, $F(1,119) = 3.85, p = .05, \eta_p^2 = 0.03$, such that both male and female targets' arguments were rated as equally legitimate when the observer disagreed with what they were saying ($M = 4.50, SD = 1.19$ vs. $M = 4.67, SD = 1.27$, respectively). However, male targets' arguments were rated as more legitimate than female targets' arguments when the observer called the target emotional ($M = 4.94, SD = 0.97$ vs. $M = 4.47, SD = 1.21, F(1,119) = 5.74, p = .02, \eta_p^2 = 0.04$). This effect occurred regardless of observer or participant gender. Contrary to my hypothesis, calling a woman emotional did not significantly delegitimize her compared to when the observer disagreed with her. Also, the male target's statements were actually seen as more legitimate when he was called emotional than when the observer disagreed with him ($F(1,119) = 5.74, p = .02, \eta_p^2 = 0.03$). Thus, in effect, calling a man emotional legitimized his statements,

rather than my expected finding that calling a woman emotional would delegitimize hers. (See Figure 1.) Note that the Bonferroni correction indicates that, because I made four comparisons, a significant p-value is $p = .013$. While my p-values approach this corrected alpha, they do not meet the standards and thus should be interpreted with some caution.

Figure 1. Mean evaluation of the perceived legitimacy of the target's arguments as a function of observer's evaluation and target gender



Only one other significant effect emerged from the ANOVA, and it was unrelated to the test of the hypothesis. Namely, an observer gender by participant gender by observer's evaluation interaction emerged, $F(1, 118) = 4.09$, $p = .03$, $\eta_p^2 = 0.04$. This effect demonstrated the effect that observer gender, participant gender, and observer's evaluation conjointly have on the perceived legitimacy of the target's arguments (regardless of target gender). To explain this 3-way interaction further, for each participant gender I conducted a 2 (observer gender: man or woman) X 2 (observer's evaluation: emotional and disagree with statement) mixed-factor ANOVA with repeated measures on the last factor. The Bonferroni correction for multiple comparisons indicated a significant p-value of $\leq .013$. For male participants, this analysis revealed no significant results. (See Table 2 for means.)

For female participants, a significant observer gender by observer's evaluation interaction occurred, $F(1,59) = 4.91, p = .03, \eta_p^2 = 0.07$, such that when the observer was female, there was a trend for the target's arguments in the disagree condition to be seen as less legitimate than the target's arguments in the emotional condition, $F(1,119) = 3.43, p = .07, \eta_p^2 = 0.06$. This effect suggests that female participants report that the female observer does a better job delegitimizing the target's arguments when disagreeing with targets rather than calling them emotional. In addition, there was a trend for female participants to report that the target's arguments in the disagree condition were less legitimate when the observer was female than when the observer was male, $F(1,119) = 4.30, p = .04, \eta_p^2 = 0.05$. In other words, female participants report that the female observer does a better job at delegitimizing than the male observer does when disagreeing with targets. Note that the analyses for these simple main effects does not meet the Bonferroni correction standards. Thus, these simple main effects must be interpreted with caution. No other significant effects emerged. (See Table 2 for means.)

Table 2

Mean Perceived Legitimacy of Target's Arguments as a Function of Participant Gender, Observer Gender, and Observer's Evaluation

	Female Participants		Male Participants	
	Female Observer	Male Observer	Female Observer	Male Observer
Obs. Evaluation: emotional	4.86 (1.27)	4.64 (0.84)	4.54 (1.11)	4.72 (1.17)
Obs. Evaluation: disagree	4.35 (1.33)	4.95 (1.19)	4.79 (0.96)	4.43 (1.31)

Note: Standard deviations are in parentheses; lower score = greater delegitimization

Exploratory Analyses. The main finding that emerged from the test of the hypothesis was that the male target's arguments are perceived as more legitimate than the female target's arguments when the observer calls the target emotional, while their arguments are seen as equally legitimate when the observer disagrees with the target. The next series of analyses are

exploratory attempts to understand why this effect emerged. I first examined each of the possible mediator variables by submitting them to a series of 2 (target gender: man or woman) X 2 (observer's evaluation: emotional and disagree with statement) mixed-factor ANOVA with repeated measures on the last factor. I only included target gender and observer evaluation because the test of the hypothesis indicated that neither observer nor participant gender contributed to effects pertinent to the hypothesis.

A potential mediator can be described as one for which a significant target gender by observer evaluation interaction emerges, such that the arguments of men in the emotional condition are judged as more legitimate than those of women in the emotional condition and men in the disagree condition. Only one variable fits these results: the perceived legitimacy of the observer's arguments. The target gender by observer gender interaction was significant, $F(1, 124) = 5.55, p = .02, \eta_p^2 = 0.04$. Simple main effects suggest the necessary pattern for mediation: Observers were rated as less legitimate when male targets were called emotional than when female targets were called emotional, $F(1, 124) = 7.88, p = .006, \eta_p^2 = 0.06$. Observers were also rated as less legitimate when male targets were called emotional than when they were disagreed with, $F(1, 124) = 9.47, p = .003, \eta_p^2 = 0.07$. Because I made two comparisons, the Bonferroni correction suggests that alpha levels should be at 0.025 to be considered significant. These two results meet the standard. All other variables either produced no significant effects or a main effect for Observer evaluation. Please see Table 3 for a list of all variables, means, and significant effects.

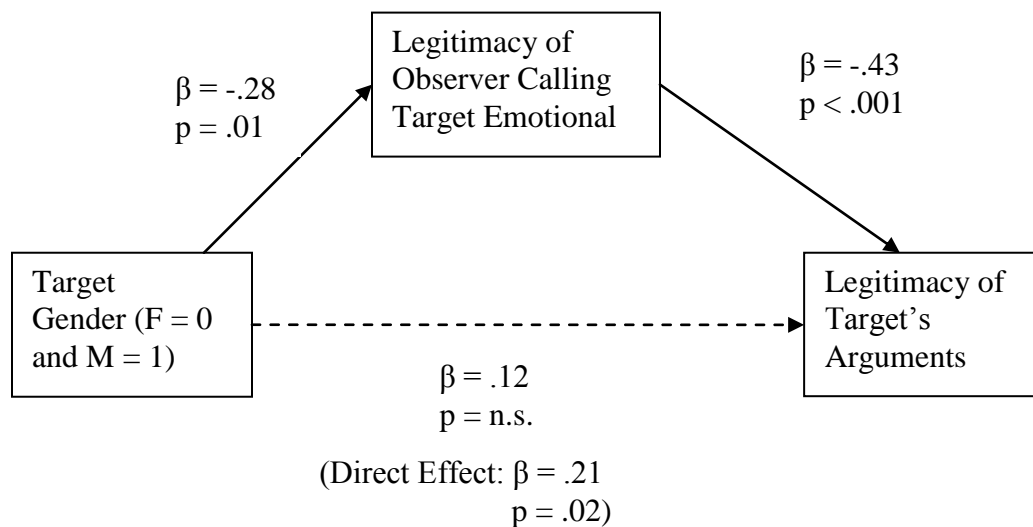
Table 3
Means and Standard Deviations of Potential Mediating Variables for the Effects of Observer Evaluation and Target Gender on Legitimacy of Targets' Arguments.

Variable	Obs. Evaluation: Emotional		Obs. Evaluation: Wordy		Significant Effects
	Male Targets	Female Targets	Male Targets	Female Targets	
Legitimacy of Observer's Arguments	3.85 (1.13)	4.42 (1.13)	4.43 (1.02)	4.37 (1.02)	Observer Evaluation Observer Evaluation x Target Gender
Target <i>Fear/Anxiety</i>	4.58 (1.37)	4.62 (1.41)	4.84 (1.22)	4.59 (1.36)	none
Target <i>Anger/Frustration</i>	5.37 (1.20)	5.62 (1.14)	5.56 (0.95)	5.55 (1.35)	none
Target <i>Happiness</i>	1.38 (0.65)	1.54 (0.87)	1.52 (0.82)	1.63 (1.01)	none
Target <i>Emotional item</i>	4.95 (1.63)	5.27 (1.49)	5.13 (1.40)	5.30 (1.43)	none
Observer <i>Fear/Anxiety</i>	4.16 (1.59)	3.95 (1.47)	4.70 (1.29)	4.20 (1.43)	Observer Evaluation
Observer <i>Anger/Frustration</i>	5.23 (1.23)	5.37 (1.38)	5.72 (1.14)	5.66 (1.04)	Observer Evaluation
Observer <i>Happiness</i>	1.48 (0.74)	1.54 (0.73)	1.43 (0.78)	1.67 (1.05)	none
Observer <i>Emotional item</i>	4.86 (1.79)	4.59 (1.73)	5.25 (1.48)	4.92 (1.66)	Observer Evaluation

Note: "Significant Effects" are any effects that emerged at the $p < .05$ levels. Standard deviations are in parentheses.

Because legitimacy of the observer's arguments emerged as a potential mediator, I followed the recommendations of Baron and Kenny (1986) and conducted a series of multiple regression analyses. I specifically focused on the extent to which this potential mediator explains the relationship between target gender and legitimacy of the targets arguments when the target is called emotional. Indeed, I found evidence of successful mediation. (See Figure 2.) Note that target gender was treated as a dummy variable (coded as 0 = Female and 1 = Male).

Figure 2. Legitimacy of observer calling the target emotional mediates the relationship between target gender and legitimacy of the target's arguments.



Target gender predicts the legitimacy of the observer's evaluation of the target as emotional such that when the target is male, the observer's evaluation is seen as less legitimate than when the target is female ($\beta = -.28$, $t(125) = -2.50$, $p = .01$). The legitimacy of the observer's evaluation, in turn, predicts the legitimacy of the target's arguments, such that the more legitimate the observer is perceived to be, the less legitimate the target is perceived to be ($\beta = -.43$, $t(125) = -5.28$, $p < .001$). Target gender directly predicts the legitimacy of the target's arguments, such that the male target's arguments are perceived to be more legitimate than the female target's arguments ($\beta = .21$, $t(125) = 2.41$, $p = .02$). However, when legitimacy of observer calling the target emotional is entered as a mediator, this relationship becomes non-significant ($\beta = .12$, $t(125) = 1.50$, $p = \text{n.s.}$), suggesting a successful mediation. Sobel's test confirmed a successful mediation, $z = 2.36$, $p = .02$. Looking at these mediational analyses suggests one potential explanation for my results. One reason why the male target may have been perceived to have more legitimate arguments than the female target is because participants did not believe the observer when he or she called men emotional.

It is important to note that when I tested for mediation in the condition in which the observer disagreed with the target, the mediation was not significant, in that target gender did not predict the legitimacy of the target's arguments. Thus, my results only extend to the condition when the observer called the target emotional.

Discussion

Results for Study 1 revealed partial support for my hypotheses. I did indeed find that the legitimacy of male and female targets' arguments were perceived differently when they were called emotional, and consequences were more favorable for men than they were for women. However, contrary to my hypotheses, I did not find that targets, regardless of gender, were seen as less legitimate when called emotional. In fact, I found that men were seen as *more* legitimate when the observer called them emotional than the observer disagreed with them. Mediation analyses show that participants believed that male targets were more legitimate than female targets because they perceived the observers' arguments to be *less* legitimate when calling a male target emotional. Presumably, participants believed that there was less validity to calling men emotional than women emotional. Interestingly, target and observer emotionality did not emerge as mediators, suggesting that 1) targets' degree of perceived emotionality did not contribute to how legitimate participants' arguments were perceived to be; and 2) observers' degree of perceived emotionality did not contribute to how legitimate participants' arguments were perceived to be. I will discuss each of these unexpected results in turn.

There are several potential reasons why women were not delegitimized when called emotional. First, perhaps the term "emotional" was not as damaging as originally expected. Rather, the observer should have said, "you are *just* being emotional," to more explicitly emphasize that the target was acting out of emotion only and not reason. Second, it may have

been more delegitimizing if the observer implied that the target's emotionality interfered with his or her reasoning or competence at the task that caused the argument in the first place (in one case, the budget and in the other, the group paper). For example, implying that a person was not capable of balancing a budget because she is too "sensitive" or emotional is more damaging to how legitimate she is perceived than simply being emotional in an argument about the budget. Third, perhaps the topic of the arguments between the characters were not perceived to be important enough for participants to make a judgment about delegitimization. In other words, the topical matter may have been perceived as trivial or silly, and thus the participants did not form a strong enough impression of the targets to make an evaluation of them. The data do not clearly show which, if any, of these potential reasons are most likely to have occurred. Thus, in future work I will test the extent to which these different possibilities contributed to the finding that women were not delegitized.

But then why were men *legitimized*? While expectations about women's emotionality may not have negatively affected the legitimacy of their arguments in this vignette, expectations about men's emotionality might have positively affected the legitimacy of their arguments. As discussed in the introduction, when men express emotions, it is believed that they do so out of competence whereas when women express emotions, it is believed they do so out of vulnerability (e.g., Timmers et al., 2003). It is possible that beliefs about men's competence in expressing and feeling emotions can explain why men's arguments were seen as more legitimate when called emotional than when disagreed with. Specifically, it may have been difficult for participants to imagine men as "emotional" because it did not fit with their expectations about how men typically behave. As a result, compared to the observer's absurd evaluation, the male target's arguments may have seemed *more* legitimate. Because it is more expected for women to

be emotional, the observer's comment may not have been perceived as absurd, and thus women did not receive the boost in legitimacy that the men received. Thus, men's perceived competence helped him instead of women's perceived incompetence hurting her.

Key to this explanation is that men and women do not have to necessarily differ in perceived *amount* of emotionality for men to be seen as more legitimate than women. Men and women could differ, instead, in the reasons believed to be behind their emotionality. Namely, it is possible that men were believed to be expressing emotion for a rational reason, whereas women were believed to be expressing the very same amount of emotion for irrational reasons. In other words, men and women may be expected to differ in the reasons behind their emotionality rather than simply the *amount* of emotionality. Thus, the reason why the scales measuring amount of emotionality did not contribute to results may be because these scales measured amount instead of the reason for the emotion (i.e., whether irrational or rationally-based).

Future work will examine these potential explanations for unexpected results. In order to test the delegitimizing effect that being called emotional has on women, I will begin by conducting further vignette studies. Specifically, I will test different versions of the vignettes to see which is more likely to be delegitimizing for women. Also, I will include measures that better examine the idea that men and women are perceived to differ in the reasons behind their emotionality (e.g., competent or not) rather than merely questions about their degree of emotionality. And finally, as with any phenomenon, it is important to examine the extent to which findings are found using different methods. One method particularly relevant to examining the delegitimizing effects of calling someone emotional would be one in which participants perceive an actual interaction between two individuals instead of merely reading

about them. This context would complement the vignette method. An advantage of vignettes is that they are ambiguous and leave room for interpretation, which means that biases are likely to influence perceptions of characters in vignettes. As a result, it should be easier to detect biases when using the vignette method. However, one drawback to using vignettes is that they instruct the individual to imagine the situation rather than experience it. Thus, the individual reading a vignette may not feel as involved or persuaded by what the character does or says. If participants were to perceive characters in real time they would perhaps be more psychologically involved. In sum, in future work I will vary the method I use to measure participants' perceptions of target legitimacy.

In Study 1 I focused on the effect that the observer calling the target emotional has on the third party observer. In Study 2 I turn to looking at the effect that the act of calling emotional has on the target him or herself. I had originally thought that men, like women, would react to being called emotional by being preoccupied when called emotional more so than non-emotion-related negative feedback. However, if, as suggested by Study 1 results, men are generally not negatively affected by being called emotional, perhaps their responses to being called emotional would not differ from other types of negative feedback. Namely, I now expect that women, but not men, will react to being called emotional by being more preoccupied with their emotions.

In addition, because I found that observer gender affected the perceived legitimacy of the target, at least for female participants, it is necessary to specify the gender of the individual who calls participants emotional in Study 2. Results suggest that, relative to several other conditions, female participants are less convinced by female observers when they call targets emotional. Thus, I will specify in study 2 that the observer is male. The alternative to specifying the gender of the individual who calls participants emotional is to leave the gender ambiguous. However,

leaving the gender ambiguous does not ensure that participants will not ensure a gender.

Research has found that individuals (at least in the American samples used) automatically sex categorize any person they perceive (Blair & Banaji, 1996; Brewer & Lui, 1989). Individuals have great difficulty imagining a person without assuming a gender for that person (Ridgeway & Correll, 2002), which means that merely leaving gender of the observer ambiguous would not eliminate effects due to observer gender. As a result, I will specify that the gender will be male in order to standardize how the participants conceptualize the individual who calls them emotional.

CHAPTER 3: Study 2

The results of Study 1 indicated that men were given relative advantage to women when they were called emotional. Given these results, the purpose of Study 2 is to examine how, relative to men, women would react to being called emotional. In this study I test the general hypothesis that, when criticized for being “emotional,” women will respond by being preoccupied with their own emotions to a greater degree than men will, and to a greater degree than those who do not receive emotion-related feedback. Originally, I had predicted that all participants would be preoccupied with their emotions when called emotional. Results of Study 1 suggest, however, that only women will respond by being preoccupied with their emotions. I thus focus specifically on the consequences that women in particular experience when being called emotional.

The independent variables for this study are similar to those of Study 1: target gender and calling a target emotional or not. I compare calling someone emotional to non-emotion related criticism, namely, saying to someone that he or she is too “wordy.” I use non-emotion related criticism as the comparison to rule out whether women’s heightened focus on emotion is simply a result of receiving negative feedback. Indeed, previous research has found that negative mood, one consequence of negative feedback, can lead to self-focused attention (Silvia & Abele, 2002) and attention to one’s own negative emotions in particular (Niedenthal, Halberstadt, & Setterlund, 1997). Thus, by using non-emotion related feedback I will test the hypothesis that being called emotional will lead to a stronger preoccupation with emotion than receiving negative feedback alone. I use the term “wordy” because writing an excessive number of words in an essay is typically not included in lists of known stereotypes of women (e.g., Schneider, 2005).

In addition, I will test the hypothesis that those who are called emotional will focus on their own emotions more than those who are called wordy will focus on their own wordiness. Because women are stereotypically believed to be more emotional than men, and because being called emotional is potentially linked to delegitimization, women may be especially likely to focus on emotion feedback when being called emotional. However, because being called wordy is not known to be stereotypically linked to women, it would not be expected to have the same meaning for women as being called emotional does.

Stated in operationalized terms, the hypotheses are as follows: 1) women who are called emotional will be more preoccupied with their emotions than men who are called emotional; 2) women who are called emotional will be more preoccupied with their own emotions than women who are called wordy will be with (a) their own emotions and (b) their own wordiness. Thus, the design for this study is a 2 (feedback: emotional or wordy) X 2 (target gender: man or woman) X 2 (target thought focus: emotions or wordiness) mixed design with repeated measures on the third factor. The dependent variables for this study are the degree to which targets become preoccupied with their own emotions and the degree to which targets become preoccupied with their own wordiness.

Using ANOVAs and follow-up tests, I will test these hypotheses by performing the following a priori comparisons. For Hypothesis 1, I will compare women versus men who are called emotional in terms of the degree to which their thoughts were about their emotions. For Hypothesis 2, Part A, I will compare women called emotional with women called wordy in terms of the degree to which their thoughts were about their emotions. And finally, for Hypothesis 2, Part B, I will compare women called emotional in terms of the degree to which their thoughts

were about emotions to women called wordy in terms of the degree to which their thoughts were about wordiness.

To test the hypothesis for this study, I performed a high-impact study in which participants wrote an essay about an issue that is important to them and then were given predetermined feedback on this essay indicating either that they did a poor job because they were “emotional” or that their essay was “wordy.” Because, as Study 1 indicated, gender of the person giving the predetermined feedback may contribute to the effects, I led all participants to believe that the person providing critique of their writing was male. To compare preoccupation with emotions versus wordiness, participants engaged in a thought-listing task adapted from Showers (1992). In this task, participants listed individual thoughts which were later coded by independent raters on specified criteria. One of those criteria is the degree to which the participant was thinking about the feelings that he or she felt while writing the essay. Another is the degree to which the participant was thinking about how well-written his or her essay was. Ratings for these two questions served as the dependent variables.

In addition to independent raters, participants also assessed their thoughts on these questions in order to provide supplementary evidence about the nature of the thoughts. Because emotion is always an inferred construct (i.e., it is an internal state), cross-referencing multiple measures of emotion increases researchers’ chances of correctly identifying emotions (e.g., Larsen & Fredrickson, 1999). Any differences between the independent raters’ and participants’ assessments should be minimal because both will rate the thoughts on the same questions.

If independent raters’ ratings are different, they should only differ in terms of how intense thoughts are across conditions. Because independent raters only have access to the written responses, and not participants’ actual thoughts, independent raters may not detect the

strength of the thought or feeling that the participants have. For example, a participant may say “I am angry,” and on paper the statement might not clearly signal the emotion intensity that participants feel. Participants may also differ from raters because participants are more invested in the task. Because participants will have just received negative feedback, they may be motivated to defend themselves. These intense reactions may, in turn, lead them to assess the thoughts as more intense than independent raters would. Independent raters are not expected to differ from participants within condition; rather, it is expected that they would only differ in terms of overall perceived strength of thought. That said, independent raters may provide insight that participants cannot because raters are trained to assess thoughts via a structured set of rules.

It is also important to note here that I am specifically making predictions about the average intensity rather than frequency of thoughts about feelings compared to thoughts about wordiness¹. Intensity, as opposed to frequency, is the more relevant dimension to measure because preoccupation may not lead individuals to think many distinct thoughts. Following Showers’ technique, participants are instructed to write down each *distinct* thought in the thought-listing task. Preoccupation may, however, lead a person to think the same thought over and over again. Rumination and worry, for example, are characterized by repeating the same negative thoughts (Segerstrom, Stanton, Alden & Shortridge, 2003). Thus, the *degree* to which individual thoughts are about emotion or wordiness may better represent preoccupation than the number of *distinct* thoughts that are about emotion or wordiness.

Method

Participants. 122 undergraduate students were recruited from the Penn State psychology subject pool for either course credit or extra credit. Six participants were excluded from analyses because they were suspicious of the feedback and two participants were excluded for failing to

follow directions. The final sample of 114 individuals (58 women and 56 men) consisted of 94 participants who identified as European American, 1 as African American, 5 as Asian American; 3 as Latino/a American, 4 as mixed racial origin, and 7 who declined to provide racial identity. The mean age for this sample was 19.02 years ($SD = 1.25$).

Procedure. Participants first wrote an essay about a social issue that was important to them and then received written predetermined feedback on this essay, either in the form of being called emotional or in the form of being told that their essay was “too wordy.” After receiving feedback participants completed a questionnaire that contained the dependent measures and the individual differences measure. Participants were tested in groups of 6-8 people.

1. *Cover Story.* Participants were told that they were participating in a study about social interactions. They were told that they would write an essay and that the experimenter would then give their essay to another participant in the room. That participant would then evaluate their essay and give them feedback.

2. *Argument Task.* Participants were told to indicate their opinions on three social issues informally pretested as important to college students: same-sex marriage, fraternities and sororities, and George Bush’s quality as president. All descriptions of the social issues were accompanied by a brief description of the issue. Participants first indicated how they stand on each issue and then indicated how strongly they feel about their stance. (See Appendix D.) I measured attitude strength as a composite of four common attitude strength dimensions related to feelings toward an issue: centrality, certainty, importance, and extremity (Pomerantz, Chaiken, & Tordesillas, 1995). Participants rated their attitude strength using a 7-point Likert scale (1 = not at all; 7 = very much) for each of the four topics: for same-sex marriage, $\alpha = 0.73$; for fraternities and sororities, $\alpha = 0.82$; and for George Bush’s quality, $\alpha = 0.83$. After participants rated the

issues, they were asked to write a short essay about the issue that they had rated as having the strongest opinion about. Participants were given the choice in order to maximize the likelihood that they would feel strongly about their essays. After spending approximately 5 minutes writing the essay, experimenters stopped the participants, collected their essays and passed the essays out to other participants.

3. *Predetermined Feedback.* All participants were led to believe that they were actually receiving feedback from each other. While they all did write feedback, participants did not actually receive feedback given by another participant. Instead, the experimenter switched the feedback sheet with a predetermined feedback sheet when participants' essays were returned to them. The feedback was structured such that participants were to give a letter grade and to write a one sentence explanation of their evaluation. The predetermined feedback sheet always indicated that participants received a C- as a grade and the sentence varied on whether it referenced wordiness or emotion.

“My main comment is that you shouldn't have been so wordy”

“My main comment for you is that you shouldn't have been so emotional.”

The feedback was always hand-written in styles of handwriting that were pretested as most likely to be a man's handwriting. A pool of ten individuals, whose handwriting had been identified by the experimenter as having masculine features was asked to write the two feedback statements in their own handwriting. The one specification given for this task was that the handwriting should be legible. Each sentence from each individual was rated on a 1 (“Not at all”) to a 5 (“Very much”) scale: “How likely is it that this is a man's handwriting?” and “How likely is it that this is a woman's handwriting?” by 28 undergraduate students from the Pennsylvania State University who completed these ratings on a volunteer basis. The two questions for each

piece of handwriting were subjected to paired samples t-tests and the five handwriting samples that were rated as most likely to have come from a man and least likely from a woman were chosen. The top five pieces of handwriting from both the “emotional” sentence and from the “wordy” sentence were written by the same 5 individuals. (See Appendix E for t-tests and means for each of the top five handwriting samples.)

It is important to note that content of feedback may influence the degree to which the handwriting is judged to belong to a man or a woman, regardless of the handwriting style. It may be the case that calling someone emotional is an act that participants believe to be something that one gender would be more likely to do than the other. To test this, I averaged ratings on the two questions (i.e., how likely written by a woman; how likely written by a man) for each handwriting sample for the wordy statement and for the emotional statement. I then subjected averaged ratings to a 2 (content of statement: wordy and emotional) X 2 (type of handwriting question: how likely from a man and how likely from a woman) repeated-measures ANOVA. I found a main effect for type of handwriting question such that, regardless of the content of the statement, the statement was perceived to more likely be a man’s handwriting ($M = 3.90$, $SD = 0.55$) than a woman’s handwriting ($M = 2.56$, $SD = 0.52$), $F(1, 27) = 68.17$, $p < .001$. This main effect was qualified by a content of statement by handwriting question interaction, such that the emotional statement was more likely to be coming from a woman ($M = 2.67$, $SD = 0.50$) than the wordy statement ($M = 2.45$, $SD = 0.56$), $F(1, 27) = 5.66$, $p = .03$. However, the degree to which the emotional ($M = 3.81$, $SD = 0.58$) and wordy statements ($M = 3.99$, $SD = 0.52$) were perceived to be coming from a man did not significantly differ from each other, $F(1, 27) = 3.17$, $p = \text{n.s.}$ It is crucial to highlight that the main effect for handwriting question persisted within both types of content (emotional, $F(1, 27) = 67.73$, $p < .001$; wordy, $F(1, 27) = 41.94$, $p < .001$). Thus, even

though there is a slight tendency for the emotional statement to be evaluated as more likely to have come from a woman than the wordy statement, both statements are still seen as significantly more likely to have come from a man than a woman. Thus, I conclude that the handwriting is sufficiently masculine for the present study.

Participants were randomly assigned to receive feedback in one of the five selected writing samples. Participants were told to begin filling out the questionnaire immediately after receiving their feedback.

4. *Preoccupation Measure.* In order to measure preoccupation, participants filled out a thought-listing technique adapted from Showers (1992) in which participants list their current thoughts. In this technique, participants themselves determine what constitutes a thought, and write each individual thought in a series of blank boxes. Twelve blank boxes (each 7 in. X 1.5 in.) were provided in a 2-page booklet. Participants were instructed to continue to add thoughts until they had added all currently occurring thoughts.

Independent raters assessed the thoughts on four categories using a 1 “Not at all” to 7 “Very much” scale. Categories were designed so that any thought would fit into at least one category. Categories referred to the degree to which thoughts were (1) about emotions (“I was thinking about what I was feeling or feelings I expressed while writing my essay”), (2) about wordiness (“I was thinking about the degree to which my essay was well-written), (3) about feedback in general (“I was thinking about the feedback that I received after writing my essay”), or (4) were unrelated to feedback (“I was thinking about something unrelated to my essay and/or the feedback I received”). The former two categories served as the dependent variables, while the latter two categories were used to ensure that all thoughts could be categorized with respect to their relevance to the feedback.

Three independent raters who were unaware of the hypotheses and conditions followed a standardized set of decision rules in order to maintain consistency². The raters answered four questions for each of up to 12 thoughts, for a total of 48 questions per participant. Interrater reliability was established via the consistency approach, which makes the assumption that the raters do not need to employ exactly the same meanings of the scales used for rating; rather, each judge needs to be consistent in classifying the phenomenon (Stemler, 2004).

In order to establish consistency, coders were given the following instructions for coding:

1) In order for a thought to count as *about emotions*, the thought has to indicate that the person was *reflecting* about his or her emotional state. The thought does not fit this category if it merely demonstrates that the individual is having an emotional response (such as using exclamation points (!!) or deep pen marks) or if they merely repeated feedback without reflecting on it (e.g., “I received feedback that I was being emotional”). The thought only fits this category if the person was thinking about emotions (e.g., “I don’t think I was being that emotional”).

Independent raters estimated the *degree* to which thoughts reflected emotions by estimating what proportion of the thought concerned thinking about emotion (i.e., was the entire thought a reflection about their emotions, or was only a portion of it relevant to emotions?). 2). In order for a thought to count as *about wordiness*, the thought has to indicate that the person was *reflecting* about how well-written his or her essay was. The thought does not fit this category if it merely demonstrates that the individual’s thoughts were poorly written (i.e., the thoughts themselves were wordy) or if they merely repeated feedback without reflecting on it (e.g., “I received feedback that I was being wordy”). The thought only fits this category if the person was thinking about how well written their essay were (e.g., “maybe I did say too many things”). Independent raters estimated the *degree* to which thoughts conveyed reflection on how well-

written the essay was by estimating what proportion of the thought concerned thinking about well-written. 3). Any thought that is related to the feedback would score high on the general feedback item, including items that are not explicitly about emotion or wordiness, such as thinking about length of time given to write the essay, “we weren’t given enough time to write our essays.” 4). Any thought not related to the feedback scored in the fourth category of unrelated to feedback.

All raters compared their assessments several times throughout evaluating the participants and then resolved any major discrepancies, defined as any single coding for which one or more raters deviated 2 or more points from each other. The few major discrepancies that emerged were resolved by clarifying decision rules.

After all participants’ thought lists had been coded and major discrepancies resolved, interrater reliability was determined via Cronbach’s alpha for each of the 48 variables. Chronbach alpha values of 0.70 are considered acceptable for consistency estimates of interrater reliability (Barrett, 2001 as cited in Stemler, 2004). All variables had an alpha of 0.70 or greater, indicating sufficient interrater reliability, and 46 out of the 48 variables had an alpha of 0.80 or greater. Due to the high alphas for each variable, raters’ responses were averaged to create a composite score for each question.

The dependent variables, degree to which the thought was about emotion and the degree to which the thought was about wordiness, were created by averaging all relevant thoughts to create one score per dependent variable for each participant. I decided *a priori* to only include those thoughts that were relevant to the feedback that participants received. I expected that participants would have a number of thoughts completely unrelated to the task, such as, “I am hungry.” In order to refrain from artificially deflating the results, I did not include those thoughts

that were scored as 5 or above on the category that assessed degree of irrelevancy. Separate composite scores were created for participants' self-ratings and independent raters' coding.

Participants also assessed their own thoughts using the same four categories that the independent raters used. Unlike the independent raters, however, participants were left to decide themselves what constituted a high or low score on each category. Participants assessed their thoughts after the thought-listing exercise.

Results

Results are divided into four sections. The first section summarizes participants' essay choice and attitude ratings. The second provides descriptive information about the content of thoughts in reaction to feedback. The third reports tests of the hypotheses. The final section describes participants' assessments of their own thoughts and provides a comparison of participants' self-assessments with the independent raters' assessments. For a table of correlations between all scales, please see Appendix F.

Essay Choice and Attitude Ratings. Participants were slightly more likely to pick same-sex marriage as their essay topic (42.2%) than George Bush (31.0%) or fraternities and sororities (26.7%). Although random assignment should ensure that essay topic was unrelated to experimental condition, to test whether essay type differed among conditions, I performed a 2 X 2 chi-square analysis on participant gender (man or woman) and feedback (emotion or wordy) for each of the three essay topics for each topic. None were significant, indicating no relation between topic chosen and assignment to feedback condition (same sex marriage, $\chi^2 = (1, N = 49) = 0.85, p = \text{n.s.}$; George Bush, $\chi^2 = (1, N = 36) = 0.42, p = \text{n.s.}$; fraternities and sororities, $\chi^2 = (1, N = 31) = 3.66, p = .07$). Please see Appendix D for a table listing the breakdown of essay topic per condition.

Participants' mean attitude strength rating for the essay they chose, which was 5.72 ($SD = 1.07$) on a 7 point scale, indicated that the participants' attitudes were sufficiently strong. In order to ensure that attitude ratings did not differ by condition, I subjected participants' attitude strength ratings to a 2 (target gender: man or woman) X 2 (feedback: emotional or wordy) between-subjects ANOVA, and found no significant main effects or interactions: main effect of target gender, $F(1,112) = 0.14, p = n.s.$; main effect of feedback, $F(1,112) = 1.01, p = n.s.$; and target gender by feedback interaction, $F(1,112) = 0.17, p = n.s.$ (See Table 4 for means.)

Table 4

Mean Attitude Strength as a Function of Target Gender and Feedback

	Emotional Feedback	Wordy Feedback
Women	5.62 (1.20)	5.61 (1.20)
Men	5.74 (0.89)	5.90 (0.98)

Note: Standard deviations are in parentheses.

Description of Thought Content. Participants wrote an average of 4.80 thoughts ($SD = 2.19$) out of 12 possible thoughts. Out of those thoughts, an average of 1.68 thoughts ($SD = 2.42$) were rated by the independent raters as a 5 or above on the category of thinking about something unrelated to the essay or feedback (e.g., "I am hungry"). Thus, the 29% of the thoughts were rated as irrelevant, indicating that 71% of the thoughts were relevant to the feedback. Because the majority of thoughts were relevant to the feedback, it can be concluded that, for the purposes of this study, the thought-listing exercise sufficiently measured reactions to feedback.

Independent Raters' Assessments of Participants' Thoughts. To test the hypotheses, I subjected the independent raters' assessments to a 2 (feedback: emotional or wordy) X 2 (target gender: man or woman) X 2 (target thought focus: emotions or wordiness) mixed factor ANOVA with repeated measures on the third factor. In other words, I tested the extent to which feedback conditions and target gender affected the degree to which participants were rated as thinking

about emotions versus wordiness. I found a main effect of target gender, $F(1,105) = 13.95, p < .001, \eta_p^2 = 0.12$, and a significant target thought focus by feedback interaction, $F(1,105) = 64.11, p < .001, \eta_p^2 = 0.38$ but these were qualified by a significant 3 way interaction: feedback by target gender by target thought focus, $F(1,105) = 3.75, p = .056, \eta_p^2 = 0.03$. To explain this 3-way interaction further, for each target thought focus I conducted a 2 (feedback: emotional or wordy) X 2 (target gender: man or woman) between-subjects ANOVA.

Looking first at the emotion thought focus, I found a marginally significant feedback by target gender interaction, $F(1,105) = 3.38, p = .07, \eta_p^2 = 0.31$. In support of Hypothesis 1, women who were called emotional wrote thoughts that were rated as more intensely about emotions than men who were called emotional, $F(1,105) = 11.21, p = .001, \eta_p^2 = 0.10$. Women who were called wordy did not differ from men who were called wordy in terms of how intense thoughts were about emotion, $F(1,105) = 0.63, p = \text{n.s.}, \eta_p^2 = 0.01$. These results demonstrate that, compared to men, women in the emotion condition in particular wrote thoughts that were more intensely about emotion. Hypothesis 2 Part A stated that women who are called emotional will be more preoccupied with their own emotions than women who are called wordy will be with their own emotions. As predicted, women who were called emotional wrote thoughts that were rated as more intensely about emotions than did women who were called wordy, $F(1,105) = 43.20, p < .001$. Incidentally, men who were called emotional were rated as having more intense thoughts about emotions than men who were called wordy, $F(1,105) = 13.35, p < .001$. These results provide evidence that merely receiving negative feedback does not lead to as much of an increase in preoccupation with emotions as being called emotional does. Because I made four comparisons, the Bonferroni correction suggests that alpha levels should be at 0.015 to be

considered significant. The significant results meet such standards. (See Table 5 for the means. Please also see Figure 3 for a pictorial representation of these same means.)

Looking next at the wordy thought focus, I did not find a significant feedback by target gender interaction, $F(1,105) = 0.91$, $p = \text{n.s.}$, $\eta_p^2 = 0.01$. However, because I did not make any a priori predictions about men versus women in the wordy conditions, these results are not inconsistent with my hypotheses.

Hypothesis 2 Part B stated that women who are called emotional will be more preoccupied with their own emotions than women who are called wordy will be with their own wordiness. To test this hypothesis, I performed a Tukey post-hoc test and found that there was only a statistically non-significant trend for women who were called emotional to write thoughts that were rated as more intensely about emotions than women who were called wordy to write about wordiness, $t(56)=1.19$, $p = .12$.

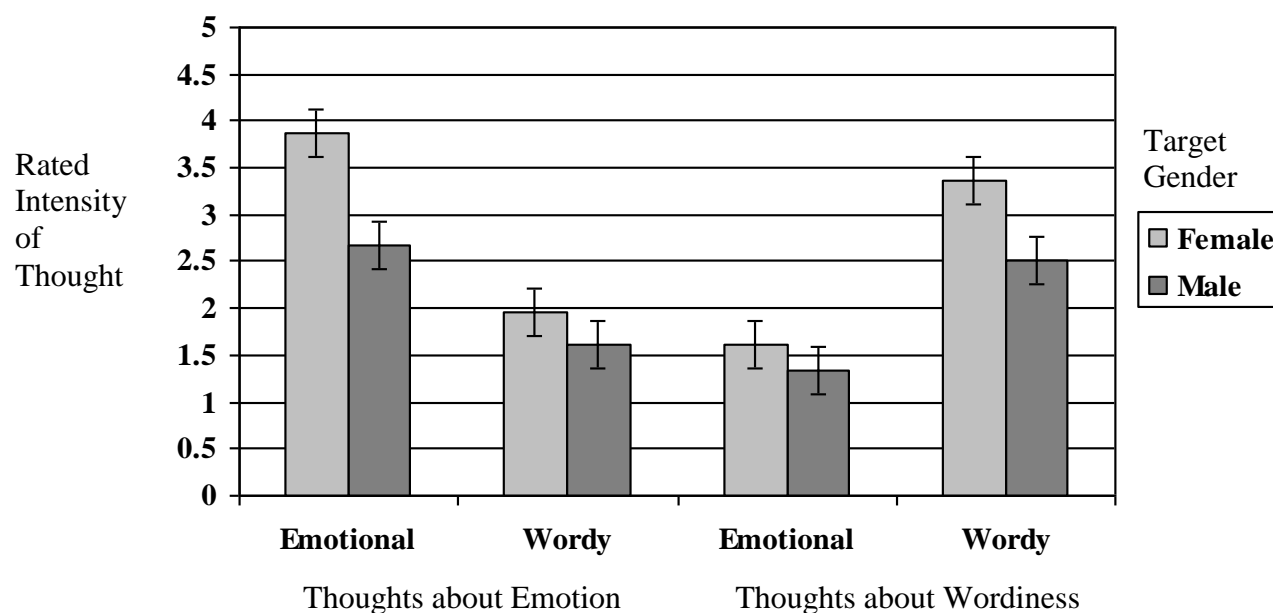
Table 5.

Mean Raters' Assessments of Thoughts about Emotion and Wordiness as a Function of Target Gender and Feedback

	Intensity of Thoughts about Emotion		Intensity of Thoughts about Wordiness	
	Emotional Feedback	Wordy Feedback	Emotional Feedback	Wordy Feedback
Female Targets	3.87 (2.10)	1.61 (0.77)	1.96 (1.50)	3.36 (1.47)
Male Targets	2.67 (1.31)	1.33 (0.44)	1.61 (1.04)	2.51 (1.37)

Note: Standard deviations are in parentheses

Figure 3. Raters' assessments of mean thoughts about emotion and wordiness as a function of target gender and feedback



Finally, I find that it is also important to point out the overall main effect of target gender, such that women were more likely to be preoccupied than men regardless of condition ($M = 2.70$, $S.D. = 1.73$ vs. $M = 2.03$, $S.D. = 1.24$, respectively). Rather than specifically being more preoccupied than men when called emotional, women were more preoccupied than men in both feedback conditions.

In summary, the results of the independent raters' assessments of participants' thoughts partially supported the Hypotheses 1 and 2A. In support of these hypotheses, women in the emotion condition were rated as particularly preoccupied with their emotions as compared to

men and as compared to women in the wordy condition. Inconsistent with Hypothesis 2B, women in the emotional condition were not significantly more likely to be rated as preoccupied with their emotions than women in the wordy condition were with their wordiness, although a trend was consistent with this pattern. In addition, women were more likely to be preoccupied than men in general.

Participants' Assessments of Their Own Thoughts. For the purpose of identifying possible corroborative support for the hypotheses, I subjected participants' assessments to a 2 (feedback: emotional or wordy) X 2 (target gender: man or woman) X 2 (target thought focus: emotions or wordiness) mixed design ANOVA with repeated measures on the third factor. Unexpectedly, the only significant result was a target thought focus by feedback interaction, $F(1,105) = 12.46, p = .001, \eta_p^2 = 0.11$, such that individuals who were called emotional rated their thoughts to be more strongly about emotions than about wordiness ($M = 4.78, SD = 1.50$ vs. $M = 4.39, SD = 1.46$), $F(1,105) = 3.81, p = .05, \eta_p^2 = 0.04$, and individuals who were called wordy rated their thoughts to be more strongly about wordiness than about emotions ($M = 5.00, SD = 1.49$ vs. $M = 4.11, SD = 1.95$), $F(1,105) = 4.93, p = .04, \eta_p^2 = 0.04$. Because I made four comparisons, the Bonferroni correction suggests that alpha levels should be at 0.015 to be considered significant. The significant results do not meet these standards, so should be interpreted with caution.

Thus, I did not find the significant three-way interaction that I found with the independent raters' assessments of participants' thoughts. Nevertheless, there are some promising trends. (See Table 6 for the means. Please also see Figure 4 for a pictorial representation of these same means.) There is a trend supporting Hypothesis 1 such that women who are called emotional report that their thoughts are more strongly about emotions than men

report them to be. A trend also supports Hypothesis 2 such that women who are called emotional report their thoughts to be more strongly about emotions than women who are called wordy.

Inconsistent with Hypothesis 2, however, women who are called emotional did not report their thoughts to be about emotions more than women who are called wordy reported their thoughts to be about wordiness.

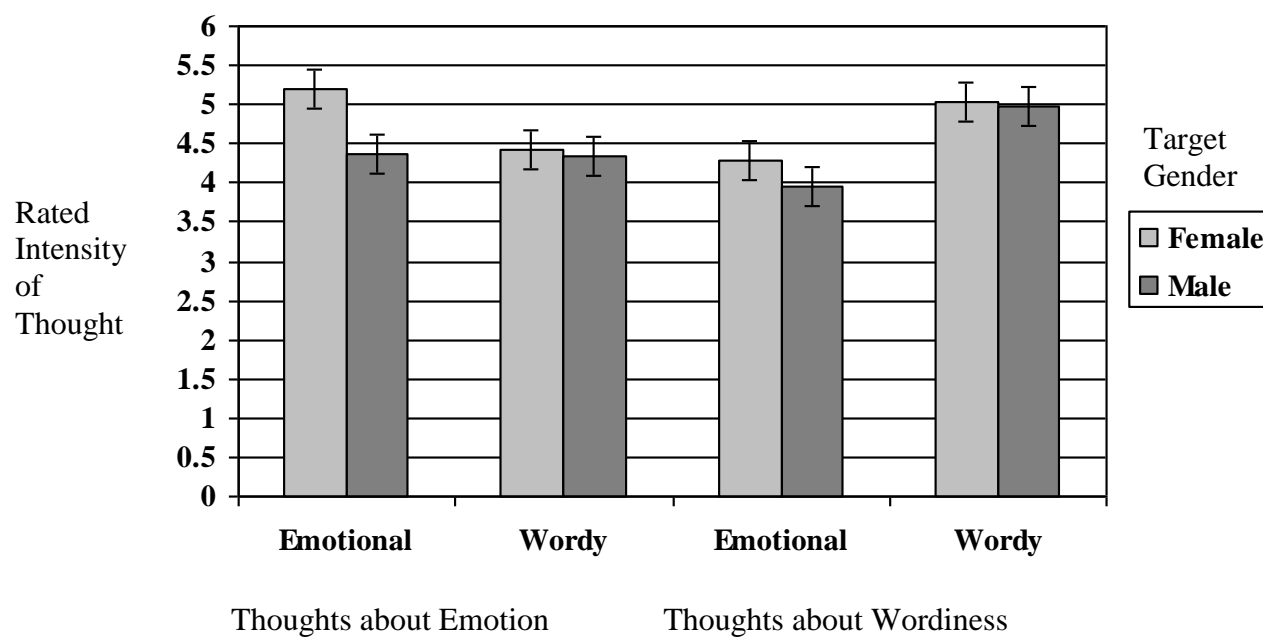
Table 6

Mean Participants' Assessments of Thoughts about Emotion and Wordiness as a Function of Target Gender and Feedback

	Intensity of Thoughts about Emotion		Intensity of Thoughts about Wordiness	
	Emotional Feedback	Wordy Feedback	Emotional Feedback	Wordy Feedback
Female Targets	5.20 (1.47)	4.28 (1.88)	4.43 (1.46)	5.02 (1.22)
Male Targets	4.36 (1.44)	3.95 (2.03)	4.34 (1.48)	4.98 (1.73)

Note: Standard deviations are in parentheses

Figure 4. Mean participants' assessments of thoughts about emotion and wordiness as a function of target gender and feedback



I had suspected that if differences between participants' assessments and independent raters' assessments emerged, participants' assessments would simply indicate more intense thoughts overall. I had not expected that differences would emerge within feedback conditions. To assess the extent to which participants' and independent raters' assessments differed, I conducted a 2 (feedback: emotional or wordy) X 2 (target gender: man or woman) X 2 (target thought focus: emotions or wordiness) X 2 (rater type: independent or participant) mixed factor ANOVA with repeated measures on the third and fourth factors. The dependent variables were participants' assessments of the degree to which their thoughts were about emotion and about wordiness. Several significant effects emerged, including a main effect of rater type, $F(1,96) = 250.78, p < .001, \eta_p^2 = 0.72$, and a target thought focus by feedback interaction, $F(1,96) = 53.85, p < .001, \eta_p^2 = 0.36$, both of which were qualified by a three-way rater by target thought focus by feedback interaction, $F(1,96) = 10.98, p = .001, \eta_p^2 = 0.10$. The simple main effects of this interaction reflect the main effect for type of rater across all conditions, and also reflect the differences within rater type that have already been discussed. To examine the means, standard deviations, and simple main effect F-tests for this interaction, please see Appendix H.

There was also a main effect of target gender, $F(1,96) = 7.23, p = .008, \eta_p^2 = 0.07$, which was qualified by a rater type by target gender interaction, $F(1,105) = 4.13, p = .05, \eta_p^2 = 0.04$. This interaction indicates that, regardless of condition and thought type, both male and female participants rated their own thoughts as more intense than independent raters did, $F(1,105) = 95.27, p < .001, \eta_p^2 = 0.62$ and $F(1,105) = 159.65, p < .001, \eta_p^2 = 0.50$, respectively. This result simply demonstrates the overall main effect of rater. Also, independent raters assessed female participants' thoughts as more intense than they rated male participants' thoughts, $F(1,105) = 15.98, p < .001, \eta_p^2 = 0.14$. This finding reflects that, averaged across conditions, the independent

raters assessed female participants as possessing more intense thoughts than male thoughts. Men and women participants' self-ratings did not significantly differ from each other $F(1,105) = 0.42$, $p = \text{n.s.}$, $\eta_p^2 = 0.004$. Because I made four comparisons, the Bonferroni correction suggests that alpha levels should be at 0.015 to be considered significant. The significant results meet these standards. Please see Table 7 for means and standard deviations.

Table 7
Mean Thought Intensity as a Function of Target Gender and Rater Type

	Participant Assessment	Independent Rater Assessment
Women	4.74 (1.55)	2.77 (1.77)
Men	4.57 (1.70)	2.03 (1.25)

Note: Standard deviations are in parentheses.

In sum, the comparison between the independent raters' and participants' assessments indicate a main effect for raters across all conditions, such that participants rated their own thoughts as more intense than independent raters did. Because the differences between independent raters' and participants' assessments were so large, interactions merely reflected variations within rater type. In the discussion section I consider why these differences between independent raters and participants emerged.

Discussion

The results of this study provide preliminary evidence about the effect that being called emotional has on male and female targets. The independent raters' assessments indicated that women who were called emotional were more likely to be preoccupied with their own emotions than men who were called emotional and women who were called wordy. However, only a non-significant trend suggested that women who were called emotional were preoccupied with their own emotions more than women who were called wordy were preoccupied with their own wordiness. Results also revealed that women were more likely to be preoccupied than men

regardless of the type of feedback. The participants' assessments, on the other hand, did not match independent raters' assessments, showing only that individuals were more preoccupied with the content of the feedback they perceived than other types of content.

One of the more pressing questions that these data elicit is what to make of the discrepancies between participants' and raters' assessments of participants' thoughts. At the outset of the experiment I had expected that participants might, across all conditions, rate their own thoughts as more intense than raters would, and this was the case. However, I had not predicted *a priori* that participants' and raters' assessments would differ within condition. Upon examining the data I believe several factors are contributing to the discrepancies. First, the independent raters were exposed to a series of rules to help them decipher the meanings of the items used as the dependent variables. The participants, on the other hand, were left to their own interpretations of the items. For example, participants who were called emotional may not have sufficiently distinguished between the feelings they were having and their thoughts *about* their feelings. As a result, their potentially strong affective responses across conditions may have masked variations within conditions regarding their reflective thoughts about their emotions. Thus, it is possible that participants' results were not as clear as they would have been if a structured set of rules were present to help guide participants in their assessments. In the future, if I were to use participants' assessments again, I would need to provide more explicit guidelines for self-rating.

If participants' assessments were in fact affected by the lack of structured rules and by their affect, then it is appropriate for me to draw my conclusions about this study more from the independent raters' assessments. I discuss potential reasons behind the unexpected results and directions for future research. Specifically, I focus on the fact that independent raters found that

women who are called emotional were not found to be significantly more likely to be preoccupied with emotions than women who are called wordy are with wordiness. I then focus on the findings that women were both more likely to be preoccupied than men not only in the emotional condition, but also the wordy conditions.

Several possibilities exist as to why women who were called emotional were not significantly more likely to be rated as preoccupied with their own emotions than women who were called wordy were with their own wordiness. One possibility is that “wordiness” was not an effective control group. I chose wordiness as a control group because it is a type of negative feedback that is not associated with stereotypes of women. One potential criticism of this choice is that wordiness could be tied to the stereotype that women are more verbally-oriented than men (Schneider, 2001). Typically, however, this stereotype is concerns girls’ earlier development of verbal skills and women’s greater verbal than quantitative intelligence. Without further study, it is unclear whether this stereotype extends to the amount of verbiage in a written essay. Another potential limitation of wordiness as the control group is that perhaps it is milder negative feedback than emotion criticism. The participants received equally negative feedback in terms of the grade, which served to minimize the differences between the intensities of the two statements. In an effort to equalize the two types of feedback, I gave participants a C- and then used the statements to justify why the C- was given. Future research would benefit from other types of control conditions that are pre-tested for being low on stereotypicality for women and for being of equal intensity as being called emotional.

In addition, women may have been equally as likely to have been rated as preoccupied with emotions and wordiness because of the way the thoughts were measured. It is possible that alternative ways to measure thoughts might be more sensitive to preoccupation with emotion.

For example, it is possible that, because of not wanting to appear emotional, women may be uncomfortable sharing thoughts about their feelings through the thought-listing exercise. As mentioned previously, one consequence of being called emotional may be emotion suppression, or reducing emotion-related thoughts and expressions in order to match situation norms (Gross & Levenson, 1994). As a result, more indirect measures of preoccupation, such as a lexical decision task, an implicit measure of attention, may be more effective. At the same time, women were assessed as expressing some preoccupation with emotions, suggesting some comfort with sharing thoughts about feelings. Perhaps the current coding schemes are not sensitive to more subtle expressions of preoccupation with emotion in which emotion is not explicitly mentioned. If so, perhaps it would be useful to devise a more structured coding scheme for participants to code their own thoughts, for they may have greater access to this preoccupation than independent raters who solely have written expressions of thoughts.

In future research I will test these variations of measuring preoccupation. The thought-listing exercise has both promise and limitations. It has promise because it offers a quantifiable measure for assessing thought content when research questions call for such a measure. It also has the potential to allow both participants and independent raters to assess thought content. Multiple perspectives, especially when assessing emotions, may increase the robustness of results (Larsen & Fredrickson, 1999). Thought listing also has its limitations. One limitation is that researchers only have access to those thoughts that participants are both willing and able to communicate verbally. In addition, in its present form the measure only asks for distinct thoughts, which makes it difficult to measure frequency of the same thought. Future iterations of the thought listing task will include directions that enable measurements of frequency. In future work I will examine these possibilities and limitations.

I turn now to the result that women were more likely to think about the feedback they received than were men. This result is consistent with the pervasive finding that women are more likely to engage in rumination than men are (Nolen-Hoeksema, Larson, & Grayson, 1999). Rumination is generally defined as conscious, recurrent thoughts about past negative information (Nolen-Hoeksema, 1991). Although the findings in this research do not directly point to rumination, preoccupation with negative feedback may be evidence of participants' focus on the causes of their negative reactions, a component of ruminatory thought (Nolen-Hoeksema & Jackson, 2001). Rumination is characterized as focusing on one's symptoms of distress and the causes and consequences of these symptoms, and differs from more adaptive coping strategies in that individuals do not actively repair negative emotions. It may be useful for future research to examine whether rumination is a feature of preoccupation, and the degree to which this rumination can account for women's responses.

Although women on average tend to ruminate more than men, being called emotional may have a particular significance for women. Even though I did not find statistical support for the hypothesis that women are likely to be more preoccupied with emotions when called emotional more so than women called wordy are with their wordiness, I found a promising trend. I also found that women who were called emotional were rated as being more preoccupied with emotions compared to women or men in all other conditions. The rationale for Study 2 was based in part on the assumption that women who are called emotional will be preoccupied as a result of a desire to maintain legitimacy. Future work will directly test whether this desire indeed predicts preoccupation.

CHAPTER 4: General Discussion

In this thesis I developed the argument that, because of emotion's potential to subvert reason, biased perceptions of emotionality can function to delegitimize individuals, and in particular, women. In order to test this hypothesis, I looked at the extent to which calling a woman emotional affected the parties involved in the delegitimization process. In the first study, I tested the extent to which calling a woman versus a man emotional functioned to delegitimize them in the eyes of a third-party observer. In the second study, I examined the consequences of being called emotional for targets, especially women. I specifically examined the degree to which calling women emotional made them preoccupied with their emotions.

The results partially supported my predictions. Specifically, in the first study I found that when the target was called emotional, male targets' arguments were seen as more legitimate than female targets' arguments. Mediation analyses revealed that observers' evaluations were rated as less legitimate when the target was male than female, suggesting that the third party observer did not believe the observer when he or she called the male target emotional. While I found that men had the advantage over women, I did not find evidence that women were specifically delegitimized. In my second study I found that women who were called emotional were judged by independent raters as being the most preoccupied with their emotions relative to men who were called emotional and both men and women who received non-emotion-related negative feedback. I did not find statistical support for the hypothesis that women who were called emotional would be more preoccupied with their emotions than women who were called wordy were with their wordiness. While I found these result for independent raters, participants' self-ratings did not follow this pattern, perhaps due to the fact that, unlike independent raters, participants were not instructed to follow specific coding rules.

In this discussion I will focus on future directions by elaborating on the program of research for which my results serve as a foundation. I begin by discussing when emotion may legitimize and when it may delegitimize. Next, I elaborate on the theoretical foundation for my program of research. Specifically, I will discuss the extent to which emotion's relation to legitimacy is associated with systems of structural inequality between men and women and also between other social groups.

When Does Emotion Legitimize and When Does It Delegitimize?

Although the present work provided evidence *that* emotion is related to perceptions of legitimacy, it remains to be tested *why* emotion influences whether a person is delegitimized or legitimized and in which contexts. As stated previously, I argue that legitimizing entails making a claim or a person seem valid in the eyes of an actual or implied reference group, while delegitimizing makes such a claim or person seem invalid (Berger, Ridgeway, Fisek, & Norman, 1998; Zelditch, 2000). As this definition of legitimacy suggests, emotions would legitimize or delegitimize if they could signify that a person's claim or argument is valid or invalid, respectively. Thus, one question to ask is what aspects of emotion are related to perceived validity. As briefly discussed at the outset of this paper, I predict that emotion's association with authenticity and humanity affect the validity of an argument.

A belief exists, at least in the North American societies that have been studied, that others' genuinely expressed emotions give us a window into their true "inner nature," or in other words, their authentic self (Johnson, Robinson, & Mitchell, 2004; Pizarro, 2000; Shields 2005). If emotions do indeed portray one's authentic self, how might this authenticity affect assessments of legitimacy? In psychological literature, authenticity is typically defined as the extent to which a person is behaving in a way that is consistent with one's true or core self (Ashforth & Tomiuk,

2000; Kernis, 2003; Wild, 1965). While authors have identified several components of what constitutes authenticity (e.g., Kernis, 2003), what is most relevant to perceptions of legitimacy is the belief that those individuals who are authentic are principled, truthful individuals who act according to their values (Crocker & Nuer, 2003; Kernis, 2003).

Most research has focused on beliefs about and perceptions of one's own authenticity. The little actual research on perceptions of others' authenticity suggests that authenticity leads individuals to be perceived as more influential and believable. In a series of studies on close relationships, Neff and Harter (2000) found that those individuals who resolved conflicts in an inauthentic way (i.e., reported that they resolved conflicts in a way that was not a reflection of their "true self") reported lower levels of validation from their partners than those who resolved conflicts in an authentic way. More specifically, individuals who reported a match between how they wanted to resolve conflicts and how they actually did resolve the conflicts also reported that partners were more likely to listen to them and take their views seriously. In other words, when participants perceived that they were being authentic their partners were more likely to be influenced by what they said. In another series of studies, this time in a political realm, Donsbach and Jandura (2003) found that politicians who were perceived to be authentic had a more positive influence on viewers' perceptions of them. In sum, these studies show that when an individual is seen as authentic, they are evaluated positively. If authenticity is indeed associated with the perception that a person is truthful, a person who is perceived to be authentic may be seen as conveying legitimate arguments. In other words, we believe that what they say or do is more valid and worthy of our attention. And thus, if we see that a person is expressing his or her true feelings, we will be more convinced by what they say or do.

Authenticity by itself, however, may not determine whether emotions will legitimize or delegitimize. In other words, a person may need to act authentically, but it is only a certain type of authenticity that will legitimize. A person could be authentically reasonable and competent or authentically irrational and mad. The former may be more likely to lead to legitimacy and the latter to delegitimacy. In her discussion of late 19th Century scientific beliefs about emotion, Shields (2005; 2007) discusses two ways in which emotion was represented, as “passions” or as “mere emotionality.” While emotionality was associated with uncontrolled, weak, and ineffectual urges, passions concerned all strong feeling that drove “creative thinking, social action, and physical prowess” (p. 5). This latter type of representation identified competence as infused with personal conviction (i.e., passion).

Shields’ work on the 19th Century resonates with Haslam’s (2006) theory of dehumanization. Haslam argues that dehumanization, or the stripping of human-like traits from individuals, involves making a distinction between what individuals perceive as uniquely human characteristics and those that do not distinguish us from other species (such as irrationality, coarseness, and childishness). Distinguishing between human and inhuman emotions enables dehumanization. Haslam cites work on inhumanization in which researchers found that emotions judged to be uniquely human were believed to be more “morally informative, cognitively saturated, internally caused rather than responsive to the environment, private... and emerging late in development” (Demoulin et al., 2004 as cited in Haslam, 2006, p. 256). Thus, those who are perceived to be “human” are seen as emoting in a way that conveys maturity, moral refinement, and civility, or in other words, rationality. According to his theory, those who are dehumanized, on the other hand, are seen as childish, lacking refinement and civility, or in other words, irrational. Thus, humanity is dictated not just by any emotion, but rather emotion

that suggests rationality. I predict that emotion will function to legitimize when it is perceived to be authentic and when it is associated with these refined qualities that represent “humanity” and emotion will delegitimize when it lacks these qualities. In future work I plan to test the extent to which the perceived humanity and authenticity that emotions convey mediate the relationship between emotions and legitimacy.

At the outset of Study 1 I predicted that calling someone emotional would have a delegitimizing effect, regardless of gender. Because results revealed that men were seen as more legitimate when emotional than when disagreed with, however, I may have underestimated the degree to which stereotypes about men’s competence serve to legitimize. I do not, however, argue that delegitimizing via emotions can only occur for women. It may be the case that in somewhat more restricted conditions, emotion labeling can also delegitimize men (in terms of the general, undifferentiated category of “man,” irrespective of other social identities).

Beliefs about men’s competence in expressing emotions (e.g., Timmers, Fischer, & Manstead, 1998) may lead men to be legitimized more often than delegitized. However, in extreme situations, where the norms are clear and unambiguous (such as at a political rally), and thus where stereotypes and beliefs are less likely to inform beliefs, men may be delegitized when violating those norms. Shields (2002) argues that visible, open displays of emotion can have deleterious consequences in situations that call for more restrained emotional displays. An anecdotal example is Howard Dean’s infamous “scream” during the 2004 Iowa Caucus, where he was perceived to inappropriately regulate his excitement about his bid for the presidential nomination, and afterwards lost his status as favorite for the Democratic nomination. While little research has explicitly looked at the effects of emotion displays on delegitimization of men, Shields and I (Warner & Shields, *In press*) found that men who were described in vignettes to be

openly crying in response to learning about a parents' impending divorce or finding out about a cheating girlfriend were rated less positively than men who displayed more subtle tears. In sum, men tend to be perceived negatively when they are perceived to violate expressive norms. Thus, in situations where calling a man emotional implies that he has violated an expressive norm, he may be delegitimized.

Consequences of Delegitimization for Women

I now turn to the consequences that being called emotional has on targets. What happens to a woman target when she is called emotional? The purpose of Study 2 was to test whether women would respond with preoccupation. As mentioned previously, future work must first establish the conditions under which being called emotional leads to preoccupation. In this section I will discuss future directions of this research by proposing possible theoretical explanations for this preoccupation and potential negative consequences it has for women.

One potential explanation for why targets engage in preoccupation is that they are engaging in a version of behavioral confirmation. Behavioral confirmation occurs when a perceiver holds an expectation about a target and, through treating the target in line with his or her expectations, elicits that very expectation from the target (Snyder, 1992; Snyder, Tanke, & Berscheid, 1977). The typical experimental situation in a behavioral confirmation study involves a perceiver and a "naïve" target in which the target is unaware of the perceiver's expectation and unwittingly behaves in accordance to that expectation. Calling someone emotional varies from typical scenarios in behavioral confirmation literature because when called emotional, the target is very aware of the expectation.

Behavioral confirmation research argues that when a target *is* made aware of a negative expectation, he or she will actively try to counter it (Stukas & Snyder, 2002). Indeed, Stukas and

Snyder (2002) found that when targets are informed of the perceiver's negative expectations (that they were introverts), targets engaged in actions to attempt to counter the expectations (by acting in more extroverted ways). When called emotional, then, a target may react by trying not to be emotional. This reaction would first involve heightened attention to their emotion response (preoccupation) in order to monitor it for emotionality. Then, targets may engage in what is known as response-focused emotion regulation, or in other words, the things we do once an emotion is already underway, after the response tendencies have been generated (Gross & Levenson, 1993). The most common form of response-focused emotion regulation is suppression of emotion, in which the individual restrains from engaging in an emotion response and overt behaviors (Butler et al., 2003). Emotional suppression is typically associated with psychological and physical consequences such as less successful social interactions (Butler et al., 2003) and physiological strain (Butler & Gross, 2004, Christensen et al., 1997). Thus, in trying to resist confirming the observer's accusation, women may engage in consequential behaviors.

However, even if women attempt to suppress their affective reactions, they may still be "trapped" into confirming the expectation. When being called emotional, the very concern over confirming an emotional behavior may ironically lead to confirming it. In other words, being called emotional during an interaction might lead to feelings of anxiety if the individual is afraid of confirming the expectation for the remainder of that interaction, and also may lead to feelings of frustration over the negative evaluation. These feelings in and of themselves may subsequently actually confirm the expectation that the woman is being emotional, perhaps to the woman and/or to the observer. When the particular situation involves women, worries or frustration about having confirmed a negative stereotype may be added to the concern that was already present.

In sum, behavioral confirmation processes may be a useful theoretical framework for defining the processes through which being called emotional has negative experiential and behavioral effects on targets and will be a focus for future research. I turn now to outlining the theoretical framework for my research program for which the present experiments serve as a foundation. Specifically, I connect my research to theories about the maintenance of structural inequality between men and women, and also to structures of inequality between intersections of multiple social identities.

Emotion, Legitimacy, and Structures of Gender Inequality

I base my assertions about the consequences for women of being labeled emotional on the assumption that gender entails structural inequality between men and women (e.g., Hare-Mustin & Marecek, 1990; Ridgeway & Correll 2004). Furthermore, I base my assertions on the idea that gender is not something one *has*, but is something that one *does*. Using West and Zimmerman's (1987) theory of "doing gender," I base my work on the understanding that social markers of "man" and "woman" are always contested, disputed, and negotiated. I argue that emotion's connection to legitimacy results in the performance of emotions as means through which inequality between men and women is maintained. One of the major consequences of the belief that emotion can be an uncontrolled, primitive reaction is how it serves to maintain the connection between thought, reason, and White Western masculinity and the connection between emotion with femininity and racial others. In this way, discussions about maintenance of emotion control are really discussions about justifying maintenance of social order in general.

For example, calling someone emotional serves not only to reduce the claim to mere emotionality, but it also takes away the contribution that emotions give to thought processes (Ahmed, 2004). Shields (2002) argues that emotions can be described in terms of what they are

about. They provide us information about our orientation towards objects in the environment and what we need to do in reaction to those objects. For example, anger involves a clear appraisal that one has been denied what one is rightly due (e.g., Lazarus, 2001). When people reduce others' anger to "oversensitivity" or inappropriate emotionality, they are effectively disputing the basis for the perception of injustice. Hence, the label chosen to name an occurrence of emotion is a value-laden act; for example, by labeling the person as "bitchy" compared to "indignant," one implies that the person has no significant basis for feeling angry or upset. When dominant group members discuss groups who are minority or considered "special interest" who make claims on the majority, the question of "Why are they always so angry?" often arises (Shields, 2005).

Although Catherine Lutz (1990) viewed emotion is irrational (without explicitly noting that emotion can also be used to bolster reason), she provides a persuasive argument about the effects that discourses about emotion have on power relations. Lutz identifies this discourse on emotion control as really about a larger discourse on controlling people (in her discussion, women in particular) who are perceived to have unregulated emotion. Lutz' work indicates that by calling someone emotionally uncontrolled, one is legitimizing social control over that person as well. Lutz found that when individuals in her interviews discussed other people's need to control themselves (usually women), references were made to these people as "dangerously" emotional. In addition, she cites work on discourses of premenstrual syndrome in women, in which people discuss how it problematic that women engage in such uncontrollable anger and irritability while they have PMS. These pieces of evidence demonstrate the belief that an individual's inappropriately regulated emotions provide justification for social control over that individual.

Lutz argues that the discourse on the control of emotions is very similar to Western discourses on sexuality, especially as described by Foucault (1980, as stated in Lutz, 1990). She argues that both emotionality and sexuality are dominated by descriptions of them as “universal, natural impulses,” that there are both healthy and unhealthy forms, and both have been discussed in terms of medical problems, or, more in the case of emotions, psychiatric problems. She says that Foucault would argue that this description of being “natural, dangerous, irrational, and physical” (p. 154) takes the attention away from the idea that they are repeatedly reproduced, that is, “doing emotion” instead of “having emotions.” In other words, one of the reasons why labeling emotions as uncontrolled or irrational is such an effective way to maintain unequal power relations is that such labels are cloaked beneath a discourse about the “natural.” Being able to attribute emotions to nature, or as Haslam (2006) argues, non-uniquely human attributes, ignores the socially constructed nature of emotions. If one can deny that emotion is socially constructed, one can also deny that those in power have the capability to influence who gets to have appropriate emotions. Thus, saying that women are emotional may seem more like common sense than as an attempt to maintain power differentials. As Lutz says, “Given its definition as nature, at least in the West, emotion discourses may be one of the most likely and powerful devices by which domination proceeds” (p. 159). Lutz, in sum, demonstrates that emotions are powerful forms of representation for the negotiation of gender hierarchies.

Does Emotion’s Relationship with Legitimacy Reach beyond Gender?

Although I have focused on inequality between men and women in this paper, I argue that my work extends beyond this particular type of inequality. Previous work in psychology has demonstrated a strong association between inappropriate emotions and women, but rarely has the gendered nature of emotion beliefs been considered in conjunction with targets’ race or other

potentially important social identities, such as age or sexuality. Thobani, who I mentioned in the introduction, was not solely a “woman” – she was a professor of Middle-Eastern descent – identities that may also have played into perceptions of her emotions as irrational. Thus, in order to better understand the relationship between emotion and legitimacy, I need to expand my focus beyond the category of “woman,” specifically by looking at how the delegitimizing and legitimizing process is negotiated through multiple identities.

Stereotypes tend to be historically contingent and applied to specific intersections of social identities (Fiske, Glick, Cuddy, & Xu, 2002), such as the “Angry Black Woman,” the “Dragon Lady,” or “hot-blooded” Latina. Thus, the aspects of emotion that will function to legitimize or delegitimize may depend on the stereotypes associated with social groups. For example, North American stereotypes of Asian women as reserved and subservient (Hess, Beupre, & Cheun, 2002) may lead Asian women to be seen as inauthentic (due to perceived lack of open, deep feeling), whereas stereotypes of Black women as inappropriately angry (Landrine, 1985) may be linked to perceptions of their emotions as lacking self-control and clear thinking. Thus, Asian women may be the most severely delegitimized in situations that warrant authenticity, whereas Black women may be the most severely delegitimized in situations that warrant rationality.

Another way to expand the discussion of emotionality and legitimacy is to examine how the identities of these different social groups intersect to create new meaning. Intersectionality is the mutually constitutive relation among social identities (Baca Zinn & Dill, 1996; Crenshaw, 1997, Nakano-Glenn, 1999). This definition is based on the idea that social group categories interact to form emergent properties not found when looking at each individually. Looking at how individuals negotiate their multiple identities may provide a more nuanced perspective on

how targets react to being labeled as inappropriately regulating emotion. For example, Weber (2004) argues that social identities often give us power and options in some places while restricting our options in others, and that it is only through looking at how these identities intersect that we can understand human experience. She argues that we tend to occupy both dominant and subordinate positions among race, class, sexual orientation, and gender hierarchies, and that we negotiate experience through them. In other words, we may negotiate our identities to buffer against parts of ourselves that we see as more vulnerable.

Specifically, individuals may negotiate their identities to represent their emotions in ways that legitimize rather than delegitimize. For example, Levitt and Hiestand (2004) conducted qualitative interviews on women who self-identified as butch lesbians and found that they performed different aspects of their identities depending on the particular emotion norms that were present in the situation: “some women allowed themselves to be more feminine when they wanted to be spiritual, to put others at ease, or to be emotionally vulnerable with a romantic partner, and they adopted a more butch presence when they felt more sexual, insecure, or playful” (Levitt & Hiestand, 2004, p. 610). In other words, they sometimes described becoming more butch, in terms of being tough, steely, and strong, as a way to guard tender emotions, and sometimes specifically as a guard against scrutiny or discrimination. Thus, their gender performance with regards to emotion seems like it varied depending on if they thought it was safe or not safe to be emotional. Bringing that to a context where emotion could either legitimize or delegitimize, one could imagine that the women described in this interview may have exaggerated the more butch aspects of themselves if they expected to be discriminated against for being “emotional.” At the same time, however, it is important to mention that while they were able to be flexible in their identities in some cases, they did struggle with others’

expectations to be tough, not to cry, or to take care of others, and thus found it hard at times to be seen as sensitive and vulnerable when the situation called for it. In other words, identity was something that these women were able to negotiate, but they faced limitations on the degree to which they could successfully do so in others' eyes.

Thus, individuals may reappropriate their representations into ways that protect them from discrimination. My future work will focus on how individuals creatively use representations to empower themselves, instead of only how such representations may hurt them. In other words, I plan to focus on how women resist the negative consequences of being called emotional.

Conclusion

The major overarching theme of my dissertation research is the paradoxical nature of emotion's association with reason. Emotion is at once seen as capable of serving reason and subverting it through the association that emotion has with authenticity and humanity. I have focused on how emotion's relation to reason functions to legitimize and delegitimize and also on the consequences that delegitimization has for individuals faced with being called emotional. I have developed this line of thinking by laying out a foundation for a program of research on the relations among emotion, legitimacy, and intersections of social identities. Experimental social psychology contributes to feminist discourses on emotion and gender by examining the causes and effects of individual social interactions – what thoughts, perceptions, and behaviors will lead to using emotions to delegitimize others in a particular instant? Social psychologists study the everyday relations and patterns of social behavior. When these everyday relations are problematic, as when they maintain inequality, research helps to figure out how the patterns work so as to interrupt them. Ridgeway and Correll (2004) point out that the focus on

interrupting these everyday patterns is one important contribution towards challenging systems of power: “The gender system [as well as other systems of oppression] will only be undermined through the long-term, persistent accumulation of everyday challenges to the system resulting from both socioeconomic change and individual resistance” (2004, pg. 528). My goal is to contribute to these accumulations of everyday challenges by demonstrating the powerful role that the perception and understanding of emotion plays in reproducing or challenging social hierarchies.

ENDNOTES

1 I performed analyses looking at frequency of thoughts as opposed to intensity and only found that participants were rated as having a higher frequency of thoughts that were relevant to their feedback than were irrelevant to their feedback.

2 Several revisions to this coding scheme were suggested: 1) coding thoughts as “well-written” is not specific enough to approximate “wordiness.” “Well-written” can include thoughts beyond those associated with “wordiness” in particular; 2) instead of coding each individual thought, I should instruct raters to give an overall rating. Independent raters conducted assessments with these two revisions. However, results were largely the same as the original assessment without the revisions. Thus, I retained the original assessments for the purpose of this dissertation.

REFERENCES

- Ahmed, S. (2004). *The cultural politics of emotion*. New York : Routledge.
- Ashforth, B.E. & Tomiuk, M.A. 2000, 'Emotional labour and authenticity', In Finman, S. (Ed.) *Emotion in Organizations*. (pp. 184-203). London, UK; Sage.
- Averill J.R. (1980). A constructivist view of emotion. In R. Plutchick & H. Kellerman (Eds.), *Emotion: Theory, research and experience, Vol. 1* (pp. 305-339). New York: Academic Press.
- Averill, J.R. (1990). Disorders of emotion. *Journal of Social and Clinical Psychology, 6*, 247-266.
- Baca Zinn, M. & Dill, B.T. (1996). Theorizing difference from multiracial feminism. *Feminist Studies, 22*, 321–331.
- Ben-Ze'ev, A. (2000). *The subtlety of emotions*. Cambridge, MA: MIT Press.
- Berger, J., Ridgeway, C.L., Fisek, M.H., & Norman, R.Z. (1998). The legitimation and delegitimation of power and prestige orders. *Sociological Review, 63*, 379-405.
- Blair, I.V. & Banaji, M.R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology, 70*, 1142-1163.
- Bosson, J.K., Haymovitz, E.L., & Piel, E.C. (2003). When saying and doing diverge: The effects of stereotype threat on self reported versus non-verbal anxiety. *Journal of Experimental Social Psychology, 40*, 247-255.
- Brewer, M. & Layton Lui. (1989). The primacy of age and sex in the structure of person categories. *Social Cognition, 7*, 262-274
- Butler, E.A., Egloff, B., Wilhelm, F.H., Smith, N.C., Erickson, E.A., & Gross, J.J. (2003). The social consequences of expressive suppression. *Emotion, 3*, 48-67.

- Butler, E. A., & Gross, J. J. (2004). Hiding feelings in social contexts: Out of sight is not out of mind. In P. Philippot & R. Feldman (Eds.), *The regulation of emotion*. Erlbaum.
- Cardinu, M., Maass, A., Rosabianca, A., and Kiesner, J. (2005). Why do women underperform under stereotype threat? *Evidence for the role of negative thinking. Psychological Science, 16*, 572-578.
- Crenshaw, L. (1997). Intersectionality and identity politics: Learning from violence against women of color. In M.L. Shanley & U. Narayan (Eds.), *Reconstructing political theory: Feminist perspectives* (pp. 178-193). University Park, PA: Pennsylvania State University Press.
- Crocker, J. & Nuer, N. (2003). The insatiable quest for self-worth: Comment. *Psychological Inquiry, 14*, 31-34.
- Donsbach, W. & Jandura, O. (2003). Chances and effects of authenticity: Candidates of the German federal election in TV news. *Press/Politics, 8*, 49-65.
- Ekman, P. (1993). Facial expression and emotion. *American Psychologist, 48*, 384-392.
- Elsbach, K.D. (2000). The architecture of legitimacy: Constructing accounts of organizational controversies. In J.T. Jost & B. Major (Eds.), *The psychology of legitimacy: Emerging perspectives on ideology, justice, and intergroup relations*. (pp. 33-53). New York, NY: Cambridge University Press.
- Fischer, A. H., & Manstead, A. S. R. (2000). The relation between gender and emotion in different cultures. In A. H. Fischer (Ed), *Gender and emotion: Social psychological perspectives* (pp. 71-94). London: Cambridge University Press.

- Fischer, A. H., Manstead, A. S. R., Evers, C., Timmers, M., & Valk, G. (2004). Motives and norms underlying emotion regulation. In P. Philippot & R.S. Feldman (Eds.), *The regulation of emotion* (pp. 187-210). Mahwah, NJ: Lawrence Erlbaum Associates.
- Fiske, S.T., Cuddy, A.J.C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82, 878-902.
- Grandey, A.A. (2000). Emotional regulation in the workplace: A new way to conceptualize emotional labor. *Journal of Occupational Health Psychology*, 5, 95-110.
- Gross, J. J., John, O. P., & Richards, J. M. (2000). The dissociation of emotion expression from emotion experience: A personality perspective. *Personality and Social Psychology Bulletin*, 26, 712-726.
- Gross, J.J. & Levenson, R.W. (1993). Emotional suppression: Physiology, self-report, and expressive behavior. *Journal of Personality and Social Psychology*, 64, 970-986.
- Hare-Mustin, R.T. & Marecek, J. (199). On making a difference. In R.T. Hare-Mustin and J. Marecek (Eds.), *Making a difference: Psychology and the construction of gender*. (pp. 1-21). New Haven, CT: Yale University Press.
- Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, 10, 252-264.
- Hess, U., Beupre, M.G., & Cheun, N. (200). Who to whom and why: Cultural differences and similarities in the function of smiles. In M.H. Abel (Ed.), *An empirical reflection on the smile. Mellen studies in psychology, Vol. 4*. (pp. 187-216). Lewiston, NY: Edwin Mellen Press.

- Hutson-Comeaux, S. L., & Kelly, J. R. (2002). Gender stereotypes of emotional expression: How we judge an emotion as valid. *Sex Roles, 47*, 1-10.
- Johnson, J.T., Robinson, M.D., & Mitchell, E.B. (2004). Inferences about the authentic self: When do actions say more than mental states? *Journal of Personality and Social Psychology, 87*, 615-630.
- Kelman, H.C. Reflections on social and psychological processes of legitimization and delegitimization. (2000). Theories of legitimacy. In J.T. Jost & B. Major (Eds.), *The psychology of legitimacy: Emerging perspectives on ideology, justice, and intergroup relations*. (pp. 54-76). New York, NY: Cambridge University Press.
- Kernis, M.H. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry, 14*, 1-26.
- Labott, S. M., Martin, R. B., Eason, P. S., and Berkey, E. Y. (1991). Social reactions to the expression of emotion. *Cognition and Emotion, 5*, 397-419.
- Landrine, H. (1985). Race x class stereotypes of women. *Sex Roles, 13*, 65-75.
- Larsen, R.J. & Fredrickson, B.L. (1999). Measurement issues in emotion research. In Kahneman, D., Diener, E. and Schwarz, N. (eds) *Well-being: Foundations of hedonic psychology*. Russell Sage, New York.
- Lazarus, R.S. (1991). *Emotion and adaptation*. New York, NY: Oxford University Press.
- Levitt, H.M. and Hiestand, K.R. (2004). A quest for authenticity: Contemporary butch gender. *Sex Roles, 50*, 605-621.
- Leyens, J. P., Paladino, P. M., Rodriguez-Torres, R., Vaes, J., Demoulin, S., Rodriguez-Perez, A., et al. (2000). The emotional side of prejudice: The attribution of secondary emotions to ingroups and outgroups. *Personality and Social Psychology Review, 4*, 186-197.

- Lutz, C. A. & Abu-Lughod L. (Eds.) (1990). *Language and the Politics of Emotion*. New York, NY: Cambridge University Press.
- Marx, D.M. & Stapel, D.A. (2006). It's all in the timing: Measuring emotional reactions to stereotype threat before and after taking a test. *European Journal of Social Psychology*, *36*, 687-698.
- Memon, A., Hope, L., & Bull, R. (2003). Exposure duration: Effects on eyewitness accuracy and confidence. *British Journal of Psychology*, *94*, 339-354.
- Morgan, C. & Averill, J.R. (1992). True feelings, the self, and authenticity: A psychosocial perspective. In V. Gecas (Ed.) *Social perspectives on emotion, Vol. 1*. (pp. 95-123). US: Elsevier Science.
- Nakano Glenn, E. 1999. The social construction and institutionalization of gender and race: An integrative framework. In M. Ferree, J. Lorber, & B. Hess (Eds.) *Revisioning gender*. Thousand Oaks, CA: Sage.
- Neff, K.D. & Harter, S. (2000). The authenticity of conflict resolutions among adult couples: Does women's other-oriented behavior reflect their true selves? *Sex Roles*, *47*, 403-417.
- Niedenthal, P.M., Halberstadt, J.B., & Setterlund, M.B. (1997). Being happy and seeing "happy": Emotional states mediates visual word recognition. *Cognition & Emotion*, *11*, 403-432.
- Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration and depressive episodes. *Journal of Abnormal Psychology*, *100*, 569-582.
- Nolen-Hoeksema, S. & Jackson, B. (2001). Mediators of the gender difference in rumination. *Psychology of Women Quarterly*, *25*, 37-47.

- Nolen-Hoeksema, S., Larson, J. & Grayson, C. (1999). Explaining the gender difference in depressive symptoms. *Journal of Personality & Social Psychology*, 77, 1061-1072.
- Oatley, K. & Jenkins, J. (1996) *Understanding emotions*. Cambridge, MA: Blackwell.
- Ong, A. (1987). *Spirits of resistance and capitalist discipline: Factory women in Malaysia*. New York, NY: State University of New York Press.
- Paladino, M.P., Leyens, J.P., Rodriguez, R., Rodriquez, A., Gaunt, R., & Demoulin, S. (2002). Differential association of uniquely and non-uniquely human emotions with the ingroup and the outgroup. *Group Processes & Intergroup Relations*, 5, 105-117.
- Pizarro, D. (2000). Nothing more than feelings? The role of emotions in moral judgment. *Journal for the Theory of Social Behavior*, 30, 355-375.
- Pomerantz, E.M., Chaiken, S., & Tordesillas, R.S. (1995). Attitude strength and resistance processes. *Journal of Personality and Social Psychology*, 69, 408-419.
- Ridgeway, C.L. (2000). The emergence of status beliefs: From structural inequality to legitimizing ideology. Theories of legitimacy. In J.T. Jost & B. Major (Eds.), *The psychology of legitimacy: Emerging perspectives on ideology, justice, and intergroup relations*. (pp. 257-277). New York, NY: Cambridge University Press.
- Ridgeway, C.L. & Correll, S.J. (2004). Unpacking the gender system: A theoretical perspective on gender beliefs and social relations. *Gender and Society*, 18, 510-531.
- Rosencrantz, P., Vogel, S., Bee, H., Broverman, I., & Broverman, D. H. (1968). Sex-role stereotypes and self-concepts in college students. *Journal of Consulting and Clinical Psychology*, 32, 287-295.
- Salovey, P., Mayer, J. D., Goldman, S. L., Turvey, C., & Palfai, T. P. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the trait meta-mood

- scale. In J. Pennebaker (Ed.), *Emotion, disclosure, and health*. Washington, DC: American Psychological Association.
- Schnider, D. J. (2005). *The psychology of stereotyping*. New York: Guilford Press.
- Segerstrom, S.C., Stanton, A.L., Alden, L.E., & Shortridge, B.E. (2003). A multidimensional structure for repetitive thought: What's on your mind, and how, and how much? *Journal of Personality and Social Psychology*, 85, 909-921.
- Shields, S. A. (2002). *Speaking from the heart: Gender and the social meaning of emotion*. Cambridge, England: Cambridge University Press.
- Shields, S. A. (2005). The politics of emotion in everyday life: "Appropriate" emotion and claims on identity. *Review of General Psychology*, 9, 3-15.
- Shields, S.A. (2007). Passionate men, emotional women: Psychology constructs gender difference in the late 19th Century. *History of Psychology*.
- Shields, S. A., & Crowley, J. C. (1996). Appropriating questionnaires and rating scales for a feminist psychology: A multi-method approach to gender and emotion. In S. Wilkinson (Ed.), *Feminist social psychologies*. Buckingham, Great Britain: Open University Press.
- Showers, C. (1992) The motivational and emotional consequences of considering positive or negative possibilities for an upcoming event. *Journal of Personality and Social Psychology*, 63, 474-484.
- Silvia, P.J. & Abele, A.E. Can positive affect induce self-focused attention? Methodological and measurement issues. *Cognition & Emotion*, 16, 845-853.
- Stemler, S.E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research, & Evaluation*, 9.

- Steele, C.M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*, 613-629.
- Steele, C.M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797-811.
- Stukas, A.A. & Snyder, M. (2002). Targets' awareness of expectations and behavioral confirmation in ongoing interactions. *Journal of Experimental Social Psychology*, *38*, 31-40.
- Timmers, M., Fischer, A. H., and Manstead, A. S. R. (2003). Ability versus vulnerability: Beliefs about men's and women's emotional behavior. *Cognition and Emotion*, *17*, 41-63.
- Vingerhoets, A. J. J. M., Cornelius, R. R., G. L., Van Heck, & Becht, M. C. (2000). Adult crying: A model and review of the literature. *Review of General Psychology*, *4*, 354-377.
- Warner, L. R. & Shields, S. A. (in press). The perception of crying in women and men: Angry tears, sad tears, and the "right way" to weep. In U. Hess & P. Phillipot (Eds.), *Emotion recognition across social groups*. Cambridge University Press.
- Weber, L. (2004). A conceptual framework for understanding race, class, gender, and sexuality. In S. N. Hesse-Biber & M. L. Yaiser (Eds.), *Feminist perspectives on social research*. Oxford University Press.
- West, C. & Zimmerman, D. H. (1987). Doing gender. *Gender and Society*, *1*, 125-151.
- Wild, J. (1965). Authentic existence: A new approach to "value theory." In J. M. Edie (Ed.), *An invitation to phenomenology: Studies in the philosophy of experience* (pp. 59-78). Chicago: Quadrangle Books.

Yoder, J. D., Schleicher, T. S., & McDonald, T. W. (1998). Empowering token women leaders: The importance of organizationally legitimated credibility. *Psychology of Women Quarterly*, 22, 209-222.

Zelditch, M. (2000). Theories of legitimacy. In J.T. Jost & B. Major (Eds.), *The psychology of legitimacy: Emerging perspectives on ideology, justice, and intergroup relations*. (pp. 33-53). New York, NY: Cambridge University Press.

APPENDIX A: Study 1 Legitimacy Scales

Perceived Legitimacy of Observer/Target Argument:

INSTRUCTIONS: Please keep the scenario in mind and indicate your agreement with each of the following statements. Please refer to the scenario if you need to be reminded of what [the target/observer] said. Using the scale below, indicate the appropriate number to the left of each item. Although some items may seem similar, please consider each item individually.

- | | | | | | | |
|----------------------|---|---|-------------------------------|---|---|-------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly
disagree | | | Neither agree nor
disagree | | | Strongly
agree |
-
- _____ 1. [the target/observer's] opinion is rational.
- _____ 2. [the target/observer's] opinion is valid
- _____ 3. [the target/observer's] opinion is illogical.
- _____ 4. [the target/observer's] opinion makes sense.
- _____ 5. [the target/observer's] opinion is incorrect.
- _____ 6. [the target/observer's] opinion is true.
- _____ 7. [the target/observer's] opinion is well thought-out.
- _____ 8. [the target/observer's] opinion is accurate.

APPENDIX B: Factor Analyses for Study 1 Emotion Scales

Targets in the Disagree Condition

Factor analysis for the target in the disagree condition yielded three factors with eigenvalues over 1.0. A scree test suggested a three factor solution which conformed to the three predicted dimensions, together accounting for 63% of the variance. The first factor, Anger (eigenvalue = 4.64), is comprised of 4 items (annoyed, mad, frustrated, and angry, $\alpha = .87$). The second factor, Fear/Anxiety (eigenvalue = 2.11), is comprised of 4 items (nervous, scared, anxious, and afraid, $\alpha = .86$). Happiness (eigenvalue = 1.12), is comprised of 2 items (happy and joyful, $\alpha = .77$). An item was included in a subscale if loaded higher than .30 on one factor, but not others. The “emotional” item was left out of the three factors because it loaded both on the Fear/Anxiety and Anger factors. Thus, it was tested separately from the other factors. Items for each subscale and factor loadings are listed in Table 8.

Table 8
Factor Loadings for the Measure of Target Emotional States When the Observer Disagrees with the Target

	Anger	Fear/Anxiety	Happiness
Annoyed	0.70		
Joyful			0.68
Mad	0.92		
Happy			0.83
Nervous		0.74	
Angry	0.85		
Scared		0.85	
Anxious		0.62	
Frustrated	0.54		
Afraid		0.87	
Emotional	0.42	0.37	

NOTE: loadings below 0.30 are not shown.

Observers in the Emotional Condition

Factor analysis for the observer in the emotional condition yielded three factors with eigenvalues over 1.0. A scree test suggested a three factor solution which conformed to the three predicted dimensions, together accounting for 64% of the variance. The first factor, Anger (eigenvalue = 5.39), is comprised of 4 items (annoyed, mad, frustrated, and angry, $\alpha = .89$). The second factor, Fear/Anxiety (eigenvalue = 1.46), is comprised of 4 items (nervous, scared, anxious, and afraid, $\alpha = .88$). Note that “anxious” also loaded on the anger factor, but because it loaded separately on the fear/anxiety scale in all other instances, I retained it as an item for this

subscale. Happiness (eigenvalue = 1.18), is comprised of 2 items (happy and joyful, $\alpha = .68$). An item was included in a subscale if loaded higher than .30 on one factor, but not others. The “emotional” item was left out of the three factors because it loaded both on the Fear/Anxiety and Anger factors. Thus, it was tested separately from the other factors. Anxiety also loaded on both factors and was not included in analyses. Items for each subscale and factor loadings are listed in Table 9.

Table 9
Factor Loadings for the Measure of Observer Emotional States When the Target is called Emotional

	Anger	Fear/Anxiety	Happiness
Annoyed	0.76		
Joyful			0.57
Mad	0.74		
Happy			0.86
Nervous		0.77	
Angry	0.81		
Scared		0.92	
Anxious	0.42	0.47	
Frustrated	0.68		
Afraid		0.82	
Emotional	0.59	0.35	

NOTE: loadings below 0.30 are not shown.

Observer in the Disagree Condition

Factor analysis for the observer in the disagree condition yielded three factors with eigenvalues over 1.0. A scree test suggested a three factor solution which conformed to the three predicted dimensions, together accounting for 67% of the variance. The first factor, Fear/Anxiety (eigenvalue = 5.29), is comprised of 4 items (nervous, scared, anxious, and afraid, $\alpha = .89$). The second factor, Anger (eigenvalue = 1.90), is comprised of 4 items (annoyed, mad, frustrated, and angry, $\alpha = .89$). Happiness (eigenvalue = 1.10), is comprised of 2 items (happy and joyful, $\alpha = .89$). An item was included in a subscale if loaded higher than .30 on one factor, but not others. The “emotional” item was left out of the three factors because it loaded both on the Fear/Anxiety and Anger factors. Thus, it was tested separately from the other factors. Items for each subscale and factor loadings are listed in Table 10.

Table 10
Factor Loadings for the Measure of Observer Emotional States When the Observer Disagrees with the Target

	Fear/ Anxiety	Anger	Happiness
Annoyed		0.73	
Joyful			0.84
Mad		0.81	
Happy			0.87
Nervous	0.75		
Angry		0.85	
Scared	0.90		
Anxious	0.66		
Frustrated		0.61	
Afraid	0.85		
Emotional	0.46	0.35	

NOTE: loadings below 0.30 are not shown.

APPENDIX C: Correlations Between All Variables in Study 1

Table 11
Correlations Between All Variables in Study 1

	Legit of Tar (Emo)	Legit of Tar (Arg)	Legit of Obs (Emo)	Legit of Obs (Arg)	Tar Anger (Emo)	Tar Anger (Arg)	Tar Scared (Emo)	Tar Scared (Arg)
Legit of Tar (Emo)	1							
Legit of Tar (Arg)	r = .08	1						
Legit of Obs (Emo)	r = -.43*	r = .09	1					
Legit of Obs (Arg)	r = .14	r = -.29*	r = .07	1				
Tar Anger (Emo)	r = -.11	r = -.01	r = .17	r = .09	1			
Tar Anger (Arg)	r = .12	r = .13	r = .07	r = .10	r = .20*	1		
Tar Scared (Emo)	r = -.15	r = -.13	r = .12	r = .06	r = .32*	r = -.04	1	
Tar Scared (Arg)	r = -.01	r = -.11	r = .09	r = .10	r = -.01	r = .36*	r = .06	1
Tar Happy (Emo)	r = .02	r = -.02	r = -.11	r = .01	r = -.31*	r = -.29*	r = -.16†	r = -.15†
Tar Happy (Arg)	r = -.10	r = -.11	r = .01	r = -.07	r = -.08	r = -.44*	r = -.07	r = -.19*
Tar Emoti (Emo)	r = -.27*	r = -.04	r = .31*	r = -.05	r = .42*	r = .09	r = .52*	r = .16†
Tar Emoti (Arg)	r = .14	r = -.20*	r = .10	r = .19*	r = -.004	r = .51*	r = -.08	r = .42*
Obs Anger (Emo)	r = -.01	r = .08	r = -.03	r = -.05	r = .08	r = .06	r = -.08	r = .14
Obs Anger (Arg)	r = .001	r = -.06	r = -.05	r = .04	r = .07	r = .15†	r = -.13	r = .37*
Obs Scared (Emo)	r = -.14	r = -.17*	r = -.06	r = .05	r = .09	r = -.14	r = .34*	r = .08
Obs Scared (Arg)	r = .01	r = .03	r = -.01	r = .002	r = -.09	r = -.09	r = .11	r = -.09
Obs Happy (Emo)	r = -.05	r = -.13	r = .08	r = .10	r = -.16†	r = -.16	r = -.09	r = -.16
Obs Happy (Arg)	r = .03	r = .02	r = .02	r = -.07	r = -.08	r = -.08	r = -.09	r = -.08
Obs Emoti (Emo)	r = -.01	r = -.03	r = -.09	r = .03	r = -.09	r = -.09	r = -.10	r = -.09
Obs Emoti (Arg)	r = -.0	r = .06	r = -.02	r = -.09	r = -.07	r = -.07	r = .001	r = -.07

Note: * signifies that $p < .05$; † signifies that $p < .10$

	Tar Happy (Emo)	Tar Happy (Arg)	Tar Emoti (Emo)	Tar Emoti (Arg)	Obs Anger (Emo)	Obs Anger (Arg)	Obs Scared (Emo)	Obs Scared (Arg)
Legit of Tar (Emo)								
Legit of Tar (Arg)								
Legit of Obs (Emo)								
Legit of Obs (Arg)								
Tar Anger (Emo)								
Tar Anger (Arg)								
Tar Scared (Emo)								
Tar Scared (Arg)								
Tar Happy (Emo)	1							
Tar Happy (Arg)	r = .63*	1						
Tar Emoti (Emo)	r = -.27*	r = -.14	1					
Tar Emoti (Arg)	r = -.23*	r = -.28*	r = .09	1				
Obs Anger (Emo)	r = -.02	r = -.01	r = .003	r = .10	1			
Obs Anger (Arg)	r = -.13	r = -.08	r = -.04	r = .10	r = .31*	1		
Obs Scared (Emo)	r = .10	r = .10	r = -.02	r = -.08	r = .59*	r = .16†	1	
Obs Scared (Arg)	r = .01	r = -.04	r = -.01	r = .19*	r = .24*	r = .46*	r = .25*	1
Obs Happy (Emo)	r = .21*	r = .24*	r = -.03	r = .04	r = -.37*	r = -.31*	r = -.26*	r = -.20*
Obs Happy (Arg)	r = .51	r = .49*	r = -.19*	r = -.08	r = -.13	r = -.50*	r = -.09	r = -.25*
Obs Emoti (Emo)	r = .06	r = -.02	r = -.26*	r = .21*	r = .61*	r = .16†	r = .51*	r = .20*
Obs Emoti (Arg)	r = -.05	r = -.01	r = -.12	r = .15†	r = .30*	r = .49*	r = .28*	r = .52*

Note: * signifies that $p < .05$; † signifies that $p < .10$

	Obs Happy (Emo)	Obs Happy (Arg)	Obs Emoti (Emo)	Obs Emoti (Arg)
Legit of Tar (Emo)				
Legit of Tar (Arg)				
Legit of Obs (Emo)				
Legit of Obs (Arg)				
Tar Anger (Emo)				
Tar Anger (Arg)				
Tar Scared (Emo)				
Tar Scared (Arg)				
Tar Happy (Emo)				
Tar Happy (Arg)				
Tar Emoti (Emo)				
Tar Emoti (Arg)				
Obs Anger (Emo)				
Obs Anger (Arg)				
Obs Scared (Emo)				
Obs Scared (Arg)				
Obs Happy (Emo)	1			
Obs Happy (Arg)	r = .21*	1		
Obs Emoti (Emo)	r = -.14	r = -.12	1	
Obs Emoti (Arg)	r = -.13	r = -.39*	r = .42*	1

APPENDIX D: Study 2 Social Issues Questionnaire

INSTRUCTIONS: For the questions below please indicate your opinions on the following controversial social issues. A brief description of the issue has been provided along with each set of questions.

1. What is your opinion about *same-sex marriage*? (A divisive debate has emerged over whether gay and lesbian couples should be able to become legally married.) *Please circle one answer.*

- a. I support same-sex marriage
- b. I oppose same-sex marriage
- c. I have no opinion on same-sex marriage.

Please indicate how strongly you feel about this opinion by indicating your answer to the left of each item using the scale below.

1	2	3	4	5	6	7
Not at all			Somewhat			Very much

- ___ 1. How strongly do you feel about your opinion on same-sex marriage?
- ___ 2. How certain do you feel about your opinion on same-sex marriage?
- ___ 3. How important to you is your opinion on same-sex marriage?
- ___ 4. How much is your opinion on same-sex marriage related to how you see yourself as a person?

2. What is your opinion about *President George Bush*? (People in the U.S. are divided in their opinions about George Bush's performance as president.) *Please circle one answer.*

- a. George Bush is doing a good job as president.
- b. George Bush is doing a bad job as president.
- c. I have no opinion about George Bush as president.

Please indicate how strongly you feel about this opinion by indicating your answer to the left of each item using the scale below.

1	2	3	4	5	6	7
Not at all			Somewhat			Very much

- ___ 1. How strongly do you feel about your opinion about President George Bush?
- ___ 2. How certain do you feel about your opinion about President George Bush?
- ___ 3. How important to you is your opinion about President George Bush?
- ___ 4. How much is your opinion about President George Bush related to how you see yourself as a person?

3. What is your opinion about *sororities and fraternities*? (Students tend to disagree as to whether they like or dislike sororities and fraternities at Penn State.) *Please circle one answer.*

- a. I support having sororities and fraternities at Penn State.
- b. I oppose having sororities and fraternities at Penn State.
- c. I have no opinion about sororities and fraternities at Penn State.

Please indicate how strongly you feel about this opinion by indicating your answer to the left of each item using the scale below.

1	2	3	4	5	6	7
Not at all			Somewhat			Very much

- ___ 1. How strongly do you feel about your opinion about sororities and fraternities at Penn State?
- ___ 2. How certain do you feel about your opinion about sororities and fraternities at Penn State?
- ___ 3. How important to you is your opinion about sororities and fraternities at Penn State?
- ___ 4. How much is your opinion about sororities and fraternities at Penn State related to how you see yourself as a person?

Table 12

Number of Participants per Essay Type by Gender and Feedback Type

Essay Type	Gender	Feedback Type		Total
		Wordy	Emotional	
Same Sex Marriage	Female	18	8	26
	Male	13	10	23
	<i>Total</i>	<i>31</i>	<i>18</i>	<i>49</i>
George Bush	Female	9	8	17
	Male	8	11	19
	<i>Total</i>	<i>17</i>	<i>19</i>	<i>36</i>
Fraternities and Sororities	Female	4	13	17
	Male	8	6	14
	<i>Total</i>	<i>12</i>	<i>19</i>	<i>31</i>

APPENDIX E: Study 2 Handwriting Analyses

To determine the top 5 pieces of handwriting, I conducted a series of paired-samples t-tests on the “how likely is this a man’s handwriting?” and “how likely is this a woman’s handwriting” for each of the two types of comments (“you shouldn’t have been so wordy” and “you shouldn’t have been so emotional”). Means ranged from 3.66-4.21 (*SDs* from 0.66-1.15) for how likely it was a man’s handwriting, and ranged from 2.00-2.86 (*SDs* from 0.65-1.15) for how likely it was a woman’s handwriting. The t-tests ranged from $t(28) = 2.60, p = .02$, to $t(28) = 10.02, p < .001$. Please see Table 13 for the results.

Table 13
Means and T-Tests for Handwriting Pretesting

Handwriting		Mean (SD) for Likelihood that Man’s Handwriting	Mean (SD) for Likelihood that \Woman’s Handwriting	T-Test
Hand-writing #1	Emotional	3.79 (1.15)	2.52 (1.15)	$t(28) = 3.14, p = .004$
	Wordy	4.17 (0.66)	2.00 (0.65)	$t(28) = 10.02, p < .001$
Hand-writing #2	Emotional	3.75 (1.11)	2.75 (1.00)	$t(28) = 2.78, p = .01$
	Wordy	4.21 (0.73)	2.41 (0.98)	$t(28) = 6.16, p < .001$
Hand-writing #3	Emotional	3.89 (0.98)	2.72 (1.03)	$t(28) = 3.41, p = .002$
	Wordy	3.79 (1.05)	2.66 (1.14)	$t(28) = 2.92, p = .007$
Hand-writing #4	Emotional	3.72 (1.10)	2.86 (0.83)	$t(28) = 2.60, p = .02$
	Wordy	4.07 (0.80)	2.48 (0.69)	$t(28) = 6.33, p < .001$
Hand-writing #5	Emotional	3.96 (0.77)	2.41 (0.91)	$t(28) = 5.17, p < .001$
	Wordy	3.66 (0.87)	2.76 (0.74)	$t(28) = 3.52, p = .001$

Note: Standard deviations are in parentheses

APPENDIX F: Correlations between All Variables in Study 2

Table 14
Correlations between All Variables in Study 2

	Raters' Assessment of Degree to which Thoughts Were About Emotion	Raters' Assessment of degree to which Thoughts Were About Wordiness	Participants' Assessment of Degree to which Thoughts Were About Emotion	Participants' Assessment of degree to which Thoughts Were About Wordiness
Raters' Assessment of Degree to which Thoughts Were About Emotion	1			
Raters' Assessment of Degree to which Thoughts Were About Wordiness	$r = -.08$ $p = \text{n.s.}$	1		
Participants' Assessment of Degree to which Thoughts Were About Emotion	$r = .43$ $p < .001$	$r = .19$ $p = .05$	1	
Participants' Assessment of Degree to which Thoughts Were About Wordiness	$r = .14$ $p = \text{n.s.}$	$r = .44$ $p < .001$	$r = .48$ $p < .001$	1

APPENDIX G: Comparing Participants and Raters in Study 2

To explain this 3-way interaction further, for each target thought focus I conducted a 2 (feedback: emotional or wordy) X 2 (rater type: independent or participant) mixed-design ANOVA. For the emotion thought focus, I found a significant feedback by rater type interaction $F(98) = 11.56, p < .001, \eta_p^2 = 0.11$. For the wordy thought focus, I also found a significant feedback by rater type interaction, $F(98) = 11.56, p = .05, \eta_p^2 = 0.04$. Please see Table 15 for the simple main effect tests of these two interactions.

Table 15

Simple Main Effects Comparing Participants and Independent Raters for the Feedback X Target Thought Focus X Rater Type Interaction

Feedback	Thought Focus	Rater Type	Means (SD)	F-Tests
Emotional	Emotional	Participant	4.78 (1.50)	F(96) = 102.67, $p < .001$, $\eta_p^2 = 0.51$
		Ind. Rater	3.35 (1.86)	
	Wordy	Participant	4.39 (1.46)	F(96) = 76.68, $p < .001$, $\eta_p^2 = 0.22$
		Ind. Rater	1.82 (1.33)	
Wordy	Emotional	Participant	4.25 (1.91)	F(96) = 27.95, $p < .001$
		Ind. Rater	1.50 (0.67)	
	Wordy	Participant	5.20 (1.29)	F(96) = 98.45, $p < .001$
		Ind. Rater	2.94 (1.46)	

Leah R. Warner
Curriculum Vita

Educational History

- 2007 (anticipated) Ph.D., Social Psychology and Women's Studies, The Pennsylvania State University, University Park, PA
- 2003 M.S., Social Psychology, The Pennsylvania State University, University Park, PA
- 2001 B.A., Psychology and Hispanic Studies, Vassar College, Poughkeepsie, NY, with honors in Psychology

Academic Honors and Awards

- 2006-2007 Teaching Fellowship, *The Penn State Psychology Department Teaching Fellowship Program*
- 2005-2006 Florence L. Geis Memorial Dissertation Award, *APA Division 35, The Psychology of Women*
- 2003, 2005 The Carolyn Wood Sherif Memorial Research Fund, *The Pennsylvania State University*
- 2005 Conference Travel Funding from the Research and Graduate Studies Office in the College of the Liberal Arts, *The Pennsylvania State University*
- 2004 Bruce V. Moore Graduate Fellowship in Psychology, *The Pennsylvania State University*
- 2001 Phi Beta Kappa, *Vassar College*

Scholarly Publications

- Warner, L.R. , & Shields, S.A. The perception of crying in women and men: Angry tears, sad tears, and the "right way" to cry. To appear in the forthcoming edited volume: Hess, U. & Philippot, P., Eds. *Group Dynamics and Emotional Expression*.
- Pinel E.C., Warner L.R., & Chua, P.P. (2006). Getting there is half the battle: maintaining diversity in higher education. *Journal of Social Issues*, 61, 181-506.
- Shields, S.A. & Warner, L.R. The emotion politics of emotional intelligence. In preparation for the edited volume: Fineman, S., Ed. *The Emotional Organization: Critical Voices*.

Manuscripts Under Review

- Warner, L.R. & Shields, S.A. (Under Review) Judgments of others' emotions are multidimensional.
- Shields, S.A., Simon, A. Warner, L.R., & Scott, R. (Under Review). The emotion evaluation effect: Naming a winner influences observers' evaluation of debaters' emotion.

Manuscripts in Preparation

- Warner, L.R. Intersectionality and interpersonal perception: Does the treatment of multiple group memberships in stereotype research contribute to intersectionality theory? In preparation for the *Sex Roles* special issue: Shields, S.A. Ed. *Gender: An Intersectionality Perspective*.
- Warner, L.R & Gasper, K. Do I feel for you? The role of mood in everyday experiences of empathy.