

Identity Relativity:  
Linking Stereotype Threat and Social Comparison as Parallel Processes

A DISSERTATION  
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF MINNESOTA  
BY

Suzanne L. Russ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

Dr. David W. Johnson  
Dr. Geoffrey Maruyama

August, 2009

© Suzanne L. Russ, August, 2009

## Acknowledgements

I acknowledge with gratitude and warmth the support, help, and contributions of many people who assisted in varying but critical ways to the completion of this dissertation.

To those I have known at UMN:

I acknowledge with warm gratitude the essential contributions of my advisors, Drs. David W. Johnson and Geoffrey Maruyama, who provided steady support, insightful feedback, helpful instruction, humor, and practical advice over the long course of this degree process. I extend thanks to the members of my committee: to Dr. Michael Harwell, for continually setting high statistical standards that enhanced this study and built my own capacity as a researcher ; to Dr. Mark Snyder, for providing stimulating coursework and facilitating rich discussions; to Dr. Teresa Swartz, for encouragement, wisdom, and calm insight. Thanks also to others from UMN: to Dr. Chris Espin, for humor and realism and insight into the mechanisms of higher education; to Dr. Michelle Everson, for warm generosity and opportunities to assist and teach her well-conceived statistics course; to Dr. Neils Waller, for an incredible class in psychological measurement; to Sharon Sawyer and Lori Boucher, who were patient, prompt, and helpful with so many email questions; and to my fellow students for rich and stimulating discussions.

To my colleagues at the University of Wisconsin:

I offer enduring respect and gratitude to my first mentor, Dr. Bertram Chiang, who offered me my first opportunities in academia and enriched them with guidance, patience, conversations, confidence, encouragement, and wise counsel. Without Bert's guidance and friendship, this path would not even have been visible to me. I also acknowledge the contributions of Drs. Craig Fiedler and Billie Jo Rylance for their instruction, patience, and support during our work together. Finally, I extend warm gratitude to Dr. Wayne Swanger, whose wisdom, warmth, and steady friendship smoothed and enriched the long years of commuting and study.

To my family:

I acknowledge with abounding gratitude the support of my family. To Jim, for his enduring confidence, steady patience, ready backrubs, tolerance of too much time on the computer, and of course the good meals that tided us all over. To Melody, for wonderful conversations that allowed these ideas to develop and for forestalling my tendency to take it all a bit too seriously. To Ben, for the helpful feedback on the soundness of the ideas in real-life people and for making a success of his high school years of his own accord. (and thanks to them both for their valuable role in data collection). To Mack, who held steady and uncomplaining through many years with a half-present mother, and who always welcomed me home with most treasured warmth and love upon my returns. To my mother Ruth, for long conversations as these ideas were being developed, for being such an important grandma when I was not around, for kind friendship. To Betty and Pete, for endless support and encouragement and warmth, and for the means to complete this degree. To Aunt Audie, whose warm and comfortable home provided a welcome respite during a particularly difficult semester.

And to my new friends here in Dickinson, whose warm and lively friendships, encouragement, conversations, laughter, and happy evenings away from the computer made the completion of this process so much easier to endure.

## **Dedication**

*I dedicate this dissertation to the memory of Betty Russ, whose too-short joy-filled presence in this world echoes daily in my thoughts and self and being.*

## **Abstract**

Academic achievement gaps across racial, socioeconomic, and gender groups have persisted in the United States despite formidable recent attention on reducing the disparities. While a wide range of social, personal, and cultural factors contribute to the disparities, the simple knowledge that one belongs to a poorly-performing, stereotyped group can impair performance for even very successful members of stereotyped groups. This phenomenon is known as stereotype threat, and has been documented extensively over more than a decade of research. This study sought to elucidate the underpinnings of stereotype threat through the lens of identity relativity, suggesting that dissonance between and among identities contribute to performance deficits under different levels of identity activation. The hypothesis that stereotype threat is a type of social comparison operating under collective rather than individual identity activation was tested: an experiment was conducted in which test performance and three levels of self-esteem (implicit, explicit, and collective) served as dependent measures after activation of either an intergroup or interpersonal comparison. Subjects included community college students, about whom mild stereotypes are shown to exist. While some predictions were supported, the overall pattern of results did not support the hypotheses. Discussion builds that case that the hypotheses were valid but the methods used to investigate the phenomenon were flawed. Recommendations for redesigning the study are proposed. If the connection between stereotype threat and social comparison can be established, then research in both areas can be used to minimize detrimental comparisons and narrow the gap associated with pervasive negative stereotypes.

## TABLE OF CONTENTS

List of Charts .....	v
List of Tables .....	vi
CHAPTER 1. INTRODUCTION .....	1
1.1 Nature of Identity.....	1
1.2 Identity Relativity .....	3
1.3 Dissonance Among Competing Identities.....	5
1.4 Identity Relativity and Performance .....	7
CHAPTER 2. REVIEW OF LITERATURE .....	10
2.1 Overview of Stereotype Activation Processes .....	10
2.2 Social Comparison: Empirical Findings .....	34
2.3. Theoretical Support for Identity Relativity .....	51
2.4 The Current Research .....	65
CHAPTER 3. METHODS .....	66
3.1 Subjects.....	66
3.2 Manipulations and Measures .....	66
3.4 Manipulation checks .....	71
3.5 Analyses.....	72
3.6 Procedure .....	76
CHAPTER 4. RESULTS.....	78
4.1 Data validation .....	78
4.2 Results of key hypothesis tests .....	83
4.3 Ancillary Analyses.....	97
CHAPTER 5. DISCUSSION .....	102
5.1 Significant and Non-Significant Findings .....	103
5.2 Implications of the findings .....	110
5.3 Future Explorations .....	116
5.4 Concluding Thoughts.....	117
REFERENCES .....	118
APPENDICES.....	125

## LIST OF CHARTS

Chart 1. Boxplot of positively valenced items by condition.....	81
Chart 2. Boxplot of group/individual activation by condition .....	83
Chart 3. Performance by condition.....	84
Chart 4. Explicit self-esteem by condition.....	87
Chart 5. Collective self-esteem by condition .....	89
Chart 6. Test percent as a function of test percent and SCO .....	94
Chart 7. Test percent as a function of condition and DID .....	95
Chart 8. Test percent as a function of condition and ITI.....	97

## LIST OF TABLES

Table 1. Positive responses to valence tests by condition .....	82
Table 2. Group identity activation responses to Twenty Statements task .....	83
Table 3. Test performance by condition .....	85
Table 4. Performance for subjects in upward, downward, and control conditions .....	86
Table 5. Explicit self-esteem by condition. ....	87
Table 6. Collective self-esteem by condition.....	90
Table 7. Signature sizes for subjects in threat conditions.....	91
Table 8. Signature sizes for subjects in boost conditions .....	92
Table 9. Signature sizes for subjects in the control condition.....	92
Table 10. Test percent for UI and all other conditions .....	98
Table 11. ACT scores and test percent for UI and all other conditions .....	98
Table 12. Mean ratings for public CSE, private CSE, ITI, and DID .....	100
Table 13. Paired t-test comparisons for public CSE, private CSE, ITI, and DID.....	101

## CHAPTER 1. INTRODUCTION

One's identity is both dynamic and stable, both individual and collective, both an influence on and reflection of the views of others (e.g., Onorato & Turner, 2002; Yeung & Martin, 2003). One's identity incorporates both individual self-conceptions, in which one experiences the self as a unique being different from all others, and social self-conceptions, in which one experiences the self in conjunction with characteristics and identities of various social roles. Tajfel (1982) envisioned this myriad of identities as concentric circles in which identities closest to the individual (e.g., one's role as family member) are enclosed by gradually more inclusive group identities (e.g., female). Each level of identity holds power to sway individual affect, cognition and behavior when activated (Onorato & Turner, 2004), therein alternately freeing an individual to perform optimally or constraining the individual within the confines of others' expectations. The development of identity, its fluctuation relative to situations, and its influence on task performance provide the foundations on which this study is developed.

### 1.1 Nature of Identity

Earlier perspectives on identity assumed that one's identity represented a core structure that is "stable and chronically accessible" (Markus, 1977; in Onorato & Turner, 2002). This core was derived of a "gradual integration of all possible identifications" (Erikson, 1980, p. 95) possessing "inner sameness and continuity" (p. 94). In a stable and static social environment, this is surely the case: individuals who remain as adults in the same social setting in which they were raised form identities that persist and solidify if unchallenged by new experiences. In a shifting and dynamic environment replete with new encounters and experiences, however, one's sense of identity becomes less stable and more complex (e.g., Bronfenbrenner, 1979). Because identity develops through interactions with other people (e.g., Erikson, 1979; Allport, 1954), expanding one's social horizons concurrently expands and challenges one's sense of identity.

Though the formation of identity is typically associated with adolescence, Erik Erikson (1980) frames identity formation as a process "that neither begins nor ends with adolescence...[but rather as] a life-long development largely unconscious to the individual and to society" (p. 122). Adopting the ideas of Ruth Benedict (1938; in Erikson), Erikson further suggested that identity 'crystallizations,' during which

individuals believe they know themselves, occur throughout the lifespan; such “self-certainty, however,” he continued, “ever again falls prey to the discontinuities of psychosocial development” (p. 122). Thus, one’s identity is composed of core elements that diversify and emerge and are silenced over the course of lifespan development—an entity that is therefore both stable and dynamic.

One’s identity is not constrained to the experience of oneself as an individual, but also consists of the myriad identities associated with one’s social roles. One may be a mother or father, for example, and when activated in that social role (perhaps in the presence of one’s child) may experience the self quite differently than when activated in the role of friend or employee or individual. One may identify as a member of the female gender, and therefore experience the feminine dimension of the self when that identity is activated—such as a situation in which she is the only woman among men.

Bronfenbrenner (1979) captured the essence of this idea when he states that “roles have a magiclike power to alter how a person is treated, how she acts, what she does, and thereby even what she thinks and feels” (p. 6). Thus, one’s identity can also be considered to be both individual and collective, with behavioral patterns that vary according to the level of activation.

Finally, one’s identity influences and is influenced by perceptions of others. An individual who believes he is seen by others as a wild, daring party-hound may find himself enacting this role and assuming this identity (even in his own mind) to align with his presumptions of others’ perceptions. Cooley (1902) termed this phenomenon the “looking-glass self,” exploring in depth the individual characteristics and situations that cause people to be more and less influenced by these presumed judgments of others. The self-uncertainty experienced during transitions, particularly during those formative transitions of adolescents, make some people more vulnerable to fluctuations based on others’ expectations. The concept of looking-glass self also makes it difficult for people to redefine themselves within a situation in which others hold expectations, particularly if the individual has had few opportunities to create an autonomous self-definition.

Although Cooley did not explore the manifestations of group identity expectations, evidence suggests that people are also influenced by expectations with regard to groups to which they belong (e.g., Steele & Aronson, 1995). Generally-held

expectations or stereotypes about groups of people can impel an individual to respond to non-elective social roles that may or may not align with how they perceive themselves as individuals. A crying infant may automatically be placed in the arms of the female in the group, for example, due to stereotypes about women and nurturing ability; the woman is therefore forced to confirm or deny expectations related to her own femininity. A task requiring conscientious attention to mathematical detail may be discretely directed to an Asian person due to stereotyped beliefs about Asian people and math, forcing the Asian person to confirm or disconfirm such academic expectations. Just as the concept of looking-glass self can create an individual from the expectations of others, so can expectations of one's group roles define an individual.

Thus, identity is considered to be stable yet changing, multi-leveled and alternately activated, and intertwined with expectations of others both on an individual and collective level. The next section will explore the mechanisms through which these complex nuances of identity are activated and empowered.

## 1.2 Identity Relativity

Onorato and Turner (2002) expanded on the concept of identity by suggesting that the discontinuities in self-schemata described by Erikson (1980) are prompted by the social composition of influential situations. "Personal identities are salient when interpersonal interactions take place," they reported, "whereas social identities are salient in intergroup settings" (p. 150-1). Thus, an Asian person among like-minded peers may be conscious of the self as an individual, but the same person attending a predominantly White university may be conscious of the self as a member of the Asian race. Both identities are associated with the same individual, but nuances of the situation cause one or the other identity to become salient.

Elaborating, Onorato and Turner (2002) further suggested that the personal identity is activated through situations in which the self and others are sufficiently similar to provide "a context for social comparison" (p. 151). When the group providing the basis for social comparison changes, the identity that the individual experiences also shifts. Likewise, when a particular social role is activated for an individual, social comparisons will tend to occur at the intergroup level (e.g., Tajfel). The comparisons themselves are dependent on relevant domains: if the situation prompts activates the

athletic dimension of the group identity, intergroup comparisons will be made in the domain of athleticism. Such comparisons rest not only on the natures of the groups themselves, therefore, but also on the domain in which the comparison is made.

To exemplify this process, consider the example of a college athlete, Jack, who sees himself both as a strong athlete and as an intelligent young man. Place him in a university at which college athletes have the reputation of athletic prowess and academic incompetence—a not-uncommon stereotype. When he is among his athletic peers, Jack may subconsciously compare himself with those peers at an individual level, and derive positive or negative self-views from the comparisons. If the comparison is made in the academic domain, and most of his peers are academically weak, Jack may feel particularly intelligent and academically skilled through the comparison. Comparisons within his group are made at the individual level, so he compares himself with his fellow college athletes. If he finds himself among a group of college intellectuals, however, Jack's group identity of "college athlete" with all of its positive and negative stereotypes may be activated. He may then subconsciously compare his group—athlete—with the competing group—intellectuals—and derive positive or negative self-views from the group comparison. If the domain of comparison is athletic, his identity as college athlete will provide a boost to his overall view of self. Positive stereotypes about athletic ability of college athletes and negative stereotypes about the poor athletic ability of intellectuals will combine to inflate Jack's sense of athletic competence. If the domain of comparison is academics, however, the same identity (college athlete) may provide a decrement to his self-view by activating the self-stereotype that college athletes are not intelligent. Both the level at which the comparisons are made (i.e., intergroup or intragroup) and the relevance of the comparison domain will serve to produce the accompanying boost or decrement.

It is through concepts such as these that the idea of identity relativity has been conceived. Identity relativity suggests that one's identity is not only multi-leveled (consisting of group and individual dimensions) and not only dynamic (changing in response to stimuli), but that it is also driven by comparisons at the individual level and at the group level. Further, the domain in which comparisons are made and the general expectations of other people combine to determine the valence of the active identity. To

fully understand the implications of identity relativity, however, the consequences of self-uncertainty and competing identities must be explored.

### 1.3 Dissonance Among Competing Identities

The shifts in identity in response to social comparisons are not without emotional impact. As identity crystallizes and dissolves through the course of new interactions, self-certainty is temporarily replaced with self-doubt as individuals redefine themselves within the framework of new comparisons. Comparisons that pose a threat to self-esteem, such as the experience of talented Cricket players who perform poorly when attempting U.S. sports, may create in those athletes a state of disequilibrium as they are forced to adjust their self-schema to accommodate the new and less favorable information about the self. As the former identity—skilled athlete—no longer fits new information, they may experience a sense of discomfort or anxiety in conjunction with the shifting of their old identities. The self-doubt and disequilibrium that accompanies the process of accommodation may, in fact, initially impair their performance in the new sports venue as they cope not only with the learning of a new set of skills (at which they may or may not excel), but also with the uncertainty of self that accompanies the experience. The extent to which an individual experiences anxiety associated with such disequilibrium would then seem to depend upon the extent to which the domain in question is integral to his or her identity.

Particular shifts may be even more anxiety-producing if the expectations of others come into play. For example, a top-performing student from a small high school who is not able to get top grades at a highly competitive college must face not only personal redefinition but also the need to contend with the expectations of others. One might conceive of the conflict between the old and new information as a sort of cognitive dissonance, replete with the arousal and discomfort that accompany such states of dissonance (Festinger, 1957). If the dimension of identity in which the dissonance is felt is particularly important to the individual, then the accompanying discomfort may be sufficiently great that the individual will seek to resolve the dissonance by disidentifying from the domain in question—(“Really, sports are not so important to me; I am more interested in music”) or by improving his skill so that there is no longer dissonance (“See, I am a good athlete regardless of the sport”). Once the shift in identity has been made

internally, an individual might find it easier to act out the new schema outside of the expectations of others; thus, physically relocation to a new social community can allow shifts in personal identity that might be more difficult when expectations of others must be constantly challenged.

Dissonance may only occur not only as personal identities shift, however, but also when identities at different levels hold contradictory expectations. If an individual possesses a personal interest (e.g., academically inclined) that conflicts with a group stereotype (e.g., student athletes are not academically inclined), she may experience dissonance when the group identity is activated. As the group identity becomes salient to the individual, the full spectrum of stereotypes about the group is activated regardless of whether or not the individual agrees with or fits the stereotypes. As she seeks to reconcile the non-resolvable discord between a personal identity and unchangeable stereotypes about her group identity, the resulting tension may in fact impair her ability to perform in the domain in question. The impaired performance due to dissonance-based arousal may then in fact serve to reinforce the very stereotype that created the tension.

Such performance-impairing arousal as a consequence of cognitive dissonance is not a speculative relationship; in fact, the relationships among dissonance, arousal, and test performance appear sound. In his original conception of cognitive dissonance, Festinger (1957) suggested that “the experience of dissonance, being cognitively uncomfortable, will motivate a person to try to reduce the dissonance and achieve consonance” (p. 3). This motivational state associated with dissonance is considered to be analogous to the motivational state associated with hunger—that is, a physiological arousal that motivates behavior and influences performance. A critical analysis of literature related to dissonance, physiological arousal, and task performance demonstrated the connectedness of these experiences: indeed, the physiological state associated with hunger resembled the physiological state associated with dissonance, and this state of arousal impaired cognitive performance (Kiesler & Pallak, 1976). Dissonance between stereotypes about one’s group may be still more challenging, in fact, than dissonance between one’s individual self-views and others’ expectations for the self. While an individual for whom negative expectations are held due to reputation may move to a new location where others’ expectations cease to create the self, stereotypes about a

social identity often permeate a culture and are therefore very difficult to escape. Thus, stereotypes about one's group identity that conflict with personal beliefs about oneself can produce consistent performance-impairing deficits with little opportunity for escape.

If one's group is presumed to be inferior, the comparison may be detrimental to identity in the same way than it would be if one was presumed to be inferior as an individual. In both cases, one must seek to disprove the presumption and resolve the dissonance between others' expectations and one's own self-beliefs. The attempt to resolve the dissonance may then occupy sufficient cognitive space that one's performance in the very domain in question is impaired.

#### 1.4 Identity Relativity and Performance

During adolescence and young adulthood, the period with greatest intrapersonal uncertainty and therefore greatest susceptibility to threats from identity relativity, individuals are also with faced with a myriad of situations for which their performance serves as a critical determinant of their futures. From the myriad standardized tests that determine or rule out future educational possibilities to individualized measures such as job interviews that determine one's employment future, these years are replete with potential for progress or barriers. In this age of accountability, in fact, performance on a single test can determine the university one may attend, whether or not graduate school will be an option, which careers will be open to an individual, and even the future of the school from which one came. Thus, a person's capacity depends in many respects upon the accuracy of such tests.

Nevertheless, one's performance on such measures is not always an accurate reflection of one's ability. On a given occasion, an individual might overperform, providing an overly optimistic view of true ability, or underperform such that one's true ability is never fully recognized. This is to be expected: two people with identical academic ability would not be expected to produce identical test scores in a given situation. One of the two might be tired and distracted during the test, while the other might have just received a confidence-boosting award. One might be newly in love, while the other was harried from missing the bus or waking up late. Such random influences on test performance are expected and are in fact incorporated into formal test reports when the measure is of particular diagnostic importance (e.g., Novick, 1966).

Some individuals or situations, however, might produce systematically depressed test scores such that the true capacity of the individual is never captured. People whose performance is always measured in the shadow of an ostensibly superior peer, for example, might make self-deflating social comparisons that consistently impair their performance. Likewise, people whose performance is always measured in the shadow of ostensibly superior social groups might make self-deflating group comparisons that impair their performance as representatives of the subordinate group. The comparison peer or group need not actually possess superior ability; the mere perception of a difference in the mind of the test-taker could produce a deficit. If an individual is constantly barraged by such real or perceived comparisons, she or he will underperform on academic measures and may disengage from academics in favor of more ego-enhancing (but less gratifying) pursuits.

Both of these phenomena have been illustrated empirically. Studies of social comparison have found performance shifts in conjunction with positive or negative individual comparisons (e.g., Johnson & Stapel, 2007b), and studies of stereotype threat have revealed a relationship between stereotype activation and individual performance (e.g., Steele & Aronson, 1995). The two phenomena have not previously been linked as parallel processes, however, instead encompassing separate strands of research with little empirical overlap.

This paper will build the case and test the hypothesis that social comparison and stereotype threat are in fact parallel processes that operate under the same mechanisms when different levels of an individual's identity are activated. In each phenomenon, the relativity of one's identity when compared with others becomes an ego-inflating or ego-deflating experience; only the level of identity activation differs (i.e., collective or individual identity). Evidence will be presented through a variety of research strands, including in-group out-group theories (Allport, 1954), self-esteem maintenance (Tesser, 2000), self-categorization theory (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987), and factors moderating and motivating both social comparison and stereotype threat. If social comparison and stereotype threat are indeed parallel processes operating under the shared theme of identity relativity, then research in one area can inform and guide the other to enrich understanding of both.



## CHAPTER 2. REVIEW OF LITERATURE

When perceptions of the self or one's group are made salient, behaviors and performance can shift in response to the perceptions. Empirical findings related to these phenomena will be explored in the following subsections: (a) overview of stereotype activation processes, including self-relevant stereotypes, implicit self-stereotypes, competing identities, and activation of non-self-relevant stereotypes; (b) overview of social comparison processes, and (c) the theoretical basis for linking the two phenomena.

### 2.1 Overview of Stereotype Activation Processes

A stereotype is a commonly held beliefs about a particular social category or, in the words of Gordon Allport (1954), "an exaggerated belief associated with a category" (p. 191). A stereotype is distinct from a social category in that a category "can be held in mind simply as a neutral, factual, nonevaluative concept" (p. 191), whereas a stereotype is a "fixed idea that accompanies the category" or a "fixed mark upon the category" (p. 192). Generally-held beliefs that Black people have poor intellectual ability (or, conversely, strong athletic ability), or that women have poor math skills (or, conversely, strong verbal skills) are examples of such stereotypes. Because "the category saturates all that it contains with the same ideational and emotional flavor" (Allport, p. 21), the activation a social category will concurrently tend to activate all of the valenced stereotypes associated with that category.

Because "the human mind must think in terms of categories" (Allport, 1954, p. 20) in order to maintain efficient and orderly living, such categorical thinking occurs automatically and often outside of the deliberate control of the individual. If a person attaches no particular importance to the category, then activation of the category (e.g., elderly) will concurrently activate a plethora of cognitive associations (e.g., forgetful) with little personally-relevant emotional content. Such associations may lead to shifts in behavior or performance as a consequence of ideomotor processes, or the increased likelihood that a person will engage in a behavior when they are thinking about it. This is not a new idea: William James described this process as early as 1890, explaining that "every representation of a movement awakens in some degree the actual movement which is its object" (p. 526; in Bargh, Chen, & Burrows, 1996, p. 231). Empirical evidence has in fact shown that ideomotor processes can impact both academic

performance (e.g., Wheeler, Jarvis, & Petty, 2001) and other more general behaviors (Bargh & Chartrand, 1999).

Self-relevant stereotypes are general beliefs about one's in-group, or beliefs about one of the many social groups with which a person shares a common identity. If an individual identifies strongly with the in-group, the reasoning suggests, then activation of the stereotype will not only activate cognitive associations such as those described above, but also a range of personally-relevant emotional responses. If the individual happens also to possess a characteristic that contradicts the stereotype, such as being an intellectually-gifted Black person, then the emotional content of the stereotype activation is increased through the experience of dissonance. And finally, if that person is then placed in a situation where his or her own performance is presumed to be diagnostic of his or her overall ability in the domain, the immediate toll may be distinctive and severe. The nuances of this phenomenon—the mechanisms and processes and conditions of the performance toll associated with stereotype threat—have been abundantly explored in recent years under the heading of stereotype threat.

Stereotype threat is a recently-coined term (Steele & Aronson, 1995) for a phenomenon documented several decades earlier (e.g., Katz, Epps, & Axelson, 1964). The term refers to the tendency for stereotyped individuals in stereotype-relevant diagnostic situations to perform in accordance with the stereotype. If a situation activates the stereotype that old people are forgetful, for example, a normally astute older person may unintentionally confirm the stereotype with faltering memory in a memory-driven performance. A woman confronting the stereotype that women have poor spatial skills may find that her performance in a diagnostic exam is lower than it would have been in a neutral situation.

Research into the effects of stereotype activation tends to follow a fairly standard structure. In general, subjects in stereotyped groups (e.g., Black people, women) and non-stereotyped groups (e.g., White people, men) are assigned to a condition in which a stereotype is activated or is not activated before a test is taken. Stereotype activation methods range from checking a box pertaining to one's race on a demographic form (Steele & Aronson, 1995), to being the only female in a group of three (Inzlicht & Ben-Zeev, 2000), to indicating agreement with a list of stereotype statements (Shih, Pittinsky,

& Ambady, 1999). Dependent measures tend to involve validated tests, such as the Graduate Records Exam, along with ancillary measures of potential mediators or moderators. Stereotype threat predicts that members stereotyped groups will perform more poorly when their collective identity is activated than when it is not activated

For the purposes of this section of the literature review, stereotype activation will focus primarily on the processes and moderators of stereotype threat. Evidence related to psychological processes mediating stereotype activation and performance deficits is beyond the scope of this review. In conjunction with the framework used in a review of literature by Wheeler & Petty (2001), both positive and negative stereotypes and both self-relevant and other-relevant stereotypes will be discussed in varying levels of detail. For the section related to stereotype threat, the term “assimilation” shall refer to situations in which the behavior of a subject conformed to the stereotype, and the term “contrast” will refer to instances in which the behavior of a subject shifted in the opposite direction of the stereotype. Detailed information about stereotype activation mechanisms, or primes, will be provided for some studies, as the different activation mechanisms can sometimes serve to explain contradictory findings.

### *2.1.1 Self-Relevant Stereotypes*

Self-relevant stereotypes, or self-stereotypes, refer to stereotypes that pertain to a group to which one belongs. Such stereotypes are often negative (e.g., “Athletes aren’t very smart”), but positive stereotypes also exist (e.g., “Asian people are smart”). Stereotypes can be explicit, pertaining specifically to one’s in-group, or implicit, in which the existence of a stereotype about an out-group implies an opposite belief about one’s in-group. Finally, one individual can have several stereotyped identities (e.g., a Black woman) or a single identity with opposing stereotypes (e.g., elderly: forgetful, wise). The following paragraphs present key findings related to negative self-relevant stereotypes, implicit stereotypes, and competing identities as a means of providing a lens into the complexity of the effect.

#### *2.1.1.1 Negative Self-Relevant Stereotypes*

Negative stereotypes exist about many social categories, but some stereotypes influence performance in the domains most critical to outward success—intellectual or academic ability measures. The following paragraphs will explore racial and gender

stereotypes, followed by other stereotypes that have also yielded performance deficits (i.e., stereotypes about student athletes, people from lower socioeconomic classes, and older people).

2.1.1.1.1 Racial Stereotypes. Stereotypes about Black people and intelligence have been shown to impede performance of Black students in at least one study dating back to the 1960s (Katz, Epps, & Axelson, 1964). In the study, Black and White students took a math test believing their results would either: (a) be compared with other same race students at their respective segregated universities, (b) be compared with only White students used to set national norms, or (c) not be compared with others. Results indicated that Black students performed significantly better when their results were only to be compared with other Black students than in either of the other conditions, and that this effect occurred only when the task was difficult. No such effect was observed for White students. These findings offer intriguing insights, though the social context in which the experiment was conducted--segregated southern universities in a pre-civil rights era—limits the applicability of the findings in different contexts.

Similar effects were found, however, in one of the most defining studies by Steele and Aronson (1995). In their study, Black and White college students took a difficult verbal test after being assigned to either a diagnostic condition (test indicative of intellectual ability) or a non-diagnostic condition (test was a laboratory problem-solving task; Studies 1, 2, and 3). The diagnostic condition was intended to indirectly activate a stereotype about Blacks and intellectual ability by creating a stereotype-relevant situation, and a manipulation check in a third experiment confirmed that the stereotype was indeed activated. Findings indicated that Black students in the diagnostic condition received significantly lower scores (after adjusting for SAT scores) than Blacks in the non-diagnostic condition and than Whites in either condition. Further, Blacks in the diagnostic condition spent significantly more time on each item than did Blacks in the non-diagnostic condition and Whites in either condition (Studies 2 and 4). The stereotype activation cue did not need to be overt; even the simple act of checking one's race in a demographic checklist was sufficient to elicit a performance decrement for Black subjects who were required to do so. These powerful findings spawned a wave of exploration into the extent of the effect, though misinterpretation of the results led to

exaggerated conclusions regarding the role of stereotype threat in moderating the academic achievement gap (e.g., Sackett, Hardison, & Cullen, 2004).

Few stereotypes exist about White males, but popular media and sports commentators “provide illuminating commentary on the ‘possibility’ that Black athletes are physically superior to Whites, but that White athletes are intellectually superior to Blacks” (Stone, Perry, & Darley, 1997, p. 293). One group of researchers tested the effects of these stereotypes in an interesting experiment comparing golf performance of White and Black athletes when the test was framed as either a measure of sports intelligence (negatively stereotyping Black subjects), natural ability (negatively stereotyping White subjects), or of the “psychological factors correlated with general sports performance” in conjunction with a demographic questionnaire (race-prime condition) or non-primed control condition (Stone, Lynch, Sjomeling, & Darley, 1999, p. 1216). Results indicated that the respective threat conditions impaired performance for both Black and White subjects when compared with a control group. Black athletes in the sports intelligence and race-prime condition performed significantly worse than those in the natural ability or control conditions, and White athletes in the natural ability or in the control condition performed more poorly than those in the race prime or sports intelligence conditions. The race-prime condition yielded a performance decrement for Black subjects and a performance boost for White subjects. This may have occurred because the test instructions, intended to be neutral, in fact was biased toward White subjects through the use of the phrase “psychological factors” in both the race-prime and control conditions (potentially associated with intelligence in the minds of subjects).

While stereotype threat has been found for Latinos (e.g., Gonzales, Blanton, & Williams, 2002), studies have been infrequently found in published articles. Gonzales’ study, emphasizing stereotype threat for Latinos and women, will be discussed along with Asian stereotypes in the section on competing identities. Studies about stereotype threat among Native Americans were not available at the time of this writing.

2.1.1.1.2 Gender Stereotypes. The other widely studied group for whom a negative stereotype appears to wield power is women, whose performance is reduced in response to the stereotype that they have poor math and spatial skills. This effect was initially documented in a study by Spencer and Steele (1999) in a series of three

experiments paralleling the design used by Steele and Aronson (1995). Experiment 2 is most meaningful to the current exploration. In the experiment, male and female undergraduates who valued math and had strong math performance records were given a difficult math test after being told that the test revealed gender differences (gender-differences condition) or did not reveal gender differences (no-gender-differences condition). Men and women performed equally well in the no-gender-difference condition, but women underperformed in the gender-difference condition relative both to men and to non-threatened women.

Other studies have illustrated the stereotype threat effect on women using different primes. In one case, math performance decrements were discovered among women whose female identity was activated through a questionnaire about coed vs. single-sex housing (Shih et al., 1999). Another study found that female subjects exposed to subliminal primes with female words in a vigilance task performed more poorly on a math test than did female subjects who were exposed to neutral words (Ambady, Paik, Steele, Owen-Smith, & Mitchell, 2004). Finally, being the sole representative of a stereotyped group, such as being the only woman in a group of men, was sufficient to elicit stereotype threat responses in women (Inzlicht & Ben-Zeev, 2000). Because each of these studies also explored some additional aspect of stereotype threat, they will be discussed in more detail in later sections.

*2.1.1.1.3 Socioeconomic status (SES) stereotypes.* Stereotype threat has been associated with diminished performance for students from lower socioeconomic status (Croizet et al., 2002). The researchers made use of a recent shift in the French education system which had broadened access to baccalaureate degrees, the prerequisite for admission to a university. The newly added students were primarily from lower socioeconomic classes who received a technical rather than the more prestigious general baccalaureate. The technical baccalaureate was associated with lower success rates at the university, and a survey of university professors revealed that students with a technical degree had a poorer reputation in the academic domain. Thus, the two socioeconomic classes provided reasonable comparison groups for studying the effect of stereotype threat. The researchers used a manipulation similar to the one used by Steele and Aronson (1995) in their study of stereotype threat among Black students—they simply

told one randomly assigned group that the test was diagnostic of intellectual ability and told the other group that the test was being used to test a psychology hypothesis. A test containing items similar to those found on the Graduate Records Exam (GRE) served as the dependent variable. As expected, low SES students in the threat (diagnostic) condition performed significantly more poorly than did students in any of the other four groups. This study provides evidence for the power of socioeconomic status to influence performance outcomes in ways beyond the broad general challenges of poverty.

*2.1.1.1.4 Student athlete stereotypes.* In an interesting study of another sort, Yopyk and Prentice (2005) studied stereotype threat among student athletes. To do so, they selected students from Princeton University who were either part of a hockey or football team or (as a comparison group) part of the a capella singing group. Subjects were required to take a math test after responding to a writing about either a recent experience in their respective extracurricular activities (activating the identity associated with the extracurricular activity), a recent successful student experience (activating the identity associated with being a student), or directions to the library (in a neutral control condition). The analysis confirmed score differences by identity activation: in the student athlete condition, significantly lower accuracy scores were found in the athlete-prime condition than in the no-prime or student-prime condition. No such differences were found among the student singers, for whom a negative academic stereotype does not exist. Additionally, academic self-regard among athletes was significantly greater in the student-prime condition than in the extracurricular-prime or no-prime conditions. No similar differences were found for singers, or for either group in general self-regard.

While the researchers emphasized the finding that academic self-regard was significantly lower in the athlete than in the student identity activation state, it must be noted that no significant differences were found between self-regard in the control group and the athlete-identity group. The self-regard differences, therefore, may have been a consequence of a positive affirmation in the student condition (writing about the successful academic experience) rather than a decline in academic self-esteem in the other two conditions. The test used to assess self-regard emphasized feelings related to academic capacity, which was likely to have been inflated after writing about a successful academic experience. Test performance was not likely to have been similarly affected,

however. This is consistent with earlier studies, which found that the positive self-concept that follows such affirmation reduces the need to salvage self-esteem through an improved performance. (e.g., Johnson & Stapel, 2007a)

*2.1.1.1.5 Stereotypes about older people.* Many stereotypes exist about older people, and the stereotypes are both favorable (e.g., older people are wise) and unfavorable (e.g., older people are forgetful). In a rather complex experiment, Levy (1996) explored the impact of both positive and negative stereotypes on performance in a variety of recall and attitude measures. The experiment used a version of the vigilance task to prime either negative (Alzheimer's, senile) or positive (wise, sage) elderly stereotypes. A control group saw neutral words of the same length as the valenced words. The complexity in this study emerges from the use of numerous dependent measures, a pretest-posttest design, and the inclusion of a range of factors (e.g., subjects were told that a lamp in the room was thought to improve memory, thus providing subjects with an external attribution for improved performance). For the purpose of this review of literature, findings related to memory tests are the most relevant and thus the only aspects included herein.

Significant differences in recall measures occurred between negatively- and positively-primed groups. Those subjects primed with positive-elderly words experienced a significant increase in their performance on three recall tasks. Conversely, subjects primed with negative-elderly words experienced a decline in several of the recall tasks. The authors refer to this as a self-stereotyping effect, and in this study it may be more aligned with a priming effect than with stereotype threat. Because the subjects were presented with trait words that conveyed a particular image of older people, it may have activated a cognitive schema rather than an identity associated with a group to which they belong. The researchers might have explored this by including a control group of younger subjects, who might have experienced a similar effect in accordance with priming effects. Nevertheless, the study is generally accepted as evidence of stereotype threat among older people, and provides strong evidence for the potential of negative stereotypes about older people to influence their performance.

*2.1.1.2 Implicit Self-Stereotypes*

While stereotypes are often conceived of as negative beliefs about a particular group, positive stereotypes about particular groups can also exist (e.g., “Asians are good at math”). Some researchers have explored the possibility that, just as subjects’ performance might decline in fulfillment of a negative stereotype, so might performance improve in fulfillment of a positive stereotype. Several experiments have explored the possibility of such an effect, finding evidence that positive stereotypes can indeed boost performance. These studies have focused on stereotypes related to Asian people (good at math; e.g., Shih, Pittinsky, & Ambady, 1999) and students at elite universities (academically competent; e.g., McGlone & Aronson, 2006). Both of these studies will be explored in subsequent paragraphs on competing identities.

More often, however, positive stereotypes about an in-group are only inferred from the existence of negative stereotypes about an out-group—implicit self stereotypes. For example, the stereotype “women have poor spatial/math skills” harbors an implicit opposite “...compared to men.” Likewise, the belief that “Asians are good at Math” is followed by the unstated suffix “...compared to non-Asians.” Despite their implicit nature, studies have found that knowledge of stereotypes about an out-group—an implicit stereotype--can generate a contrast effect for an individual for whom the stereotype is not relevant. In other words, the stereotype that Asians have good math skills or that women have poor math skills can impair or enhance respective math performance for White males (Aronson & Lustina, 1999; McGlone & Aronson, 1996), although no specific negative or positive stereotype appears to exist that pertains to White males and math.

When performance is enhanced through implicit self stereotypes, it is termed stereotype lift by Walton and Cohen (2003), who describe such an effect as a “performance boost caused by the awareness that an out-group is negatively stereotyped” (p. 456). No particular term identifies the phenomenon in which performance of a non-stereotyped group suffers after activation of a positive out-group stereotype. In an extensive meta-analysis of stereotype threat studies, Walton and Cohen found that non-stereotyped subjects performed better on a stereotype-relevant test than did non-stereotyped controls or stereotyped individuals. In other words, meta-analysis of the forty-three studies meeting their criteria for inclusion revealed significant performance benefits for members of non-stereotyped groups (e.g., males) when a stereotype about an

out-group (e.g., females) was activated. The vast majority of the studies included in this analysis dealt with White males as the group identified with the lift effect; in fact, only two other non-stereotyped groups were involved in the studies meeting criteria--high-SES students relative to low-SES students (Croizet, Dutrevis, & Desert, 2002), and younger people relative to older people (Ewing & Smith, 2001; in Walton & Cohen). Because no stereotype exists about those experiencing performance boosts, the performance boost experienced by non-stereotyped subjects may only be attributed to the implicit comparison with a negatively-stereotyped out-group.

This finding that stereotype lift was not consistently present in studies demonstrating stereotype threat could be explained by the reduced likelihood of social identity activation for non-stereotyped groups (Marx & Stapel, 2006c) For strongly stereotyped groups, they hypothesized, a threat such as a diagnostic intellectual test would automatically activate a group identity and impair performance, but social identity activation for non-stereotyped groups would be less automatic. They confirmed this hypothesis by comparing math performance for males and females after activating individual or group identities, finding that “the lift effect increased as the manipulation of social self became more direct” (Marx & Stapel, p. 784). Similar results were found by Shih and others (2002), who discovered that Asians but not non-Asians experienced a performance boost when Asian terms were subliminally activated.

Research has also shown that implicit comparisons with positively-stereotyped out-groups can generate performance declines—an effect precisely the opposite of stereotype lift. One study illustrated this effect by activating a stereotype about Asians (good at math) before asking White males to take a difficult math exam (Aronson & Lustina, 1999). Their subjects were twenty-three White males who were students at Stanford University and were highly identified with the math domain. The threat condition was generated by requiring randomly assigned subjects to skim a series of news articles about Asian math superiority. They were then told that they would be taking a math test that sought to understand the phenomenon of Asian mathematical superiority. Subjects in the control condition were exposed to neither the news articles nor the test instruction related to Asians. Both groups took the same test, which consisted of sample items from the Graduate Records Exam (GRE). Participants in the threat condition solved

significantly fewer items than those in the control condition, and also reported that they had expended significantly higher effort than those in the control condition. No differences in anxiety were reported. This study demonstrates that comparison of one's in-group with a reputedly superior out-group caused a stereotype threat effect even though the stereotype itself was not self-relevant.

### *2.1.1.3 Competing Identities*

Although most of the studies of positive stereotypes fell into the area of multiple identities, some identities have multiple differently-valenced stereotypes. When different aspects of an identity are activated, performance can vary upward or downward. Levy's (1996) study of stereotype activation among older people, for instance, revealed that positive stereotypes (wise) enhanced performance just as negative stereotypes (forgetful) impaired performance. Likewise, in the study by Stone and others (1997), Black and White students contended with conflicting stereotypes associated with their racial group (e.g., Black people have good athletic ability; Black people have poor intellectual ability). In both cases, the performance consequences of identity activation depended on which facet of the identity becomes salient.

In other cases, however, people experience differently-valenced stereotypes in conjunction with different self-identities. Asian women, for instance, must contend with the conflicting stereotypes about their Asian identities (good at math) and their female identities (poor at math). Females attending elite universities must contend with the discord created by negative stereotypes about their female identities (poor at math) and positive stereotypes about their student identity (good at academics). McGlone and Aronson (2006) explored this discord and found that activation of the identity associated with a positive stereotype (student at elite college) can offset the problems associated with a negatively stereotyped in-group (female). After activating subjects' identities as elite students or as female/male (or control, subjects took a test involving spatial ability. Analyses revealed that women in the student-identity condition performed significantly better than those in the control condition, while women in the gender-identity condition performed more poorly than controls. A stereotype lift effect was evident for males: those in the gender-identity condition performed significantly better than those in the control or the student-identity conditions.

A similar problem with conflicting identities exists for Asian women, for whom stereotypes related to math ability contradict one another. In a study of this effect, conflicting identities (Asian, female, or neutral) were activated before subjects took a difficult math test using questions from the GRE) (Shih et al., 1999). Asian identity activation involved a questionnaire about topics such as their grandparents and their language in the home; the female activation condition involved questions about single-sex and coed dorm life; and the control condition offered an entertainment survey. Results supported the expectation of stereotype activation. When their female identity was made salient, accuracy was lowest; when their Asian identity was activated, accuracy was highest, and scores for the control group were in the middle of the other two groups. A later study found similar results for Asian females under subtle and blatant stereotype activation (Shih, Ambady, Richeson, Fujita, & Gray, 2002). Subtle activation involved the same questionnaire described above, and blatant activation involved requesting that subjects provide Likert-type responses to a series of statements about Asian stereotypes. Analysis revealed that Asian women in the subtle activation condition had significantly greater accuracy on a math exam than those in the blatant or no-prime condition.

Different results were found in a similar study of Asian female stereotypes by Cheryan and Bodenhausen (2000). In the experiment, they used the same independent and dependent variables as Shih and others (1999), but activated the stereotype using a far more direct mechanism: they adapted items from a Collective Identity Esteem test, in which subjects responded to statements such as “I am a worthy member of the race/gender I belong to” and “My race/gender is considered good by others” (Cheryan & Bodenhausen, p. 400). A control group was included in which subjects answered questions about their individual identities. Unlike the results by Shih and colleagues, women whose Asian identities were activated had lower math accuracy than controls. No differences between gender-activated women and women in the control group were observed, thereby also contradicting the abundant findings that women suffer performance decrements when their female identity is made salient. Cheryan and Bodenhausen attribute the performance decrement among Asian students to a “choking effect,” in which a poor ability to concentrate was thought to mediate the relationship between the salient identity and their performance. “Positive stereotypes,” they suggest,

“can place a considerable burden on members of the stereotyped group, adversely affecting their performance in the stereotyped domain” (p. 401). No explanation was offered for the absence of a typical female-activation decrement.

Lastly, a study of stereotype threat among Latino men and women explored the dual stereotypes related to Hispanics (poor intellectual ability) and to women (poor math ability; Gonzales, Blanton, & Williams, 2002). They hypothesized that the dual effects of two negative stereotypes would have a cumulative effect on Latino women, such that the effects would be more pronounced for Latino women through a double-minority effect than for Latino men. Results confirmed the existence of a double-minority effect, in that “Latinos experienced greater gender-based stereotype threat effects than Whites” (Gonzales et al., p. 665). Although the impact of stereotype threat was experienced by Latino women, no similar effect was found for White women, for whom no differences existed across conditions. This unexpected finding will be discussed along with others in the following paragraphs.

*2.1.1.3.1 Exploring contradictions.* Several contradictions confound the clarity of the effects of competing identities. For example, Cheryan and Bodenhausen (2000) found a contrast effect for Asian women under Asian-identity-activation conditions, while most studies (e.g., Shih et al., 1999) found assimilation effects. Further, the expected gender-based stereotype threat for White women was absent in the study by Gonzales and others (2002). Finally, the study by Shih and others (2002) found that Asian women exposed to blatant stereotype activation did not experience the expected assimilation effect. The following paragraphs will consider these disparities as they relate to the overriding theme of this paper.

Treatment effects may have accounted for some of the contradictions. In one of the studies by Shih and colleagues (2002, Study 1), for example, the blatant prime condition did not elicit the expected performance boost. The prime itself may have created the problem, in that it required that subjects indicate the extent to which they agreed with well-known positive and negative Asian stereotypes. The researchers contended that this priming mechanism would directly activate schemas related to Asian stereotypes. In order to respond to the stereotype statements, however, subjects most likely considered the applicability of the statements to themselves and to other Asians. As

such, the prime would most likely have activated an individual identity (“I am different from other Asians”) rather than a group identity (“we Asians are very similar”). Further, evidence has suggested that blatant primes allow time for cognitive corrections and thus nullify the effect of the prime (Gilbert, Giesler, & Morris, 1995) . Thus, the finding that no performance differences existed between blatantly primed and non-primed Asian subjects could be attributed to individual identity activation and/or to cognitive corrections rather than to the effect of stereotype activation suggested by the authors.

A similar prime effect may also have created the contrasting findings in the study by Cheryan and Bodenhausen (2000). The self-identity scale used to prime the Asian-identity condition may in fact have instituted intra-group comparisons in which Asian subjects compared themselves unfavorably with other Asian people. When subjects were asked to respond to statements such as “I am a good member of my racial group,” they may have activated their individual (rather than collective) identities as they assessed the accuracy of the statement. Individual identity activation might have lead to in-group comparison with superior in-group members, resulting in the expected performance deficit following upward comparison. Rather than a choking effect under activation of high group-based expectations, it is possible that the effect was one in which the subjects’ individual identities were activated within the framework of their own highly-performing in-group. The expectation for success was not the result of what out-group members believed about Asian people, but about their own relative individual status within the talented group to which they belong. When subjects were primed in such a manner that emphasized the math superiority of Asians compared to other groups (e.g., emphasizing their family’s accomplishments), the expected performance boost was found. The “choking” effect, therefore, may in fact have been a consequence of individual intragroup comparisons.

The absence of gender-based stereotype threat effects among White women in the study of Latino experience might also be explained in this way, although the authors do not address it (Gonzales et al., 2002). Because White and Latino subjects were tested simultaneously, two primes were actually present: the test diagnosticity and the presence in the test room of racial in-group and out-group members. The potential threat associated with the female identity activation might readily have been offset by the boost possible

through exposure to negatively-stereotyped out-group members. For White women, the presence of many Latino people in the test center might in fact have activated a stereotype lift effect that nullified the potential stereotype threat associated with being female. Latino women could not have experienced such a boost, as both of the cues pointed consistently in the same negative direction. The authors suggest that “being a Latino focuses an individual on group more generally, which then causes them to be affected by other relevant group stereotypes, such as those related to gender” (p. 666).

### *2.1.2 Activation of Non-Self-Relevant Stereotypes*

Thus far we have explored (a) the potential impact on performance that negative and positive stereotypes about one’s in-group can generate, and (b) the performance lift associated with activation of implicit self-stereotypes in the face of out-group stereotype activation. In both areas, the majority of the cases yielded assimilation with the self-stereotype (stereotype threat) or the implied stereotype (stereotype lift). Other studies, however, have found different results. For instance, some studies have found that stereotype activation yields assimilative effects in people who are not members of the stereotyped groups (e.g., Wheeler et al., 2001) or that outcomes vary when different experimental primes are used (e.g., Shih et al., 2002). A brief overview of these studies is necessary to help in distinguishing stereotype-based identity activation from ideomotor effects.

#### *2.1.2.1 Priming Effects*

John Bargh and colleagues have conducted extensive explorations of the impact of exposing subjects to cues or “primes” outside of conscious awareness and measuring a specific behavior (e.g., Bargh & Chartrand, 1999; Bargh, Chen, & Burrows, 1996; Higgins, Bargh, & Lombardi, 1985). Primes are thought to influence behavior through ideomotor action, in which “internal (through ideation) activation of a representation (i.e., through imagining the behavior) increases its accessibility of likelihood of activation (Bargh et al., 1996a, p. 231). Mental constructs, or the representation to which the previous statement refers, are thought to be activated through the additive function of “one’s goals, external environmental events, long-term use, and recent thought” (Bargh et al., 1996a, p. 231). In relation to stereotypes, then, priming a particular social category could activate a network of traits associated with that category, which would, in turn,

increase the likelihood of an individual's behaving in conjunction with the activated stereotype.

One of the most compelling studies involved a series of three experiments in which subjects were exposed to non-self-relevant stereotypes and then measured on stereotypic behaviors (Bargh et al., 1996). Embedded indirect primes were used in the first two experiments through the Scrambled Sentence Task, in which subjects were presented with 5-word clusters which could be unscrambled to form sentences as an ostensible part of a language task. Sentences either included trait words related to the stereotype of interest or included neutral words. In Experiment 2a, negative attributes associated with older people were included in the scrambled sentences for subjects in the experimental group, and the time it took participants to walk from a doorway to a desk served as a dependent measure. The researchers hypothesized that subjects for whom the idea of elderly was activated would (in response to ideomotor theory) walk significantly more slowly than those in the control condition. Analysis confirmed their expectations: subjects who were primed with elderly traits walked significantly more slowly than those in the control condition. Further, an additional experiment (Experiment 2b) confirmed that the differences could not be attributed to mood and must therefore be the result of the activation of a cognitive schema.

In a study important to this paper's hypothesis, Dijksterhous and others (1998) replicated and expanded on the findings of Bargh and others (1996) using stereotypes and exemplars of professors (intelligent) and supermodels (not intelligent). Exemplars are well-known individuals that represent the social category, such as "Albert Einstein" as an exemplar of the category "professor." Previous tests had indicated that the intelligence stereotypes are commonly known. Categories were activated by asking subjects to write for five minutes about either one of the two categories or an example of one of the two categories. They found that performance on a 30-item trivia task assimilated to the stereotype in the category activation condition: subjects who had described professors performed significantly better than those who had described supermodels. This is consistent with the findings of Bargh and others (1999). However, contrast effects were found for the exemplar activation: subjects who described Einstein performed significantly more poorly than those who described a well-known supermodel Claudia

Schiffer. They verified this effect using different prime mechanisms and measuring walking speed, as in the experiment of Bargh and others (1999). This experiment supports the notion that category activation yields assimilation effects while individual comparisons with extreme others yield contrast effects.

Another study that appeared to yield assimilation in non-self-relevant stereotype activation conditions was conducted by Wheeler, Jarvis, and Petty (2001). An embedded stereotype prime was used on White subjects, who were asked to write a paragraph about either “A Day in the Life of Tyrone” or “A Day in the Life of Erik” for the ostensible purpose of exploring dominant/non-dominant handedness. Students assigned to write about Tyrone (which was expected to activate a stereotype of a young Black man) performed significantly more poorly on sample GRE items than did those who wrote about Erik (which was expected to activate no stereotype or a White stereotype). Although these findings are compelling, one might reasonably construe that either an actor effect occurred in which assuming the role of a stereotyped individual caused a performance deficit, or that an ideomotor effect occurred in which the behavior shifted in response to cognitive activation.

Finally, in an exploration of subliminal and supraliminal prime conditions, Shih and others (2002, Study 2) compared test performance of subjects who were differently exposed to Asian words. Both prime conditions and the control condition used a computerized vigilance task, in which Asian words (e.g., chopsticks) or non-Asian words (e.g., water) were flashed on the screen in different time increments under the guise of a speed test for identifying location of the flash. In the subliminal condition, words were flashed for 80 milliseconds—long enough to detect a flash but not long enough to detect a word. In the supraliminal condition, the same words were flashed for 1000 milliseconds—long enough to both detect a flash and read the word. A manipulation check verified that the prime did indeed activate the Asian category. Results showed that Asians in the subliminal condition outperformed those in the control condition, but Asians in the supraliminal condition performed worse than controls. Non-Asians in the supraliminal prime condition outperformed controls, but no differences were found in the subliminal condition. These results illustrate the differences between category and identity activation, in that conscious activation yielded assimilation for nonstereotyped

subjects (category activation) but contrast for stereotype subjects (identity activation). Conscious identity activation for Asian subjects may have activated individual comparisons with a high-achieving in-group, yielding a contrast effect, but unconscious activation might have activated the category without the accompanying intragroup comparison.

#### *2.1.2.2 Distinguishing Priming Effects From Stereotype Threat*

Empirical evidence exists that distinguishes priming effects from stereotype threat (Marx & Stapel, 2006a, 2006b). In one particularly relevant study, Marx and Stapel (2006a, Study 1) randomly assigned 50 male and 60 female students to a prime/no prime condition and to a diagnostic/not diagnostic condition before giving them a test containing items similar to those on the Graduate Records Exam (GRE) and a scale of threat based-concerns. The prime was given through the twenty scrambled sentence items, for which twelve of the twenty phrases in the primed group suggested “dumb,” while the non-primed group unscrambled only filler sentences. The diagnostic/non-diagnostic condition were similar to those used by Steele and Aronson (1995), suggesting in the title that the test was either a diagnostic aptitude test or a reasoning exercise. They hypothesized that the unprimed condition would yield typical stereotype threat effects, such that women in the diagnostic condition would underperform and report higher levels of threat-based concerns than women in the non-diagnostic condition and men in either condition. They further predicted that men and women in the primed condition would underperform relative to their unprimed peers, but that threat based concerns would only be present for the women in the diagnostic-primed group. Results confirmed their hypotheses: analysis of variance found main effects for gender ( $F(1, 102) = 28.67, p = .05, \eta = .22$ ) and stereotype threat ( $F(1, 102) = 21.86, p = .05, \eta = .18$ ), and interactions within the expected pattern. This experiment directly illustrates the difference between priming effects and stereotype threat.

The priming effects eliciting shifts in behavior and performance for nonstereotyped subjects are not the main focus of this study, and may in fact add unnecessary complication to the idea of comparison effects on multiple levels. It is important to note the different processes that might have prompted the assimilation effects, however. Wheeler & Petty (2001) attributed the assimilation effects for

nonstereotyped subjects to either ‘hot’ processes “emphasizing feelings and motivational states” (p. 803) or ‘cold’ ideomotor processes involving cognitive activation without conscious emotional underpinnings. This distinction can account for the observation that both self-stereotypes and other-stereotypes led to assimilation effects. Primes that activate categories irrelevant to the self can lead to unconscious behavioral assimilation as a consequence of cognitive activation, while primes that activate a stereotype about oneself might also lead to assimilation due to concurrent cognitive and emotional activation. The importance of this distinction lies in its implications: a cognitive category can be de-activated when one becomes aware of it, but activation of a stereotyped identity activates emotional responses that are less easily controlled by the subject. One effect may be stopped by the performer, while the other operates only marginally within of individual control.

### *2.1.3 Moderators of Stereotype Threat*

In order for the stereotype to evince power over an individual, several situational conditions must be present. First, the domain in which the test is given must be relevant to the individual (e.g., Steele & Aronson, 1995). A Black person who is disengaged with school will not be threatened by a negative stereotype about the intellectual ability of his or her racial group. Second, the individual must be conscious of the stereotype (e.g., R. P. Brown & Pinel, 2003). One who is unaware that Asians are perceived to be good at math cannot benefit from the stereotype. Third, the individual must identify strongly with the stereotyped group. A woman who is strongly identified with her female in-group will suffer more in stereotype-threatening situations than one who does not identify strongly with her in-group (e.g., Schmader, 2002). Finally, one’s collective identity must be activated at either in order for the stereotype to influence performance (e.g., Ambady, Paik, Steele, Owen-Smith, & Mitchell, 2004). A woman among other women will not feel the threat of stereotyped expectations unless her gender identity is made salient. Evidence supporting these moderators is provided in the following paragraphs.

#### *2.1.3.1 Domain identification*

Domain holds a pivotal role in stereotype threat. Steele (1997) points out that “to sustain school success one must be identified with school achievement in the sense of its being a part of one’s self-definition” (p. 613), adding that “for those who identify with

the domain to which the stereotype is relevant” (Steele, p. 614) stereotype threat can undermine performance. Aronson and Lustina (1999) elaborate on this by noting that “to be threatened by the self-evaluative implications of a stereotype that alleges low ability of some kind, a person probably needs to either care about having the ability or at least care about the social consequences of being seen as lacking the ability” (p. 35).

Several studies have illustrated this effect. For example, in a study by Aronson and Inzlicht, (1999, Study 2), male undergraduates were divided into groups based on the extent to which they were personally identified with the domain of mathematics. They then took a test which was described as either an exploration of the increasing math superiority of Asians or simply as an exploration of mathematical processes. An interaction occurred: among subjects who identified with the mathematics domain, those who were reminded of Asians’ math superiority performed more poorly than those who did not have the Asian stereotype activated; the opposite pattern occurred for those who identified only moderately with the math domain.

Stone and others (1999) also explored the importance of domain in eliciting stereotype threat experiences. In a second study in their exploration of athletic stereotype threat, they compared the threat experience for White students who were highly identified with the athletic domain (athletically engaged) with that of White students who were only marginally identified with the athletic domain (athletically disengaged). They found that performance declined for those who were highly engaged in athleticism under threat conditions (test framed as a measure of natural ability). When athletic disengagement was experienced, subjects experienced no significant difference in high- and low-threat conditions. In fact, disengaged participants outperformed engaged participants when the threat was imposed.

Domain identification is so critical to stereotype threat responses that researchers often select subjects based on their established identification with the stereotyped domain (e.g., Shih et al., 1999). When exploring the stereotype threat phenomenon for Black people, for instance, Steele and Aronson (1995) selected as experimental subjects Black students whose grades and SAT scores funneled them to a premiere university. Presumably, Blacks who were not identified with the domain would not experience a similar response to the threat situation. Studying the stereotype of women’s inferior math

skills, only women whose preferences and abilities led them to major in math in college were selected for participation (e.g., Shih et al., 1999). In both cases, reduced performance for the stereotyped groups was generated.

#### *2.1.3.2 Stereotype Awareness and Stigma Consciousness*

Empirical studies have illustrated the relevance of stereotype awareness in stereotype threat studies. One such study compared math performance for Asian students from Canada and from the United States (Shih et al., 1999). Canadian students were less aware of stereotypes (or the stereotypes were less prevalent), and the consequential impact on their performance was minimized. The lack of significance found in the study of stereotypes about verbal ability of women and Asians (Inzlicht & Ben-zeev, 2000) might have been created, in part, by the lower prevalence of the stereotypes themselves. Without verification that the stereotypes were known to the group, one could not be threatened by it.

Stigma consciousness is a somewhat more extreme form of stereotype awareness, according to Brown and Pinel (2003). Individuals who are high in stigma consciousness have a tendency to attribute life events to the stigma associated with a stereotype, perceiving “more discrimination directed at them personally and ... more concrete examples of instances in which they have been treated in a prejudicial manner” (p. 194). Brown and Pinel hypothesized that individuals high in stigma consciousness would have a greater risk of performance decrement in stereotype threatening situations. In their experiment, women that who identified with the math domain and who were either high or low in stigma consciousness were placed into a high-threat condition (told that the test explored gender math differences) or a low-threat condition (told that a math test was gender neutral). Results supported their hypothesis: subjects high in stigma consciousness performed more poorly than those low in stigma consciousness. Thus, both direct and indirect evidence supports the idea that stigma consciousness moderates stereotype threat.

Finally, a study by Aronson & Inzlicht (2004) explored stereotype vulnerability as a factor moderating the response of Black students to stereotype threat situations. Stereotype vulnerability was defined as “the tendency to expect, perceive, and be influenced by negative stereotypes about one’s social category” (pp. 829-30). Using the RS-Race scale, which measured race-based sensitivity, they placed subjects into high- or

low-stereotype vulnerability groups before initiating a stereotype threatening situation. Subjects who were high in stereotype vulnerability tended to have a lower performance on a difficult academic test. That is, the greater one's susceptibility to race-based rejection sensitivity, the poorer one's performance in stereotyped domains.

### *2.1.3.3 Identification with Stereotyped Group*

Individuals who have internalized their stereotyped identity appear to be more susceptible to stereotype threat (e.g., Davis, Aronson, & Salinas, 2006; Schmader, 2002). In a study of gender-based stereotype threat, Schmader (2002) found a direct relationship between the extent to which women identified with their female identity and their susceptibility to stereotype threat. Male and female subjects took a difficult mathematics test after gender was either activated (subjects told male/female math scores would be compared) or not activated. By analyzing scores in conjunction with a gender identification measure taken earlier in the semester, it was discovered that women who were highly identified with their gender experienced a notable performance decrement when their gender was made salient before the test was taken. A reverse effect was found for men, such that "higher levels of gender identification...predicted better performance among men when gender was relevant to the test" (p. 198). Thus, gender identification can be viewed as a moderator of performance in gender-relevant situations.

The relationship between group identification and stereotype threat vulnerability is not necessarily linear. Just as high identification with the stereotyped group might diminish interest and engagement with the domain, so might high identification with the domain diminish identification with the stereotyped group. Women who wish to advance in an academic realm might distance themselves from typical feminine presentation (e.g., makeup, high heels) and interests (e.g., shopping) as a means of eluding the negative aspects of feminine stereotypes. The difficulty of maintaining one's persistence in a stereotyped domain is illustrated in the following excerpt:

For the advanced female math student who has been brilliant up to that point, any frustration she has at the frontier of her skills could confirm the gender-based limitation alleged in the stereotype, making this frontier, because she is so invested in it, a more threatening place than it is for the nonstereotyped. Thus, the work of dispelling stereotype threat through performance probably increases with the difficulty of work in the domain,

and whenever exemption is gained has to be won at the next new proving ground” (Steele, 1997, p. 618).

Steele and Aronson (1995) explored the dynamic nature of the disidentification process for Black students after discovering that Black students in a threat condition declined to identify their race more often than those in a non-threat condition. A subsequent experiment confirmed that the differences were significant: Black subjects rated their enjoyment of stereotypic activities (e.g., rap music, basketball) lower than non-threatened Black subjects ( $t(61) = 3.61, p < .001$ ), and only 25% of threatened Black subjects were willing to indicate their race on a demographic form compared with 100% of non-threatened Black subjects and of White subjects in either condition. Just as disidentification with the stereotyped domain can protect an individual’s self-esteem, it appears, so can disidentification with the stereotyped group protect self-esteem if the domain must be pursued.

The same relationship was found for women. Pronin, Steele, and Ross (2004, Study 1) explored relevant and irrelevant feminine characteristics for women with different levels of math identification. Relevant feminine characteristics included those that were applicable to the stereotype, such as “flirtatious” and “desiring to have children” (p. 159), while irrelevant feminine characteristics included such traits as sensitivity and nurturance. Traits were grouped according to earlier survey results. They found that female students who had taken more math classes (and were presumably more exposed to stereotype threat) had lower identification with relevant feminine characteristics than those who had taken fewer math classes. A second study explored this effect experimentally, finding that women who had been exposed to a stereotype-threatening article identified less with relevant feminine characteristics than those who had not been so exposed. Once again, irrelevant feminine characteristics were not disavowed, supporting the authors’ proposition that “identity bifurcation” occurs in stereotype threatening situations.

#### *2.1.3.4 Collective or Individual Identity Activation*

Stereotype threat necessitates that a particular collective identity is made salient, and a number of studies have clarified this relationship by distinguishing between

performance results under individual and collective identity activation. One such study sought to explore this phenomenon by exposing subjects to either female or neutral words during a Vigilance Task, then placing them into either an individuation or non-individuation condition (Ambady, Paik, Steele, Owen-Smith, & Mitchell, 2004). In the individuation condition, they wrote about a number of personal interests before listing positive or negative traits, while in the non-individuation condition they responded to neutral questions (Study 1). In Study 2, the individuation condition asked only for negative traits so as to avoid the possibility of self-affirmation effects. In both studies, significant interactions were found. Women whose female identity was primed performed more poorly than controls, but only in the non-individuated condition. If their individual identity was activated after the gender prime, their performance was no different from those who had not been primed. While some questions might arise as to the importance of the sequence of the primes (in both cases, the gender prime preceded the individuation prime), we can tentatively presume that activation of a group identity can be outweighed by individuation.

Another study that blends in-group social comparison with stereotype threat can be found in the work of Blanton, Crocker, and Miller (2000). In the study, sixty African American students who had just taken a bogus IQ test indicated their self-esteem related to their performance after exposure to either a high-performing or low-performing confederate who was either a member of their racial in-group (Black) or out-group (White). Verbal information given after the test indicated that it (a) explored natural math ability, (b) had previously only been standardized by White students, and (c) required that subjects indicate their race on a demographic form. Each of these should heighten race awareness and create a potentially threatening experience, and all subjects were exposed to the threat. Subjects then ‘overheard’ interactions between a Black or White female confederate and the experimenter. In one condition, the “overheard” interaction suggested that the confederate had performed very poorly on the test, and in the other condition the experimenter became excited when the confederate ostensibly performed in the 99<sup>th</sup> percentile. It was expected that performance self-esteem would vary in accordance with the in-group or out-group status of the successful and unsuccessful confederate.

Results confirmed expectations. Subjects who were exposed to the superior-performing African American target revealed higher performance self-esteem than those who had been exposed to a poorly-performing African American confederate. When the confederate was White, the opposite was true: subjects revealed lower self-esteem in comparison to a superior-performing target. State self-esteem (which had also been measured through reports of current affect) did not differ significantly. This suggests that upward comparison with an in-group member has an assimilative effect when a negative stereotype about that in-group has been activated. Conversely, upward comparison with an out-group member under negative stereotype activation condition yields a contrasting effect on performance self-esteem. The absence of performance measure as a dependent variable and lack of a control group with which to compare present some serious limitations to this study in terms of its applicability to this study.

#### *2.1.4 Summary of Stereotype Activation*

Thus it is shown that (a) negative self-stereotype activation elicits assimilation effects for members of stereotyped groups, including Black people, women, low-SES people, and student athletes, (b) positive self-stereotypes yield assimilation for students from elite universities and Asian students, (c) implicit self-stereotypes are activated when out-group stereotypes are made salient and lead to performance shifts for non-stereotyped group members, (d) activation of different identities within the same person can yield assimilation to the most salient identity, (e) the effects of different primes can explain contradictory findings, and (f) cognitive category activation can generate assimilation effects for nonstereotyped subjects. Moderators of stereotype threat have included domain identification, stigma consciousness, identification with stereotyped group, and collective identity activation. Findings in social comparison will provide the next basis for exploration of parallel processes.

## 2.2 Social Comparison: Empirical Findings

Research related to social comparison will be discussed in the following sections: (a) general overview, (b) factors influencing response to social comparison, and (c) impact of social comparison on performance.

### 2.2.1 General Overview

Since its crystallization as an area of study by Leon Festinger in 1954, social comparison theory has undergone extensive exploration, tweaking, formalization, measurement, and development. In its original conceptualization, social comparison theory was framed as a mechanism through which people in ambiguous situations “evaluate their opinions and abilities by comparison respectively with the opinions and abilities of others” (Festinger, p. 118). Framed by Festinger as a deliberate, conscious process, numerous studies explored situational and personality differences in the decisions of when and to whom individuals compare.

For the purposes of this paper, however, attention will not be placed on the conscious comparison choices described by Festinger (1954) but will instead concentrate on the consequences of forced or automatic--or *nondeliberate*--comparisons. Such nondeliberate comparison information presents itself unsolicited continually in an individual’s life—more frequently, perhaps, than the deliberate comparisons described by Festinger (1954). A more attractive person enters the bus, the neighbor gets a new car, a friend’s child is accepted into Harvard, a colleague receives an award--such daily events are believed to be the fodder for ubiquitous social comparisons that frame one’s affect, self-esteem, behavior, and performance.

Wood (1989) explained the shift in research emphasis from deliberate to automatic comparisons by observing that “the social environment provides comparisons that impinge on the individual, whether or not he or she has “selected” them” (p. 233). Stapel and Blanton (2004) emphasized it even more strongly, indicating that self-knowledge derives from social comparisons, and that such comparisons must be automated in order to be useful. “After all,” they pointed out, “social comparison can be of little benefit to individuals if it is too much of a burden for their limited mental resources” and that “incidental exposure to a target person could be sufficient to influence self-evaluation” (Stapel & Blanton, p. 468). Incidental exposures and their consequences are of greatest interest in the current thesis, as they are most analogous to the incidental and non-deliberate exposure to stereotype threat.

Studies of social comparison tend to be organized fairly systematically. Subjects are generally randomly assigned to a condition in which they are exposed to an upward

comparison, a downward comparison, or a neutral stimulus. Comparison primes are often quite overt, such as asking subjects to read and reflect on an article about a highly successful target (e.g., Lockwood & Kunda, 1997) or to read such an article under the guise of assessing newspaper style (e.g., Stapel & Koomen, 2000), but are occasionally more subtle, as in the Vigilance Task activation used by Dijksterhuis and others (1998). Following exposure, some sort of dependent measure is taken. Social comparison studies have tended to lend considerable emphasis to self-evaluation, stemming most likely from Festinger's (1954) original theory emphasizing the role of social comparison in enhancing self-knowledge, and measurements have often relied on validated or non-validated self-evaluation reports. Nevertheless, some studies have incorporated ambiguous performance measures into their designs either in addition to self-evaluation. An example of such a measure is the Remote Associates Test (RAT), in which subjects needed to find a common feature among series of three distinct words, such as "car, elephant, attic (trunk)." It is useful in experiments because it can be cast as a test of logic, creativity, verbal skill or in any other light deemed necessary to the experimenter. (Johnson & Stapel, 2007a; Johnson & Stapel, 2007b).

In this section, the term "assimilation" will again refer to a shift in behavior in the direction of the comparison target, and "contrast" will describe the opposite effect. The phrase "upward comparison" will refer to a comparison with another person whose ability is presumed to be superior to one's own, and "downward comparison" denotes a comparison with a presumably inferior other. "In-group comparisons" will refer to situations in which the comparison target is a member of a group in which the subject views himself as a member, and "out-group comparison" pertains to targets not belonging to groups in which the subject is a member. "Comparison target" or "target" refers to an individual with whom a subject has made a comparison, and "comparer" is the individual making the comparison.

### *2.2.2 Factors Influencing Response to Social Comparison*

Festinger's (1954) original conceptualization of social comparison suggested that individuals would choose comparison targets whose performance was slightly superior to their own, then adjust their own performance upward in order to make themselves slightly superior to their competitors—a "unidirectional drive upward" (p. 124). In other

words, he suggested that upward comparisons will produce performance increases (assimilation). Viewing the question in terms of self-evaluation rather than performance, however, other studies have found that upward comparisons tend to produce declines while downward comparisons produce boosts to self-evaluation (contrasting effects). Still others suggest that both contrasting and assimilative effects are produced, in an upward comparison might produce an initial decline in self-esteem or affect followed by increased effort to boost one's performance. Factors that influence these outcomes will be addressed in the following paragraphs, organized in terms of (a) the comparison target, (b) the comparer, and (c) the situation.

#### *2.2.2.1. The Comparison Target*

In order to be considered a viable target for comparison, an individual must have personal characteristics and abilities similar to one's own. Personal characteristics that resemble one's own, including ascribed characteristics like age or gender and achieved characteristics like proximity, promote the greatest likelihood of comparisons (Major, Testa, & Blysm, 1991). Broadening the definition of similarity, Major and colleagues further explain that comparisons with "individuals who share distinctive attributes with oneself" (p. 243) in a variety of domains produce the most pronounced effects on the comparer. Individuals with these characteristics might be considered to be in-group members whose ascribed status renders them similar and therefore relevant to the comparer (Tesser, 1991).

Targets must also have ability levels not too markedly different from those of the comparer, or abilities not too different from the norm. Comparison with extreme targets—targets who are markedly different from normative standards—tends to yield contrast effects. For example, Stapel and Blanton (2004) measured subjects' attribute self-ratings after subliminally exposing them to extreme comparison standards using the Vigilance Task. In each instance, contrast effects were revealed such that subjects' self-ratings of their own youthfulness was lower after exposure to a very old than to a very young target (Study 2a), attractiveness ratings were lower after comparing with a very attractive than a very unattractive target (Study 2b), and intelligence ratings were lower after comparing with Albert Einstein than with a clown (Study 2c). To ensure that these effects were not simply a consequence of anchored rating scales or of self-awareness,

they conducted a final study in which self-ratings were preceded with an implicit self-esteem measure based on signature size (Study 4). Results of the implicit self-esteem measure paralleled those taken with the explicit measures and suggested that the findings were not simply an anchoring artifact.

Other studies have found similar effects in measures of self-reported aggressiveness and athleticism after comparison with extreme and moderate standards (Mussweiler, Ruter, & Epstude, 2004) and using exemplars rather than categories (Dijksterhuis et al., 1998). In the study by Mussweiler and others (2004a), for example, exposure to moderate standards led to assimilation and exposure to extreme standards led to contrast in self-evaluation. In Study 1, subjects who were subliminally exposed to either a moderately high (a well-known TV detective) or moderately low (a well-known pop star) aggression standards tended to assimilate self-evaluations of their own aggressiveness toward the direction of the standard. In Study 2, subjects' estimates of their own athleticism (e.g., number of sit-ups) after comparison with an extreme or moderate upward standard (Michael Jordan/race car driver Nicki Lauda) or with an extreme or moderate downward standard (Bill Clinton or Pope John Paul II) tended to evaluate their own performance in the opposite direction of extreme standards (contrast) and in the same direction as moderate standards (assimilation), providing the highest self-evaluations of athleticism after exposure to Pope John Paul and lowest self-evaluations after exposure to Michael Jordan. Finally, professional water polo players estimated their own water polo ability (speed with which they could swim various distances) after comparison with a moderate or extreme upward comparison (Study 3). Even though the water polo players should have good knowledge of their swim speeds and receive regular feedback regarding their performance, they estimated their own speeds as higher when comparing with a moderate than with an extreme upward standard.

In some cases, targets with extreme abilities might serve as role models whose achievement inspires and motivates the comparer. Role models might be described as in-group members whose age or status renders them non-threatening to subjects. To serve as a role model rather than comparison target, subjects must believe that the accomplishments of the target are at least feasibly attainable by the subject. A professional football player might serve as a role model for an aspiring high school

quarterback, or a successful professional engineer might serve as a role model for engineering students. Leon Festinger (1954) points out that if another's performance is perceived to be markedly different from one's own, as in the situation where an amateur golfer encounters a professional golfer, subjects exposed to the performance difference "ceased comparison with him" (p. 133). Thus, targets who serve as role models would not elicit comparative processes with subjects and might, therefore, generate assimilation effects by way of improved performance in the relevant domain.

Illustrating this effect, Lockwood and Kunda (1997) conducted experiments to explore the inspirational impact of role models. College undergraduates who aspired to be either teachers or accountants were exposed to a description of either a highly successful teacher or accountant. In both cases, those who were exposed to the domain with which subjects identified reported higher self-ratings on career-relevant attributes (e.g., bright, capable) than did those who were exposed to an irrelevant target or to controls. The target served as a role model because their achievements were attainable and the domain important to the targets, who were themselves in college in the same area.

In another experiment within the same study, Lockwood and Kunda (1997) explored self-ratings of first- and fourth-year college undergraduates after exposure to a target described as a highly successful fourth-year student who had achieved highly in academic, social, and leadership areas. The relationship of the subjects to the target varied according to the years of schooling of the subjects. Subjects with one year of schooling were thought to view the target as a role model who was not a part of their peer group—an "attainable" target, as students might reasonably be equally successful in their own first year of college. Those with four years of schooling, however, might view the other fourth-year student as a member of their peer group—an "unattainable" target whose success could no longer be achieved by the subjects. Different results were expected according to this relationship between subject and target. Results confirmed a difference: self-evaluations for first-year students were not significantly different from those of control subjects, but fourth-year students provided significantly lower self-ratings than first-year students or fourth-year controls. Thus, a contrast effect occurred when the target was a member of one's peer group and thus posed a threat, whereas an

assimilation effect occurred when the target was not a member of the peer group and thus posed no threat.

Confirming Festinger's (1954) original suggestion that people tend to compare with slightly superior in-group members, the studies described in the previous paragraphs indicate that the manner in which one responds to these comparisons varies according to whether the target meets the criteria for similarity.

#### *4.2.2.2 Characteristics of the Comparer*

In addition to the target, comparers will be more greatly affected by comparisons if they possess particular mental states during the experience. For example, individuals who are high in self-uncertainty, low in self-esteem, or low in optimism have a greater susceptibility to the effects of comparisons than do people with the opposite states. These characteristics often represent states of being rather than enduring personality traits, suggesting interactions between the comparer and the general situation in which the comparison is made.

Festinger (1954) originally proposed uncertainty as a primary motivation for social comparisons. When one is uncertain, he suggested, one seeks to compare with others to gauge whether one's opinion or performance is adequate. When considering the consequences of such comparisons, self-uncertainty is similarly influential. This has been tested experimentally. Stapel and Blanton (2004), for example, created a state of uncertainty in some subjects by providing them with fictitious test results indicating that their personality profile was ambiguous. Subjects in the uncertain state experienced a decline in implicit and explicit self esteem following exposure to an upward comparison (Albert Einstein). These results did not occur for subjects in a high-certainty state or for subjects in a low-certainty state who compared themselves with a clown.

Individuals with low self-esteem (LSE) tend to have a more pronounced contrast response to comparisons than do individuals with high self-esteem (HSE) (Major, Testa, & Blysm, 1991; Wills, 1991). For example, Gibbons and McCoy (1991) found that LSE students who had received failure information about themselves (high-threat) derived a greater mood boost from downward comparison than did HSE students in the same situation. Crocker and Gallo (1985; in Major et al.) found similar results: LSE subjects who had failed experienced a more positive mood after downward comparison (contrast),

but HSE subjects experienced a boost after upward comparison (assimilation). Finally, Gibbons and Gerrard (1991) reported that LSE subjects who were having trouble adjusting to college experienced a mood boost after hearing about another student whose adjustment had been more difficult but no mood boost in hearing about a well-adjusted student (downward contrast). HSE subjects, on the other hand, experienced a mood boost when reading about the well-adjusted students but not about the struggling student (upward assimilation). Thus, self-esteem may determine both the target with whom one compares and the impact of this comparison on mood and other affective states.

Optimism also appears to play a role in one's responses to social comparison. One study found that subjects low in optimism were more likely to shift comparison targets downward following a decline in grades, and that they experienced a greater increase in grades after shifting targets than did high optimism subjects (Gibbons, Blanton, Gerrard, Buunk, & Eggleston, 2000). Another study found similar results for subjects who were depressed: depressed subjects reporting a mood boost from downward comparison that did not occur for non-depressed subjects (Gibbons, 1986).

Thus, subjects who have greater uncertainty, lower self-esteem, and lower optimism/greater depression were more likely to produce contrasting effects to downward comparison, whereas subjects with the opposite characteristics were more likely to experience assimilative effects from upward comparisons.

#### *4.2.2.3 Characteristics of the Situation*

While characteristics of the comparison target and the comparer influence the consequences of the comparison, the situation itself may interact with and influence these factors. Major, Testa, and Bylsma (1991) provide evidence for a comprehensive framework in which the relevance of the comparison to the individual's self esteem (esteem relevance) and the extent to which the comparer believes they can control his or her status (perceived control) influence his or her response to the comparison. Additionally, Brewer and Gardner (1996) observed that the extent to which one is oriented toward his or her individual or group identity (self-construal) will also influence responses to comparisons. While these may seem to be characteristics of the comparer and thus more appropriate in the previous section, they are each strongly influenced by

situations in which the comparer is found and thus more appropriately placed in a separate section. Each of these factors will be discussed in the following paragraphs.

*4.2.2.3.1 Esteem relevance and perceived control.* Major and colleagues (1991) proposed that a comparison dimension must be relevant to one's esteem and out of one's control in order to impact the comparer. Esteem relevance, they contend, consists of (a) the similarity between the comparer and target and (b) the importance of the dimension to the comparer. Perceived control consists of (a) the extent to which the comparer believes comparison dimension is alterable, and (b) the comparer's expectancy to be able to alter his or her status (p. 239). These two factors interact to provoke a response to a comparison situation.

Esteem relevance (or domain identification) is one of the foundations on which social comparison stands, however. Festinger (1954) observed that "if an opinion or ability is of no importance to a person there will be no drive to evaluate that opinion or ability" (p. 130), and evidence exists to suggest that one is not threatened by superior targets in domains not valued by subjects. A doctoral student in psychology, for instance, would not feel threatened when encountering a professional actor unless her own acting aspirations had been forestalled.

Evidence exists in several studies. In one series (Johnson & Stapel, 2007b), the domain of target and test were matched or mismatched to the test on which the subjects were to perform, but this was not necessarily a domain of personal value to the subjects. The researchers suggested that the presence of a performance boost but absence of a performance decrement most likely reflected the fact that the domain was not important to the individuals. The alternate domain provided a chance to boost self-regard, but subjects were not sufficiently threatened by the target to experience the anxiety that might produce a subsequent decline in performance.

To further demonstrate the importance of esteem relevance, Major and colleagues (1991) cite the work of a range of researchers. Tesser, Millar, and Moore (1988), for example, found that individuals tend to experience boost in self-evaluation if a close other performs well on a task that is irrelevant to the comparer, but a decline in self-evaluation if the superior performance represented a domain of importance to the esteem of the comparer. The consequences did not end with self-evaluation, however; subjects

who compared downward in esteem-relevant domains tended to either (a) distance themselves from the target by rating them as dissimilar or by physically sitting farther away from the target (Pleban & Tesser, 1981), (b) distance themselves from the domain by rating the domain as “less self-relevant than when they had performed better than their partner” (Major et al., 1991, p. 250), or (c) experience more negative affect and anxiety than subjects for whom the domain was not relevant (Salovey & Rodin, 1984). Each of these cases presents a contrasting effect, such that exposure to an inferior or superior other produces the opposite effect in the individual’s self-esteem.

In another study, subjects were led to believe that they were similar to a co-worker or not similar to a coworker, and then watched the same coworker fail an anagram task. Those who believed they were similar to the failing coworker tended to have reduced expectations for their own success on a similar task (assimilation). Subjects than did subjects who did not believe they were similar to the coworker did not experience these effects, perhaps because they did not make the comparison (I. Brown & Inouye, 1978).

Even when the domain and ability are similar between comparer and target, however, some comparisons produce assimilative effects and others produce contrasting effects. These differences seem to relate to the extent to which the dimension is changeable and within control of the comparer. To be specific, the determinant is not the actual extent to which the situation is changeable and controllable, but the comparer’s perception of these factors. These perceptions relate, according to Major and colleagues (1991), both to outcome expectancies (Can the dimension be changed?) and efficacy expectancies (Do I have what it takes to change my status?). When the dimension is changeable and controllable by the comparer, the consequences of comparisons will differ.

For example, cancer patients who believed they had control over their illness responded positively to an image of a successful cancer survivor but did not respond negatively to a cancer patient with a less favorable outcome. Cancer patients who felt they had little control over their outcome did not experience a similar boost following the upward comparison (Buunk, Collins, Taylor, & VanYperen, 1990). Likewise, people in smoking cessation programs, whose efforts to seek help suggests that they feel at least a

modicum of control over their fate, responded positively to a description of a person who had overcome their smoking addiction (Meichenbaum, 1971). Thus, the extent to which one has control influences whether one will benefit from or be harmed by upward comparisons.

The effect of downward comparison is influenced by similar factors. Downward comparisons of relevant dimensions to another person who is similar to the comparer will produce a boost to self-esteem and effort if the individual feels they control their own fate (contrast). If the dimension is not changeable or if the comparer does not feel as if s/he has control over the dimension, then the downward comparison will produce decline in affect, stress, reduced effort and/or disidentification from the dimension (assimilation). Such downward comparisons seem to highlight a worst case scenario that then becomes possible for the comparer. To summarize, perceived control is likely to produce favorable outcomes from both upward and downward comparisons, while absence of control produces unfavorable outcomes for the comparer regardless of the direction of the comparison.

To illustrate this model, Major and colleagues (1991) developed a framework in which the outcomes of comparisons rely on a series of components. They first proposed that an upward comparison will generate different responses based on (a) similarity of target and comparer, (b) relevance of the comparison dimension to the comparer, (c) changeability of the dimension, and (d) extent to which individual can control the dimension. If the comparer is similar to the target and the dimension is self-relevant, unchangeable, or uncontrollable, then a threat is produced by an upward comparison which will result in reduced self-esteem and effort on the part of the comparer (contrast). If the dimension is changeable and the comparer has control over the change, however, the same comparison may produce a boost in self-efficacy and effort (assimilation). A visual depiction of this model can be seen in the chapter by Major and colleagues (1991), 241-2.

*4.2.2.3.1 Self-Construal.* The identity that perceives and responds to stimuli at a given time is referred to as one's self-construal, and bears a powerful influence on behavior. The primary construal orientation that exists in research involves the "distinction between the individuated or personal self (those aspects of the self-concept

that differentiate the self from all others) and a relational or social self (those aspects of the self-concept that reflect assimilation to others or significant social groups; Brewer & Gardner, p. 83). Although individuals and groups may have different tendencies toward group or individual orientations, different self-construals “coexist within the same individual, available to be activated at different times or in different contexts” (Brewer & Gardner, p. 83).

Activation of different selves has ramifications for perceptions and behavior. When the collective self is activated, for instance, “the content of the self-concept is focused on those characteristics that make one a “good” representative of the group .... [such that] the most salient features of the self-concept become those that are shared with other members of the in-group” (Brewer & Gardner, 1996, p. 84). The in-group, therefore, plays a critical role in one’s self-concept, providing both a frame of reference for our own individual performance (e.g., “How am I doing in comparison to my peers/colleagues/race/gender?”) and social identity which relies on the relative status of our group compared to others (“How is my in-group doing compared to our out-group?”) In fact, Brewer and Gardner suggest “the collective self-concept is determined by assimilation to the prototypic representation of the in-group, with self-worth derived from the status of the in-group in intergroup comparisons” (p. 85).

The shift from personal to social identity can be sufficiently powerful to result in self-stereotyping (Simon & Hamilton, 1994), in which individuals for whom a particular identity is activated see themselves as more representational of their group. This phenomenon was well-illustrated in an experiment by Hogg and Turner (1987), in which female subjects who were placed in a discussion with one woman and two men (making salient their gender identity) subsequently reported themselves as more typical of their sex than women who had engaged in a similar discussion task with only another woman. Even a contrived group situation in which successful individual was framed as either an in-group or out-group member yielded a significant difference in the subjects’ subsequent agreement with ambiguous statements (Brewer & Gardner, 1991).

The power of self-construal has been explored in many social comparison studies. In addition to the previously mentioned work by Brewer and Gardner (1991), other studies have examined the impact of different types of self-activation on self-evaluation

and performance following social comparison (Marx, Stapel, & Muller, 2005; Schwinghammer, Stapel, & Blanton, 2006), and on the group membership of a successful other's group membership on self-evaluation (Gordijn & Stapel, 2006).

When the self is activated in a neutral manner, individuals will have greater interest in social comparison information (Stapel & Tesser, 2001) and a greater likelihood to engage in defensive processing (Stapel & Koomen, 2001). Defensive processing is a protective process through which subjects in threat situations (e.g., upward comparison in relevant domain) protect their self-concept by denigrating the comparison target (e.g., rating their attractiveness less favorably than controls) and raising their own self-estimations. Schwinghammer, Stapel, and Blanton (2006) extended this research by exploring the impact of negative, neutral, and positive self-activations. Using blatant priming tools (e.g., list and give examples of three negative traits), they activated positive or negative self-construals before asking subjects to form an impression of a very attractive other. In agreement with their findings, those whose positive selves were activated did not engage in defensive processing and thereby experienced the greatest reductions in self-evaluation following upward comparison. Negative self-activation, on the other hand, resulted in the highest self-ratings and lowest other-ratings, while neutral activation fell between both.

An equally compelling example of the power of self-construal was depicted in an interesting experiment by Stapel and Koomen (2001, Study 5). Both self-construal (I vs. we) and importance (important vs. unimportant) were manipulated, and an upward comparison target was presented as one who had excelled in an extremely difficult test subjects had just completed as part of the experiment (and on which they most likely did poorly). They found assimilation effects in the "I" condition and contrast effects in the "we" condition; both effects were more pronounced when the task was important than unimportant. Thus it is shown that a threat situation in which an upward comparison exists in an important domain elicits pronounced effects for self-evaluation, and that the direction of the effects is influenced by whether a social construal (integration mindset) or individual construal (comparison mindset) is activated. Unfortunately, this experiment does not evaluate performance as a consequence of the manipulations, limiting its applicability to this model.

### *Studies using Performance as DV*

Relatively few studies in social comparison have used performance as a dependent variable. Most tend to limit their explorations to self-evaluation or expectancies. Several studies that did use performance as dependent variable are explored in the following paragraphs.

In a study critical to the current hypotheses, Marx, Stapel, and Muller (2005) demonstrated that stereotype threat activates a collective self-construal that alters the way in which social comparison information is processed. This is one of few studies that both addresses social comparison and stereotype threat within a single framework and includes measures of performance and self-activation. The authors justify the segue into performance as a dependent measure by observing that “because reminding stereotyped targets about a group-based stereotype likewise increases accessibility of their stereotyped identity ...and thus their collective self ..., we argue that the same type of assimilation effects could occur for performance as they do for self-evaluations” (p. 433).

In this study, Marx and colleagues (2005) constructed a series of experiments to explore the effects of collective identity activation on academic performance and various self-evaluation measures. In Experiment 1, the researchers randomly assigned male and female subjects to a diagnostic or non-diagnostic condition (as in Steele & Aronson, 1995), then administered a self-activation test to determine whether a collective or individual identity was more salient. The self-activation test was the Wezwe Language Translation task, in which subjects are given a piece of writing in a foreign language (Wezwe) with about twenty words highlighted. Subjects are told that people unfamiliar with the language can sometimes identify the pronouns being used, and asked to attempt to translate the pronouns. By coding the number s of individual or collective pronouns suggested by each subject (e.g., “I,” “we”), researchers can determine whether collective or individual identity is activated. See Dijksterhuis & von Knippenburg (1998). Tests of simple effects following significant ANOVA findings ( $F_{(1,38)} = 4.37, p = .04$ ) indicated that female subjects in the diagnostic condition ( $M=6.83$ ) used significantly more collective pronouns than females in the non-diagnostic condition ( $M=4.90, F_{(1,38)} = 5.75, p=.02$ ) or males in either condition ( $M=4.80, F_{(1,38)} = 6.36, p=.02$ ) . Experiment 2 expanded these findings by illustrating that female subjects in a stereotype activated

condition identified significantly more strongly with their gender group and had lowered expectancies of success.

Finally, Marx and colleagues (2005) conducted a third experiment in which they hypothesized that math performance for women in a gender-activated condition (diagnostic math test) would assimilate to a positive female social comparison target because a collective identity had been activated, whereas the performance of those in a non-diagnostic condition would contrast with a high-performing target. This is precisely what they found: math performance for women who read about a positive female peer in the diagnostic condition were higher than those who read about the positive female peer in the non-diagnostic condition ( $F(1, 71) = 36.36, p < .01, \eta = .58$ ), but the opposite was true for those who read about a negative female peer ( $F(1, 71) = 16.83, p < .01, \eta = .44$ ). Therein it is illustrated that social comparison information is processed differently when the collective self is activated, and that without collective activation, performance contrasts with comparison targets.

Johnson and Stapel (2007b) expanded on these findings by including a performance measure and a more nuanced design. They presented college undergraduates with a comparison target who was described as a student who had won a highly-valued academic award because of either his analytical (mismatched domain) or verbal skills (matched domain), and that he was either older (attainable) or the same age (unattainable) as the subjects. The attainable target might be analogous to a role model, and was expected to boost self-esteem. Dependent variables included both self-esteem measure and performance on the Remote Associates Test framed as a verbal assessment. Analysis revealed that subjects exposed to the same-age (unattainable) target in the alternate domain (analytical) performed better than those who were exposed to the attainable target in either the matched or mismatched domain or to the same-age target in the same domain. Self-evaluations, on the other hand, were lower after subjects were exposed to the unattainable target than to an attainable target. The self-esteem boost following exposure to the attainable target supports the role model findings. The performance boost in the analytical domain after comparison with the same-age peer was framed as a recovery effect in which threatening comparisons prompt superior performance in an alternate domain to restore self-concept.

Taking the relationship a step further, Johnson and Stapel (2007a) conducted a similar experiment to determine whether a series of studies included performance as a dependent measure and tested whether a change in self-evaluation mediated performance shifts (Johnson & Stapel, 2007a). To find out, they exposed subjects to a moderate (peer in top 25% of class) or extreme (peer in top 5% of class) upward comparison before giving them a general self-evaluation measure and an ambiguous RAT test framed as a measure of “integrative orientation intelligence” (p. 1053). Subjects exposed to the extreme target had significantly lower self-evaluations and higher test performance than controls, while those exposed to the moderate target performed more poorly than controls. In Study 2, Johnson and Stapel explored the impact of self-affirmation on performance following comparison with a successful same-age or older peer. Results were similar in the non-affirmed group: comparison with the same-age peer reduced self-evaluation and enhanced performance in the alternate domain. In the affirmed group, these effects did not occur. The underlying processes may have been the same: reduced self-evaluation when comparing with a very superior in-group member might have produced a recovery effect when subjects were provided with an opportunity to excel in a different area.

Finally, Tesser (1991) revealed that tasks with greater complexity are more susceptible to threat than are simpler tasks. Subjects took a test in a relevant domain and were provided with fictitious results suggesting that either a friend or stranger had outperformed them, or that the friend or stranger had performed more poorly than they had. They were then given a second task involving memory for simple or complex strings of numbers. Subjects who believed that the friend had outperformed them (threat) performed more poorly on the complex task or than subjects in any of the other conditions. Tesser suggests that the threat of being outperformed by a close other created arousal, which facilitated good work on an easy task but impaired competence on a complex task.

The group membership of a successful other has also been shown to influence self-evaluation and performance. To explore this relationship between group membership and self-construal, Gordijn and Stapel (2001, Experiment 1) first primed a group or individual focus directly by asking subjects to list and describe either personal

characteristics or characteristics of groups to which they belong. Subjects (all of whom were undergraduate psychology students) were then asked to read a story about a highly successful or unsuccessful psychologist before taking a Trivial Pursuit type of test. Performance of subjects in the intrapersonal focus contrasted with the target (i.e., those with an upward comparison performed more poorly than those with the downward comparison), and the performance of subjects in the interpersonal focus assimilated to the standard.

A slightly different model, in which the target was presented as either an in-group or out-group member (attended same university or a competing university) and the manipulation of person-group focus was omitted (Gordijn & Stapel, Experiment 2). As expected, assimilation occurred in the performance of subjects who thought the target attended the same university (i.e., higher scores for the upward comparison and lower scores for the downward comparison), and contrast occurred for those who thought he attended the opposing university. In a third study, Gordijn and Stapel (2001) explored the influence of the number of targets presented to subjects (Study 3), predicting that “when an intergroup context is salient and the group comprises out-group exemplars, contrast effects ... should be obtained. Assimilation should occur when no intergroup context is salient ...or with a group of in-group exemplars” (p. 725). Applying the same target descriptions used in Study 2 to a group of five rather than a single subject, they found an identical effect: more correct items when the comparison targets were members of the subjects’ in-group.

#### *Summary of Social Comparison*

The preceding paragraphs have demonstrated that several conditions must exist for a comparison to influence outcomes for an individual. First, the domain in which the comparison occurred must be relevant to a person. A person who is not a golfer, for instance, will not be threatened by exposure to a highly successful golfer because the golfing domain has no importance to their identity. Second, the ability of a comparison target must be somewhat similar to the one who compares. An amateur golfer, for example, will not suffer a loss of self-esteem when comparing with Tiger Woods but may experience an effect when comparing with the more-successful colleague who began at the same time as himself. Third, influence of social comparisons is greater when the self

is activated. If individuals are focused on their own unique characteristics or are self-reflecting, they are more likely to experience a response to a forced comparison. Fourth, individuals with some level of uncertainty or in an uncertain situation have an increased tendency to compare and are more greatly influenced by such comparison. Finally, comparers who feel no control over their status will contrast with upward and assimilate to downward comparisons. This framework is offered as an overview, with detail of some of the studies presented in support of parallel processes later in the document.

### 2.3. Theoretical Support for Identity Relativity

Evidence from within and outside of stereotype threat and social comparison theories provide evidence for parallel processes. First, many of the primary effects and moderators are parallel for both stereotype-driven comparisons and individual comparisons. Second, both stereotype threat and social comparison appear to be influenced by the desire to maintain one's self-esteem. Finally, self-categorization and comparative processes intertwine to generate performance deficits. These theoretical underpinnings will be outlined in the following paragraphs.

#### *2.3.1 Parallel Phenomena*

Both stereotype threat and social comparison have revealed relationships in which subjects' self-evaluation or performance varies in the direction of the target or stereotype (e.g., an individual's self-esteem increases when faced with a superior comparison target or decreases when faced with a negative stereotype), and situations in which subjects' self-esteem or performance varies in the opposite direction of the target or stereotype (e.g., an individual's performance declines when comparing with a superior colleague or improves when faced with a negative stereotype), called contrast.

The terms "assimilation" and "contrast" take on somewhat hazy meanings if dwelled on through the lens of comparison. While assimilation is generally defined as the tendency for one's behavior to move toward that of a target, like when one's self-esteem rises when presented with a superior target, one might also use to describe a situation in which one's behavior moves toward a belief about one's in-group, as in stereotype threat. The same applies to contrast: in social comparison, it refers to a process in which one's self-estimation or behavior differs from that of another person or target, while stereotype activation might produce behavior that might differ from that of a stereotype (e.g., Bargh et

al, 1999). Both upward comparisons and activation of negative stereotypes lead to performance decrements, but one effect is considered to be a contrast effect (upward comparison) and the other is considered to be an assimilation effect (stereotype threat). One might therefore question: if in one case, the performance decrement is assimilative and in the other case contrastive, then how can a parallel be drawn? Clearly, these terms muddy the clarity with which the parallels might be observed.

If we reconceptualize the problem somewhat, however, the parallels become clear and unambiguous. Research in social comparison has consistently found that intragroup comparisons yield contrast effects, so that upward comparisons often lead to reduced self-evaluation and performance and downward comparisons lead to enhanced self-evaluation and performance. Similar findings have been found in stereotype threat research, though they must be cast somewhat differently. Rather than defining stereotype threat as an assimilative effect from a unidirectional statement about one's in-group (stereotype), we must redefine it as an intergroup comparison in which one's in-group is compared to an out-group. For this argument, it is not necessary to stipulate that the stereotypes are not true or not applicable to the individual; this argument needs only rely on perceived status of one's in-group relative to an out-group for the comparison process to be activated. When one compares oneself to a superior in-group member, one's performance declines. When one compares one's in-group to an (ostensibly) superior out-group, one's performance declines. In both situations, comparisons have been made; in both cases, performance has suffered following the comparison. Thus, we might describe both the consequences of stereotype threat and upward comparison to be contrast effects.

If we redefine stereotype threat responses as contrast effects, as in the preceding paragraph, then what cases of individual and group activation lead to assimilation? In the case of social comparison, assimilation takes place when no comparison must be made, as in the case of a much older successful target or role model (e.g., Lockwood & Kunda, 1997), in the case of categorical activation without identity activation (e.g., Mussweiler et al., 2000), or in the case in which an individual can control his or her status. Likewise, stereotype activation yields assimilation when no comparison is made, such as in the studies of priming effects by Bargh and others (1996) or a situation in which the status of

the group is changeable . When a social category is activated that is not part of one's identity, assimilation tends to occur through ideomotor processes.

### *2.3.2 Common Moderators*

Further support for the parallel aspects of the phenomena is provided by the common moderators affecting both processes. These moderators include domain identification, awareness of status, uncertainty, and self-esteem maintenance.

#### *2.3.2.1 Domain Identification*

Domain identification is such an integral component of both social comparison and stereotype threat that many studies often include only subjects who have been predetermined to identify with the domain. Extensive exploration of the importance of domain identification was provided under the moderators section of each overview. Domain identification influences one's susceptibility to the effects of both stereotype threat and to social comparison. Evidence from studies in both fields has shown that if the domain is important to the individual, self-evaluation and performance deficits can occur in the face of comparisons (e.g., Steele & Aronson, 1995; Johnson & Stapel, 2007a).

#### *2.3.2.2 Awareness of Status*

Just as domain identification strengthens the relationship between comparisons and performance, so does awareness of beliefs relating to the superiority of another group or individual increase the likelihood of performance shifts. Neither stereotype threat nor social comparison effects are likely to occur if one is not aware of the ostensible superiority of the comparison target or out-group. A skateboarding teen who is not aware that the category "skateboarder" activates the category "shoplifter" in the mind of storekeepers will not feel paranoid about security guards when walking through the local department store. Likewise, if one is not aware that the person sitting next to her in class has just been awarded a dissertation fellowship, she cannot suffer a decrement in her own self-evaluation of performance. A high school valedictorian who has not understood that his future college peers had all been valedictorians cannot feel threatened by them.

In order to experience the full effects of the stereotype or comparison, one must identify with the group in which the comparison is being made. A woman who strongly identifies with her femininity, for example, will experience greater threat than one who is less enveloped by that identity (Schmader, 2001), and a Black person with high in-group

identification will more likely experience stereotype threat than one who does not identify with his or her racial group (Steele & Aronson, 1995). Likewise, one will not feel threatened by a social comparison if her identity is not linked to the entity under comparison (though studies have apparently not tested this effect). One who has lost money at a Las Vegas Blackjack table will more likely experience a decline in self-evaluation if he identifies considers himself to be a gambler than if he only dabbles in the activity on an occasional vacation.

Empirical evidence for the importance of awareness of status (or stigma consciousness) in eliciting stereotype threat has already been established (e.g., Brown & Pinel, 2003), and evidence supporting this idea for social comparison can be found in research related to competitive versus cooperative classroom environments. When comparisons are not made, as in the case of cooperative (as opposed to competitive classrooms), self-esteem and performance cannot suffer as a consequence. In an experiment with high-achieving elementary school, students randomly assigned to cooperative or competitive environments for a six-hour social studies unit, both academic self-efficacy and performance (measured through a textbook-based multiple choice exam) were significantly higher in the cooperative than in the individualistic condition. A meta-analysis of ten studies including self-esteem as a dependent measure suggested that self-esteem was highest in cooperative settings, and that the effect was most pronounced when cooperative settings were compared with individualistic settings. Because shared goals in cooperative environments make social comparisons less important, individuals in such environments will be less likely to be aware of (and thus suffer from) social comparisons.

Although the impact of competitive environments has not been systematically studied in stereotype threat research, a suggestive study by Peterson, Johnson, and Johnson (year) offers the possibility that relationship may exist. Through the framework of social interdependence theory and social judgment theory, they hypothesized that the relative status of males and females would become nullified in cooperative environments. Their data supported this hypothesis: boys and girls randomly assigned to cooperative conditions reported generated more equal gender-based status scores (based on peer nominations) than did those in individualistic conditions, where males retained a superior

status. The effect was significant in the area of math and science, in which status ratings for both males and females increased significantly in cooperative settings (compared to male-only improvements in individualistic settings). This leads to the suggestion that stereotype salience might be reduced if cooperative environments are used in early grades.

### *2.3.2.3 Situational and Personal Uncertainty*

Finally, both situational uncertainty and personal uncertainty appear to play a role in stereotype threat and social comparison. Situational uncertainty refers to aspects of a given situation that make the subjects' ability to perform uncertain, such as a high-diagnostic test or an ambiguously framed test. In such situations, a subject cannot know what to expect or how well he or she might measure up. Personal uncertainty refers to a trait characteristic—a tendency for an individual to fluctuate in self-efficacy estimations. Both appear to play a role in both phenomena under investigation.

*2.3.2.3.1 Situational uncertainty.* Situational uncertainty underlies many of the experimental designs exploring the outcomes of social comparison. Subjects in experimental designs rarely know the purpose of the experiment, nor do they have information about their relative skill levels at the task in which they are asked to engage; thus, uncertainty is great (though the domain is not likely salient unless experimentally manipulated). In the non-experimental designs exploring social comparison, subjects have been selected because they are in situations that are inherently uncertain. For instance, Buunk, Collins, Taylor, and VanYperen (1990) selected as subjects cancer patients (Study 1) and people in troubled marriages (Study 2), both involving situations high in uncertainty. In another study, comparison patterns and outcomes were explored in gifted students (at the top of their traditional classes) enrolled in a challenging program with other equally gifted students (Gibbons & Buunk, 1999) or freshmen entering high school. In each case, social comparisons were made salient due to both the importance of the domain to the individuals and the uncertainty of the situation.

The original conception of stereotype threat also did not define a situational uncertainty component, but referred instead to an “immediate situational threat that derives from the broad dissemination of negative stereotypes about one’s group—the threat of possibly being judged and treated stereotypically...” (Steele & Aronson, 1995,

p. 798). Such uncertainty will be enhanced if it has been aggravated by lifetime experience. Aronson & Inzlicht (2004) explain that stereotyped individuals' "performance, the feedback they receive, and their interpretations of such will leave room for uncertainty about their competence" (p. 829). Though not termed uncertainty in the initial research, one might reasonably draw a parallel between the situation described by Steele and Aronson—the possibility that one will be judged according to a stereotype—and the analogous social comparison situation in which one may find that she or he is inferior to another. In both cases, a fear of inferiority in an uncertain situation dictates a reaction due to a group comparison (i.e., women are perceived to have poorer math abilities than men; I am a woman) or an individual comparison (i.e., I have or may be perceived as having poorer math abilities than Jessica).

*2.3.2.3.2 Personal uncertainty.* In addition to situational ambiguity, personal uncertainty appears to play a role in both social comparison and stereotype vulnerability. Summarizing an expanse of social comparison literature during development of a social comparison scale, Gibbons & Buunk, (1999) explained that "interest in social comparison is associated with uncertainty [and other traits with] significant uncertainty components" (p. 130) such as neuroticism, low self-esteem, or depression. Adding to this knowledge, Stapel and Tesser (2001) explored the relationships among self-certainty (the degree to which one's self-concept is well-defined and stable), self-activation (the degree to which self-schemas are accessible), and social comparison orientation. Their explorations revealed that subjects low in self-concept clarity (measured through the Self-Concept Clarity Scale; Campbell, 1990) experienced higher levels of self-activation (measured through the Wezwe Language Translation Task) than those high in self-clarity (Study 5), that higher self-activation corresponded with greater interest in social comparison information (Studies 2a and 2b), and that low self-certainty lead to higher social comparison orientation regardless of manipulations of self-activation (Study 3). Additionally, Stapel and Tesser created an experimental design in which they manipulated self-uncertainty by providing either clear or ambiguous feedback on an ostensible personality test (Study 4), resulting in findings paralleling those in Study 3. Finally, subjects low in self-certainty and intolerant of uncertainty had higher social

comparison orientations (Butzer & Kuiper, 2006) and experienced different self-appraisal reactions to social comparisons (Bui & Pelham, 1999) subjects high in self-certainty.

As with social comparison, personal uncertainty (as opposed to self-confidence or self-esteem) was also shown to impact the influence of a stereotype threat situation. In two experiments, Aronson and Inzlicht (2004) explored the characteristics of and outcomes of what they termed “stereotype vulnerability,” or the “tendency to expect, perceive, and be influenced by negative stereotypes about one’s social category” (p. 830). Stereotype vulnerability, they hypothesized, contributes to poor academic self-knowledge and unstable efficacy due to mistrust of performance feedback and inconsistent performance due to stereotype threat. Indeed, their studies confirmed their hypotheses. Subjects were asked to indicate the likelihood that their answer was correct while responding to test questions from the Graduate Records Exam (GRE). Those that were higher in stereotype vulnerability were “more miscalibrated—and more overconfident—that Blacks with low stereotype vulnerability” (p. 831). To clarify the nature of the results, the researchers conducted a second experiment in which they gathered information about the stability of students’ self-efficacy by requiring Black and White students to complete self-efficacy measures twice daily for 8 days before completing the RS-Race Scale to measure stereotype vulnerability. They found that, indeed, Black students high in stereotype vulnerability reported “higher highs and lower lows [in academic self-efficacy] than all of the other participants combined” (p. 833). This inconsistency was not found for athletic efficacy or for global self-esteem, however, suggesting that the instability existed for stereotyped individuals only in stereotyped domains.

It must be stipulated that stereotype threat is not about one’s doubt about individual ability. As Steele and Aronson directly stipulate in reference to an experiment exploring math stereotype threat for women, “these women were selected for being very good at math and for reporting high confidence in their ability” (p. 620). Stereotypes, as framed in this paper, however, are not about one’s individual ability relative to other individuals but about the perceived ability of one’s group relative to another group. As such, higher individual ability would increase the importance of the domain and heighten

the impact of stereotype-driven comparisons (especially such exasperatingly persistent and uncontrollable social comparisons as stereotypes) on performance.

Thus we might conclude for the sake of this line of thinking that individuals with high levels of trait uncertainty in ambiguous situations are susceptible to what might be aptly termed “social comparison vulnerability.” Likewise, individuals who have unstable self-efficacy—or personal uncertainty—experience greater stereotype vulnerability.

#### *2.3.2.4 Maintenance of Self-Esteem*

The extent to which one identifies with an in-group or comparison domain appears to shift through the persistent desire of individuals to maintain their self-esteem. If one’s self-esteem is threatened by either a group or individual comparison, that individual will take steps to preserve his or her self-esteem through shifts in performance, domain disidentification, or disidentification with the in-group. Turner and colleagues note that not only are people motivated to belong to groups that are positively distinctive, but that they will “attempt to leave that group (psychologically or in reality) ... when social identity in terms of some group membership is unsatisfactory” (p. 30). Empirical support for these tendencies can be found below.

Within social comparison research, an explanatory theory has emerged through which the tendency to disengage from a threatening domain appears to be guided by the need to maintain one’s self-esteem within a larger social framework (Tesser, 2000). In uncertain situations, individuals will choose comparisons that are self-enhancing or boost self-esteem, and react with a variety of defensive mechanisms if a social comparison is threatening. This process is described as one which “assumes that people are motivated to maintain a positive self-evaluation” (Tesser, Pilkington, and Macintosh, 1989, p. 442), and is termed the self-esteem maintenance model.

Supporting the self-esteem maintenance model, several studies have demonstrated that individuals will take steps to protect self-esteem in the case of negative feedback. In one study, gifted students who typically performed at the top of their classes made “downward shifts” in their comparison targets and reduced the number of comparisons they made when placed in a situation where their performance was merely average (Gibbons & Buunk, 1999). In another study, students who could not make an external attribution for their own poor performance tended to reduce effort and downrate the

importance of the subject area (Thompson & Perry, 2005). People with tenuous self-esteem employed downward social comparisons more frequently than those whose self-esteem was stable (Thompson & Perry).

In stereotype threat, this same process has been described as a fight or flight response. If one cannot fight a stereotype in a given domain, she or he must either abandon the stereotyped identity or disengage from the stereotyped domain (Davis, Aronson, & Salinas, 2006). The process operates, suggests Steele (1997), “by preventing or breaking a person’s identification with school, in particular, those domains of schooling in which the stereotype applies” (p. 622). The effort to protect self-esteem—a key effort in response to both unfavorable social comparisons (Tesser, 2000) and stereotype threatening situations—leads stereotyped individuals to (among other things) generate “a pattern of devaluing domains or attributes on which one's group is disadvantaged” (Crocker & Major, 1989, p. 622).

Further, like subjects whose self-esteem in a given domain is threatened by social comparisons, global self-esteem appears not to suffer. Rosenberg (1979; in Steele, 1997), for instance, suggests that the global self-esteem for stereotype-threatened Black students is higher or as high as White students—thus, preservation of their overall self-regard can be attributed to “disidentification with domains in which their evaluative prospects were poor ... and identification with domains in which their prospects were better (i.e., their peers)” (Steele, p. 623). In the case where female test takers were led to believe that their poor performance was due to innate gender-based ability differences, they dissociated with the math subject area to a greater degree than those who believed that socialization prompted gender differences in math ability.

*2.3.2.3.4 Self-Handicapping.* Buunk and Gibbons (2007) explored a range of defensive mechanisms to which people might subscribe in the face of unfavorable comparisons. One such mechanism involves self-handicapping, in which individuals make excuses for what they perceive as their own inferior performance (by indicating that they had not slept well the previous night, for instance). In social comparison, self-handicapping was viewed in increased tendencies to choose a clearly superior other as a comparison target to increase the basis of one’s poor performance (Shepperd & Taylor, 1999) or by labeling a better-performing other as a “genius” (Alicke, LoSchiavo, Zerbst,

& Shaobo, 1997). Mussweiler, Gabriel, and Bodenhausen (2000) showed that, after performing worse than a target, participants, especially those with high self-esteem, deflected the comparison by strategically emphasizing aspects of their identity that differentiated them from the standard. Self-handicapping processes also have been reported in studies of stereotype threat, indicated in the increased likelihood that those who were threatened were more likely to respond with agreement to statements that might excuse their performance (e.g., “I am tired today”; e.g., Steele, 1997).

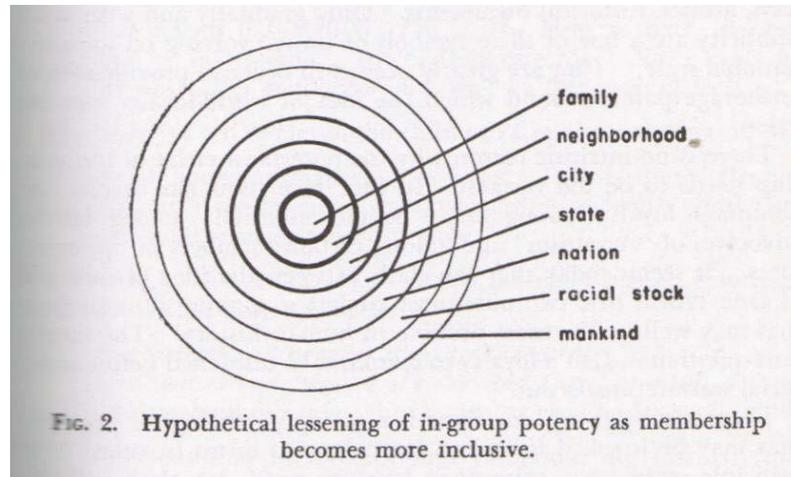
### *2.3.3 Multiple Selves*

Within each individual exists multiple facets. Just as light penetrating a crystal highlights different aspects of its core depending on the facet through which it enters, so do environmental stimuli activate different identities depending on the facet through which they are perceived. Just as light hitting a crystal reflects to different locations depending on the facet from which it bounces, so do human responses to stimuli differ according to the facet from which they are projected. In crystals, the facets are differently angled edges through which light is infused and reflected; in humans, the facets are different identities through which stimuli are perceived and reacted. The identities represented by the facets are multitudinous. Some of the identities are individual, in which one’s unique self creates the framework through which response are formed; other identities are social, in which one (or more) of the groups to which one belongs is predominant.

The nature of an individual’s identity structure is debated. For example, Marcus and others (1979; in Onorato) argued for the existence of a single dynamic self-schema that integrates individual and collective identities to generate behavior, while others argue that identities exist within other identities with varying levels of importance at each level. A figure taken from a book by Gordon Allport (1954) illustrates the hierarchical and concentric nature of an individual’s various identities (see Figure 1). The darkness of the lines with which the circles are drawn corresponds to the relative power of the in-group, such that “in-groups grow weaker and weaker the larger their circle of inclusion” (p. 43). The potency in Allport’s example refers to loyalty to the in-group, and does not include a point in the center denoting the identity of the individual him or herself. Nevertheless, the

image presents a strong illustration of an organizational structure relevant to identity theory—most notably, self-categorization theory (Turner et al., 1987).

Figure 1. Concentric circles of loyalty (Allport, 1954)



### 2.3.3.1 Self-Categorization Theory

Self-categorization theory represents a hierarchically organized identity structure, much as is illustrated in Figure 1. The framework derives from an integration of social identity theory of intergroup processes (Tajfel, Turner, Jost, & Sidanius, 2004) and social identity theory of intragroup processes, or self-categorization theory (Turner et al., 1987). Social identity theory of intergroup processes offers an exploration of relations between and among groups, suggesting that groups in which one perceives oneself as a member form an aspect of an individual's identity that links the characteristics of the group with the self-concept of the individual. Self-categorization theory explores the psychological mechanisms of group memberships and their resulting behaviors, suggesting that individuals possess hierarchical levels of identity in which lower levels (e.g., personal characteristics) are contained within higher levels (e.g., race) which are contained within still higher levels (human).

The integration of these two theories provides the mechanism through which stereotype threat can be construed. Because individuals desire positive evaluations of themselves, and because group memberships are incorporated into the individual's self-concept, as social identity theory suggests, they will desire and even create conditions in which they can have positive evaluations of their groups or categories (Tajfel et al.,

2004). The model suggests that self-concepts related to various social categories “tend to be activated ... in specific situations producing specific self-images” (p. 44). The model also notes the fluidity of self-perception, observing that one’s sense of similarity or differences with in-group and out-group members fluctuates in predictable and inverse directions according to the situation. If one’s social identity is made salient, then in-group similarities and out-group differences are maximized, but the inverse is true when one’s individual identity is made salient. Turner (1987) notes the occurrence of depersonalization, explaining it as “a process of ‘self-stereotyping’ whereby people come to perceive themselves more as the interchangeable exemplars of a social category than as unique personalities defined by their individual differences from others” (p. 50). Thus, individuals’ identities are structured systematically into a hierarchy of levels that can be situationally activated and in which can enmesh with the characteristics and identity to a group to which they belong. Finally, identities that are closer the center are more likely to produce comparative effects than those toward the outside of the circle.

#### *2.3.3.2 Shifting Standards*

While positive social identities are highly motivating to individuals, determination of positive or negative valence may be relative rather than stable. Kahneman and Miller (1986; in Ross & Nisbett, 1991) suggest that “every stimulus recruits comparison stimuli from memory against which it is judged” (p. 63). Although the emphasis of their work was on decision-making, they suggest that “people’s assessment of themselves is inherently a comparative one is part of the grain of social psychology by now” (Ross & Nisbett, p. 65). Social comparison at the individual level has been accepted as a staple of social psychology, such that Tesser (1980; in Ross & Nisbett, 1991) suggests that people “believe themselves to be talented or untalented, rich or poor, healthy or unhealthy by comparison with others” (p. 65).

Although stereotype threat has not been framed in a comparative perspective, Turner and others (1987) suggested that “subjects define and evaluate themselves in terms of the imposed social categories, compare in-group with out-group in terms of the only available dimensions”(p. 30). . This idea is advanced through the “shifting standards model of judgment”(Biernat, Manis, Stapel, & Suls, 2007), which suggests that individuals judge that the performance of other people in accordance with a “frame of

reference [that] shifts for different categories or types of target” (p. 76). This phenomenon was illustrated empirically in a study showing that comparative estimates of targets’ attributes varied according to the expected comparison group (Biernat & Manis, 1994).

Illustrating this phenomenon, Biernat and Manis (1994) described stereotypes as representations of the “mean level of an attribute that members of a given group possess” (p. 76), therein imbuing stereotypes with a statistical flavor from which further expansions can be made. If the stereotype represents the mean level of a given attribute for a particular category of people, then such attributes can be compared to the same attribute for people in other categories. The stereotype that women have strong verbal skills, for instance, can be construed as a comparative statement “women are more verbally skilled than men” (Biernat et al., p. 77). Other stereotypes can be presented in a similar form; as was discussed earlier, ‘Asian people are good at math’ might derive its character from the ending ‘relative to non-Asian people.’ The strength of the stereotype can thus be attributed to the implicit comparisons contained within the stereotype rather than to a non-comparative belief about a group.

Stereotypes take on additional statistical properties in this model, in that its definition includes not only the mean level of a given attribute but also “a likely range [of that attribute] that members of the group will exhibit” (Biernat et al., 1994, p. 76). Since a stereotype contains information about both the expected mean level of an attribute and the expected range of manifestations, certain comparisons are suggested. We can not only compare the mean level of one group to the mean level of another group, but we might also judge the behavior or attribute of an individual based on the range of that attribute for the group to which s/he belongs. “She is a very good manager” we might observe, “for a woman.”

To what extent do people engage in such comparisons? The tendency to make comparisons is thought to be a natural and automatic component of cognitive processing, spontaneous in its occurrence and consequences. Mussweiler (2007) write that comparisons “appear to occur so naturally that they are even engaged with standards that are clearly irrelevant or—at least phenomenologically—not even there” (p. 180). Illustrating the shifting standards paradigm, for instance, Biernat and Manis (1994) found

that while females were judged to be more competent for feminine jobs (and vice versa for males), an individual male in a feminine job received HIGHER competence ratings than an equally competent female in the same feminine job. Thus, the standards used to judge an individual's behavior shifted on the basis of the general standard for the group to which the person belongs. Thus we conclude that (a) stereotypes are activated automatically when categories are activated, (b) stereotypes are made of group-to-group and individual-to-group comparisons, (c) comparisons in general are made automatically and without conscious awareness whenever cognitive processing occurs.

#### *2.3.3.3 Self-Stereotyping*

Self-categorization theory offers a view of individual identity structure that can be used to understand different responses to comparison information. Considering identity structures, Onorato and Turner (2004) compared self-schema theory (in which individuals are presumed to have a single conception of themselves that may shift over time) with self-categorization theory. Their method involved activating a group or individual identity in a prime condition, then requiring that male and female subjects indicate their degree of agreement with adjectives related to independence/dependence (shown in earlier studies to reflect males and females differently). Prior measures had been taken to determine their individual dependence/independence schematics. Results supported the self-categorization theory, indicating that “independent schematic females and dependent schematic males exhibited self-concepts diametrically opposed to their personal self-schema when the context evoked a self-categorization in terms of their respective social identities” concluding that “social identity can sometimes wipe out the effects of personal identity” (Onorato & Turner, p. 266).

Comparing self-categorization with self-stereotyping, Simon and Hamilton (1994) found that subjects tended to self-categorize (e.g., indicated belongingness to an in-group ) and self-stereotype (i.e., take on the characteristics of the in-group ) based on the size of the in-group . Specifically, subjects were placed into either a majority (80% of population) or minority (10% of population) high- or low-status condition in a contrived response to their preference for one or another painting. When their choice apparently put them into a majority group, no significant differences existed between those who thought their group was high or low status. When their choice left them in the minority group,

however, those who believed their group to be high status expressed significantly greater similarity with group members than did those who believed their group to be low status. Further, members of the contrived high-status group demonstrated an increased likelihood to assume the presumed characteristics of the group (introverted or extraverted). These findings support the connection between self-categorization and self-stereotyping.

## 2.4 The Current Research

An experiment will be undertaken to establish that intergroup comparisons can produce performance and self-esteem consequences similar to those produced by interpersonal comparisons. The study will compare the effects of intergroup and interpersonal comparisons on the performance and self-esteem of community college students about which a weak negative stereotype exists. The study will use performance on an ambiguous test and three measures of self-esteem (explicit, implicit, and collective) as dependent variables to explore both phenomena, and will explore previously obtained social comparison orientation and group identification data as moderators. The following hypotheses are proposed, based on the literature findings:

1. Both intergroup and interpersonal comparisons (stereotype threat and forced social comparison) will yield performance shifts that contrast with the comparison target or group.
2. Both intergroup and interpersonal comparisons (stereotype threat and forced social comparison) will both produce shifts in self-esteem, but contrast will occur at the individual level under intrapersonal comparison conditions and at the group level for interpersonal comparison conditions.
3. Individuals most likely to make individual comparisons (high social comparison orientation, or SCO) will also be most susceptible to the effects of intergroup comparisons or stereotype threat.
4. Individuals who identify most strongly with a comparison dimension will be most susceptible to contrasting effects due to stereotype threat and social comparison.

## CHAPTER 3. METHODS

This experiment employed a completely between-subjects experimental framework using cross-sectional data. The population of interest was, broadly, members of negatively stereotyped groups, and specifically, students who attended community colleges. Voluntary subjects were randomly assigned to one of five treatment conditions invoking comparisons: individual-upward, individual-downward, group-upward, or group-downward, or to a neutral control condition. Four dependent variables, including an academic performance measure, implicit and explicit individual self-esteem measures, and a collective self-esteem measure, were analyzed separately.

### 3.1 Subjects

The subjects for this study included undergraduate students who aspired to earn a Bachelor's degree or beyond but who were currently attending a two-year community college. Evidence supports the existence of a negative stereotype about students from community colleges. One survey-based study found that students attending both two- and four-year colleges perceived a community college education as "less valued by society in general" (Caporrimo, 2008, p. 25), and another study reported that the media have created negative images of community colleges despite their well-conceived foundational tenets (DeGenaro, 2006). For this reason, community college students served as a valuable subject pool for a study about stereotype threat and social comparison. Details about the final subject pool are provided in the Results section following description of the manipulation checks.

### 3.2 Manipulations and Measures

Students were randomly assigned to one of the five experimental conditions before taking a test described as a measure to predict college success, and before providing several measures of self-esteem. The primes served to activate either an individual identity relative to another student, a collective identity as a community college student relative to students at other colleges, or a neutral identity. Primes in both the interpersonal and intergroup comparison conditions were either upward, prompting comparison with a superior individual (Condition 1: Upward Individual - UI) or a superior social category (Condition 2: Upward Group - UG), or downward, prompting comparison with an inferior individual (Condition 3: "Downward Individual - DI) or an

inferior social category (Condition 4: Downward Group - DG), or to a control group (Condition 5 - Control). Three other dimensions were explored as potential moderators for the effects: social comparison orientation, domain identification, and identification with the group of community college students. The following paragraphs describe the independent, dependent, and moderator variables in more detail.

### *3.2.1 Independent Variables*

Independent variables are described first for subjects in group conditions, then for subjects in individual conditions, and finally for subjects in the control group.

*3.2.1.1 Intergroup comparison.* Students assigned to one of the two intergroup comparison conditions were exposed to two manipulations to activate their collective identities:

First, group comparison processes were activated by randomly assigning subjects to read one of two fictitious newspaper articles about the quality of students attending another college. (see Appendix A for all of the news articles). The article used in the upward intergroup comparison condition (we-up) extolled the academic skills of students attending a highly-regarded neighboring university, stressing the increased likelihood of degree completion for students attending that university. The article in the downward intergroup comparison condition (we-down) bemoaned the academic challenges faced by students attending technical college, implying that students attending community colleges have a decided academic advantage. The newspaper articles were embedded in a larger task named the “Newspaper Styles Survey for the Department of Journalism,” in which the prime article and a neutral article were presented for the ostensible purpose of exploring how accurately subjects can distinguish Internet-based from newspaper-based articles. This is a modified version of a prime used in many social comparison studies (see Lockwood & Kunda, 1997).

Second, students’ identities as community college students were activated through a demographic form on the first page of the test. (See Appendix B for all demographic items). The demographic form for students in the group comparison condition included an item on which students were asked to check the box next to their college from a list containing only their own college and the comparison college described in the news article. In previous studies, checkboxes such as this have been used to activate a

particular social identity (e.g., Steele & Aronson, 1995). Students were also asked to indicate how long they have attended that college.

*3.2.1.2 Individual comparison conditions.* Individual conditions were also activated through mechanisms similar to those for activating intergroup comparisons. First, subjects completed the Newspaper Styles Task (described above), in which they were asked to read a fictitious newspaper article about a successful or unsuccessful target “Casey Halverson” under the ostensible purpose of identifying the media source of the story. The upward individual (UI) condition described a successful, well-liked student at their own college who recently won an award for outstanding academic performance (see Appendix A). The downward individual (DI) condition described a low-average student at their college whose performance was mediocre and who feared non-degree completion. To ensure that an individual comparison process was activated, the demographic form included an item asking students to record the initials of either a high- or low-performing student in their high school, and write two indicators that the particular student had excelled or failed (See Appendix B).

*3.2.1.3 Control.* The control condition was identical to the other conditions, except that the two articles addressed neutral topics including stories about a bridge opening and a ship (Appendix A). Further, the demographic form requested only general information about miles traveled to campus and the size of the high school graduating class (Appendix B).

### *3.2.2 Dependent variables.*

*3.2.2.1 Performance.* The performance measure included 32 items similar to items from an IQ test (see Appendix C). The content of the test was ambiguous so that it could be framed as generally predictive of college success, and it was entitled “Measure of Intellectual Potential for College Success.” The test instructions indicated that the new test items were created to explore differences in intellectual skills across universities, and that the items have been predictive of academic success.

*3.2.2.2 Implicit self-esteem.* The implicit self-esteem measure included the signature size test (Stapel & Blanton, 2004; Webb, Campbell, Schwartz, & Sechrest, 1966) as an indicator of shifts in implicit self-esteem over the course of the experiment. Signatures were obtained at three points: (a) just before the newspaper task, (b) just after

reading the newspaper articles, and (c) just before the academic test. The smallest rectangular area in which the signature could be contained was measured for each of these signatures, and the differences in size served as a dependent variable measuring implicit self-esteem.

To gather these measures, all of the signatures were scanned into a PDF document and copied and pasted into a Word document. Within the Word document, each signature was selected in a drawing box and cropped to fully enclose the signatures within a rectangle. The precise dimensions of each box were then taken from the “properties” of the text box listed in Word. To compare the sizes, a dependent t-test was used within each comparison condition to determine whether a significant change in size had occurred between signatures 1 and 2, and between signatures 2 and 3.

*3.2.2.3. Individual and collective self-esteem.* Individual self-esteem was gauged in a single measure including five items interspersed among other items described in the following paragraph. The five Likert-scaled items were taken from a state self-esteem scale developed by Heatherton and Polivy (1991) and used by Cohen and Garcia (2005). The items read “I feel smart,” “I feel confident in my abilities,” “I am concerned about the impression I am making,” “I feel that others respect me,” and “I am worried about what other people think of me.”

Collective self esteem was measured with four items from the Collective Self-Esteem (CSE) scale (Luhtanen & Crocker, 1992). Some of the items were designed to measure private collective self-esteem, described by its authors as measuring “global evaluations of one's ascribed social groups . . . the group-level equivalent of global personal self-esteem” (p. 62). These items read, ““I often regret that I am a community college student,” “In general, I'm glad to be a community college student,” Two other items measured public collective self-esteem, or how others evaluate one's social groups. The items read “Most people consider community college students to be more effective than students from other schools,” and “In general, others think that community college students are unworthy.” These items were grouped for the purposes of analysis. See Appendix D.

### 3.2.3 Moderators

Items related to social comparison orientation (SCO), importance to identity (ITI) and domain identification (DID) were also included on the scale for exploration as potential moderators, and are included in Appendix D. The basis for their use was established through earlier studies suggesting that subjects who identified highly with the domain in question (in this case, academic performance), who identified strongly with the group in question (in this case, community college students), and who had a tendency to compare often were more susceptible to threat effects. These data were gathered at the same time as the explicit self-esteem measures immediately after the test, so were subject to possible contamination by the manipulations in the study. Their usage was therefore primarily exploratory.

ITI data, or the extent to which participants identified with their role as community college students, were gathered through the Collective Self-Esteem (CSE) scale (Luhtanen & Crocker, 1992) with items reading “Overall, the fact that I am a community college student has very little to do with how I feel about myself” and “Being a community college student is an important reflection of who I am.”

Domain identification, or the extent to which participants valued their success as students, followed a protocol established by Johnson and Stapel (2007a). It contained items reading “Being a student is important to me,” “My academic performance means a lot to me,” and “I don’t really care that much about how well I do in college.”

Social comparison orientation, or the degree to which participants tended to compare themselves with others, was measured through the Iowa-Netherlands Comparison Orientation Measure (INCOM: Gibbons & Buunk, 1999) This measure included the following items: “I often compare myself with others with respect to what I have accomplished in life,” “If I want to find out how well I have done something, I compare what I have done with how others have done,” “I always pay a lot of attention to how I do things compared with how others do things,” “I am not the type of person who compares often with others,” and “I often compare how I am doing academically with other people.”

### 3.4 Manipulation checks

Several steps were taken to ensure that the manipulation was effective. These steps included a measure of identity activation and a check of the prime manipulation.

#### *3.4.1 Identity activation and valence.*

Activation of the expected identity level was assessed using the Twenty Statements Task (Kuhn & McPartland, 1954) to determine whether individual or group identity was activated and whether a positive or negative valence was active. The task required subjects to complete the statement “I am \_\_\_\_\_” fifteen times with the purpose of explaining who they are through their responses. (The measure was shortened from twenty items originally for the sake of time). See Appendix E. This task was completed immediately following the performance measure.

Responses were coded along two dimensions: valence of the response (e.g., positive, negative, or neutral), and level of identity (group or individual). Coding was conducted by the principal researcher and by an undergraduate psychology student according to pre-established criteria. Both coders were blind to the condition of the respondent during coding. Any disagreements were discussed and a new rule was agreed upon. Once rules were all stated, the researcher reread each coded sheet to ensure that the new rules were applied equitably to each response form.

*3.4.1.1 Level of identity.* Responses were coded as group-activated if they suggested a social role or relationship, or if they indicated a group membership. Some responses were quite straightforward (e.g., I am a girlfriend, daughter, granddaughter, friend, and so forth). Others were less clear (e.g., I am a hunter, a skateboarder, a fisherman); these types of responses were coded as group activation if they were often associated with an in-group with which one engages in the activity or a group with shared objectives. Hunter, therefore, was coded as group activation because a strong subculture revolved around the activity, and “Democrat” was also coded as group because of identifiable, shared objectives. “Artist,” on the other hand, was coded as individual because it necessitated neither a shared subculture nor shared objectives. All responses not coded as group responses were coded as individual.

*3.4.1.2. Valence.* Valence coding was carried out according to the following criteria: Positive statements included attributes that would be judged to be favorable in

the dominant culture, such as “athletic,” “responsible,” “a good student,” “unique,” “upbeat,” “sociable,” or “generous.” They also included accomplishments (e.g., vice president of \_\_\_\_ club, straight-A student), favorable relationships (e.g., good friend, loyal boyfriend), and loving states (e.g., an animal lover, a lover of books). Neutral statements included social roles not qualified with an adjective (e.g., daughter, student, a football player), present states of being (e.g., tired, hungry, sick of taking tests), and characteristics that might be construed as either positive or negative (e.g., argumentative, mischievous, picky, single, thin). Negative statements included statements perceived negatively in society (e.g., fat, moody, ex-drug addict). The descriptors “shy” and “not rich” were included in the neutral category, but “poor” was included in the negative category.

See Attachment F for this document.

#### *3.4.2 Prime manipulation*

Among the final measures in the study, subjects were asked to rate the overall intellectual skills of their own comparison subjects. Students in the group comparison condition and in the control condition were asked to rate intellectual capacity of community college students, and those who compared in individual conditions were asked rate the success of Casey Halverson, the subject about whom they read in the newspaper article. Ratings were gathered on a 6-point Likert scale, and were obtained as the very last item in the study, immediately prior to debriefing. (The items are also contained in Appendix F).

#### *3.4.3 Suspicion*

Suspicion about the nature of the activity was also measured. During the debriefing process, subjects were asked if they believed that the newspaper story influenced their performance on the test, and if so, in what way. Data from subjects reporting suspicion regarding the study were scrutinized for further anomalies. See Appendix F.

### **3.5 Analyses**

Planned comparisons were used to test specific hypotheses about differences among the groups, specified in the following paragraphs. The Bonferroni correction procedure was used to minimize the potential for Type I error. For all four of the

dependent variables, alpha was set at .10 and divided across the comparisons. A sample size of 84 subjects was sought for a power of .80 with  $\alpha = .03$  to detect an effect size of .3, which was predicted on the basis of standards established in the meta-analysis conducted by Walton & Cohen (2000). The actual number of valid subjects was 76; thus, the power was slightly lower than desired. The following paragraphs describe the comparisons verbally and through statistical representations.

### 3.5.1 Performance

Hypotheses 1a and 1b: Because comparisons with an inferior group were expected to produce the same effect on performance as comparisons with an inferior individual, and because comparisons with a superior group were expected to produce the same effect as comparison with a superior individual, it was hypothesized that performance for subjects in the upward group (UG) condition would be equivalent to performance of subjects in the upward individual (UI) condition ( $H_{1a}$ ), and performance for subjects in the downward individual condition (DI) would be equivalent to performance of subjects in the downward group (DG) condition ( $H_{1b}$ ). With an estimation of 30 degrees of freedom, the critical t-value for a two-tailed test would be 2.28.

$$H_{01a}: \mu_{UI} - \mu_{UG} = 0 (\alpha = .02)$$

$$H_{01b}: \mu_{DI} - \mu_{DG} = 0 (\alpha = .02)$$

$$H_{a1a}: \mu_{UI} - \mu_{UG} \neq 0$$

$$H_{a1b}: \mu_{DI} - \mu_{DG} \neq 0$$

Hypothesis 1c: Because those who compared themselves with ostensibly superior groups or individuals were expected to produce performance decrements relative to those who compared with inferior groups or individuals, it was hypothesized that performance for subjects in a group consisting of the two downward conditions (DI+DG) would surpass performance of subjects in a group consisting of the two upward conditions (UI+UG). With 60 degrees of freedom in a one-tailed test, the critical  $t$  value would be -2.10.

$$H_{01c}: (\mu_{UI} + \mu_{UG}) - (\mu_{DI} + \mu_{DG}) = 0 (\alpha = .02) \quad t_{crit}(60) = -2.10$$

$$H_{a1c}: (\mu_{UI} + \mu_{UG}) - (\mu_{DI} + \mu_{DG}) < 0$$

Hypothesis 1d and 1e: Performance of subjects in the control condition were expected to surpass the performance of subjects in the UI+UG condition ( $H_{1d}$ ), and performance of subjects in the DI+DG condition were expected to surpass the

performance of subjects in the control condition ( $H_{1e}$ ). Estimating 45 degrees of freedom in a one-tailed test, the critical value will be -2.12.

$$H_{0_{1d}}: (\mu_{UI} + \mu_{UG})/2 - \mu_{Control} = 0 (\alpha = .02) \qquad H_{0_{1e}}: (\mu_{DI} + \mu_{DG})/2 - \mu_{Control} = 0 (\alpha = .02)$$

$$H_{1_{1d}}: (\mu_{UI} + \mu_{UG})/2 - \mu_{Control} < 0 \qquad H_{1_{1e}}: (\mu_{DI} + \mu_{DG})/2 - \mu_{Control} > 0$$

### 3.5.2 *Explicit individual self-esteem:*

Explicit self-esteem (ESE) was measured through responses to Likert-scaled statements such as “I feel smart.” It was hypothesized that ESE would vary according to the condition to which subjects were assigned.

Hypothesis 2a: Because subjects threatened at the individual level were expected to experience self-esteem decrements at the individual level, it was hypothesized that ESE for subjects in the DI condition would surpass scores for subjects in the UI condition.

$$H_{0_{2a}}: \mu_{DI} - \mu_{UI} = 0 (\alpha = .03) \qquad H_{a_{2a}}: \mu_{DI} - \mu_{UI} > 0 \quad t_{crit} (30) = 1.95$$

Hypotheses 2b and 2c: ESE for subjects in the DI condition was expected to surpass scores for subjects in the control condition ( $H_{2b}$ ), and scores for subjects in the control condition were expected to surpass scores for subjects in the UI condition ( $H_{2c}$ ). A one-tailed test ( $df=30$ ) will require a critical value of 2.15.

$$H_{0_{2b}}: \mu_{DI} - \mu_{Control} = 0 (\alpha = .02) \qquad H_{0_{2c}}: \mu_{UI} - \mu_{Control} = 0 (\alpha = .02)$$

$$H_{1_{2b}}: \mu_{DI} - \mu_{Control} > 0 \qquad H_{1_{2c}}: \mu_{UI} - \mu_{Control} < 0$$

Hypotheses 2d, 2e, and 2f: Because subjects threatened at the collective level and non-threatened subjects were not expected to experience state self-esteem threats at the individual level, it was hypothesized that ESE for subjects in the UG condition would be no different from ESE for subjects in the DG condition ( $H_{2d}$ ) or from subjects in the control condition ( $H_{2e}$ ). ESE for subjects in the DG condition was expected to be no different from to subjects in the control group ( $H_{2f}$ ) ( $\mu_{UI} = \mu_{UG} = \mu_{Control}$ ;  $\alpha = .01$  for each test).

### 3.5.3 *Collective self-esteem (CSE)*

Collective self-esteem, or the extent to which subjects identified with their role as community college students, was expected to vary according to the condition to which subjects were assigned.

Hypothesis 3a: Because subjects threatened at the collective level were expected to experience self-esteem threats at the collective level, it was hypothesized that CSE for subjects in the DG condition would surpass scores for subjects in the UG condition.

$$H_{03a}: \mu_{DG} - \mu_{UG} = 0 (\alpha = .03) \quad H_{a3a}: \mu_{DG} - \mu_{UG} > 0 \quad t_{crit} (30) = 2.15$$

Hypotheses 3b and 3c: CSE for subjects in the DG condition were expected to surpass scores for subjects in the control condition ( $H_{3b}$ ), and CSE for subjects in the control condition were expected to surpass scores for subjects in the UG condition ( $H_{3c}$ ). With 30 degrees of freedom in a one-tailed test, the critical  $t$  value is -2.15.

$$H_{03b}: \mu_{DG} - \mu_{Control} = 0 (\alpha = .02) \quad H_{03c}: \mu_{UG} - \mu_{Control} = 0 (\alpha = .02)$$

$$H_{a3b}: \mu_{DG} - \mu_{Control} > 0 \quad H_{a3c}: \mu_{UG} - \mu_{Control} < 0$$

Hypotheses 3d, 3e, and 3f: Because subjects threatened at the collective level and non-threatened subjects were not expected to experience self-esteem threats at the collective level, it was hypothesized that collective self-esteem (CSE) for subjects in the UI condition would be no different from CSE for subjects in the DI condition ( $H_{3d}$ ) or from subjects in the control condition ( $H_{3e}$ ). CSE for subjects in the control group were also expected to be no different from CSE for subjects in the DI condition or from subjects in the control group ( $H_{3f}$ ) ( $\mu_{UI} = \mu_{UG} = \mu_{Control}$ ;  $\alpha = .01$  for each test).

#### 3.5.4 *Implicit individual self-esteem (IISE)*

Implicit self esteem was measured through the collection of signatures at three points during the study: (a) upon entering the test situation, (b) after reading the newspaper article, and (c) after reading the test instructions. This process was described in detail earlier in the methods section. Dependent t-tests were used to determine whether the signature size for each subgroup shifted between the first and second measure, and between the second and third measure. Alpha of .05 was used for one-tailed tests according to predictions within each condition.

Hypothesis 4a: Because subjects who compared with inferior individuals or groups were expected to experience implicit self-esteem boosts that would manifest in signature size increases, it was hypothesized that signature size would increase significantly between the first signature (Sig1) and the second signature (Sig2) for subjects in a downward comparison group consisting of the DG+DI conditions.

$$H_{04a}: \mu_{dif} = 0 (\alpha = .05) \quad H_{a4a}: \mu_{dif} > 0 \quad t_{crit} (30) = 1.70$$

Hypothesis 4b: Because subjects who compared with superior individuals or groups were expected to experience implicit self-esteem deficits that would manifest in decreased signature size, it was hypothesized that signature size would decrease significantly between Sig1 and Sig2 for subjects in an upward comparison group consisting of the UG+UI conditions.

$$H_{04b}: \mu_{\text{dif}} = 0 (\alpha = .05) \quad H_{a4a}: \mu_{\text{dif}} < 0 \quad t_{\text{crit}} (30) = -1.70$$

Hypotheses 4c: Because an absence of a threat condition for subjects in the control group was expected to produce no effect on signature size, it was hypothesized that signature size would not change significantly between Sig1 and Sig2 for subjects in the control group.

$$H_{04c}: \mu_{\text{dif}} = 0 (\alpha = .05) \quad H_{a4c}: \mu_{\text{dif}} \neq 0 \quad t_{\text{crit(two-tailed)}} (30) = 2.13$$

No predictions were made for the differences between the second and third signatures (Sig3); these differences served as exploratory data.

### *3.5.5 Moderators*

Social comparison orientation, group identification, and domain identification served as moderators for this activity. To check whether higher levels of each of these variables had a moderating effect, all respondents were placed into either a high or low group based on the average score for the group. All subjects whose score for the specified measure were equal to or higher than the mean were listed as “high” on the measure in question, and all others were counted as “low” on that measure. Each of the hypotheses tested above was retested separately for subjects in the high and low conditions.

## **3.6 Procedure**

### *3.6.1 Participant selection*

Participants were recruited by means of extra credit through courses in the sociology, psychology, and history departments. Their professors explained the project to them according to a script (see below) about two weeks before the study was conducted (Appendix G). Students then signed a form and received informed consent forms (Appendix H) which they were asked to read, complete, sign, and bring to the study if they wished to participate.

*3.6.1.4. Materials.* Packets were prepared for each student containing the following materials: Signature Form 1, the Newspaper Styles task, Signature Form 2, Language Translation Task, Signature Form 3, Test Instructions & Questions, Test Items, Personal Responses, and Manipulation Check items. The Newspaper Styles Task, Test Instructions & Questions, and Manipulation Check differed according to randomly assigned conditions. All of the other materials were identical.

*3.6.1.4. Activities.* The experiment was conducted in a single day through a series of 50-minute sessions for which students had previously signed up. When they entered the room, they turned in their informed consent documents and received a packet of materials containing all of the items above. Packets had previously been randomly sorted according to conditions. Subjects were instructed both verbally and in writing that they should not begin the items or turn pages until instructed to do so. At that time, general parameters of the informed consent were reiterated (see Script, Attachment C).

The experimenter was a female psychology undergraduate with a male undergraduate assistant. The female experimenter greeted the participants, read the script, collected signatures, and answered questions. The assistant ensured that materials were properly organized and collected at the end of each session.

Subjects participated in the following activities in the sequence in which they were presented in the preceding paragraph. The experimenter introduced each item according to the script. Once the study was complete and packets were collected, students were completely debriefed and received an informational form entitled “About the Study” (see Attachment B). The researcher also gave the subjects a cookie to help to alleviate any bad feelings subjects may have experienced as a result of the study.

## CHAPTER 4. RESULTS

Results will be presented in three separate sections. First, the results of the data validation process will be described. Second, the results of the key hypotheses will be presented. Finally, the results of ancillary analyses will be presented.

### 4.1 Data validation

A process of data examination was conducted to ensure the validity of the data. This process included (a) examination of data from one session for which anomalies occurred during testing, (b) comparison of manipulation check items verifying the understanding of the primes, (c) consideration of open-ended responses indication suspicion, and (d) coding of a subjective measure determining whether the group or individual activation was effective. These topics are addressed in the following sections:

#### *4.1.1 Exploration of problem session*

An anomaly occurred during one experimental session involving eleven students. Most of students in this session entered the testing room in one large group behaving in a jovial and rather rowdy manner. During the testing session, one student's cell phone vibrated a number of times and prompted giggling and laughing among subjects during the test time. Rather than draw further attention to the vibrating cell phone by asking the student to turn it off, the experimenter ignored the phone, and it did stop eventually stop vibrating. However, the situation was sufficiently unusual that it was necessary to ensure that the results from this group did not differ significantly from other sessions.

To test this concern, a variable was created in which the problem session was one factor, and all of the other sessions were coded as another factor. Independent t-tests were then conducted on all of the dependent measures to ensure that significant differences did not exist.

Analysis revealed no significant differences between Session 1 and the other sessions on any of the dependent measures. T-values ranged from .34 to .847, with significance levels no less than .819. On the manipulation check related to group activation, however, Session 1 differed significantly from other groups. This difference was eliminated once the data were cleaned, as described in the next sections.

#### 4.1.2 Manipulation check 1—Prime comprehension

To verify that the subjects recognized the superior or inferior characteristics of the group or individual being primed in the newspaper article, subjects in the group activation condition were asked to rate the success of community college students through a 6-point Likert response to the question “According to the newspaper article you read at the start of the study, how successful are community college students?” (1=Not successful; 6=Very successful). Students in the individual condition were asked to rate the success of the comparison target, Casey Halverson, by responding to the question “According to the newspaper article you read at the start of the study, how successful was the student Casey Halverson?” Subjects in the control condition were asked to rate the success of community college students without reference to a newspaper article.

Of the 85 participants in the study, eight subjects responded in the opposite direction of what was expected. For example, two subjects who read about a very successful student in community college indicated that the target had a very low or low level of success in college. Because the prime appeared to be misunderstood, these subjects were excluded from all portions of the study that relied on understanding the news article. Of the eight subjects who were eliminated for responding in the unexpected direction, four were from the problem session discussed in the preceding section.

Of the remaining subjects, the primes appeared to have been correctly perceived. Subjects exposed to the down-group prime rated community college students as significantly more successful ( $M = 4.85$ ,  $SD = 1.444$ ) than subjects exposed to the up-group prime ( $M = 2.0$ ,  $SD = .756$ ,  $t(26) = -7.868$ ,  $p < .0001$ ). Similarly, subjects in the down-individual condition ( $M = 5.58$ ,  $SD = .515$ ) rated the comparison target as significantly more successful than subjects in the up-individual condition ( $M = 1.85$ ,  $SD = .801$ ;  $t(23) = 13.746$ ,  $p < .0001$ ). Six subjects in the control group indicated that they did not know how successful community college students were; the remaining subjects ( $n = 10$ ) rated community college students as very successful ( $M = 5.85$ ,  $SD = .483$ ). Interestingly, subjects in the UG condition rated community college students significantly lower than subjects in the control group ( $t(31) = 12.198$ ,  $p < .0001$ ), but those in the DG condition rated community college students no differently than subjects in the control group ( $t(29) = 1.172$ ,  $p = .174$ ).

#### *4.1.3 Manipulation Check—Suspicion*

Subjects were asked a series of questions immediately after the study to determine whether they had any suspicion about the true purpose of the news articles. Specifically, subjects were asked why they were asked to read the newspaper articles (example of non-suspicious response: “to see if we could find internet sources”). One subject indicated suspicion in the following response: “Make community college kids sound like losers.” Because the same person was already questionable because he might have seen the first page of the debriefing form before participating in the study, his data were excluded from the analysis.

#### *4.1.4 Manipulation Check 3—Group/Individual activation.*

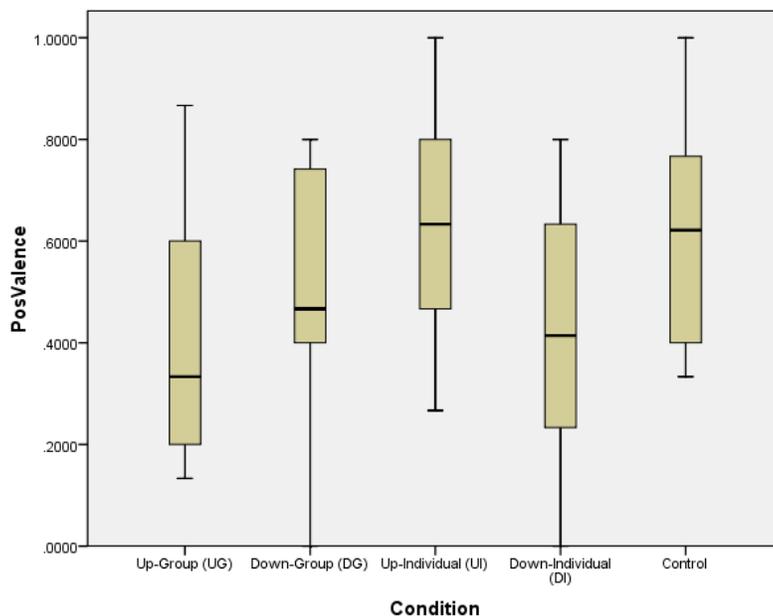
Valence and positive/negative activation were coded for the fifteen responses to the TST. Description of the coding process is presented in the Methods section. To calculate valence and activation level for each subject, the number of relevant responses (e.g., positive responses) was divided by the total number of accepted responses to produce a percentage. The percentage was used for statistical comparison. For example, if the subject produced a total of 15 responses of which 5 suggested individual activation and 10 suggested group activation, their group activation score would be 10/15, or .66. This score was used for analysis purposes. Overall, slightly more than half of the responses had a positive valence (suggesting positive feelings about the self), and fewer than 10% had a negative valence (suggesting negative feelings about the self). About 28% of the responses suggested group roles, and 72% suggested individual roles (See Table x).

It was expected that subjects in both the group and individual upward comparison conditions would produce more negatively valenced responses than those in downward comparison conditions, and that subjects in the group conditions would produce more group identifiers than subjects in the individual conditions. A one-way ANOVA was conducted using condition as independent variable; assumptions for normality and homogeneity of variance were checked using both visual inspection and statistical tests. Performance data for all of the factors appeared to meet the assumptions for ANOVA except for the data in the UG condition, which appeared to be non-normally distributed (Kolmogorov-Smirnov(17) = .211,  $p = .043$ ). The challenge to normality appeared to be

the result of a platykurtic distribution; because skewness was below .8 (skewness = .573) and kurtosis below 3 (kurtosis = -1.030), however, data transformations were not recommended (citation). Thus, results will be viewed with caution.

As expected, significant differences existed in the proportions of positively valenced responses produced by subjects in different conditions ( $F(4, 71) = 2.899, p = .028$ ). Visible inspection of boxplot suggested the following potentially significant comparisons:  $UG < DG$ ;  $UG < UI$ ;  $UG < Control$ ;  $UI > DI$  (see Chart 1); these four comparisons were tested for statistical significance using the Bonferroni correction for post-hoc testing ( $\alpha = .05/4 = .125$ ).

Chart 1. *Boxplot of Percent Positively Valenced Items by Condition\**



Post-hoc testing revealed that subjects in the UG condition produced significantly fewer positively valenced responses than did subjects in the control condition ( $t(31) = 2.535, p$  (one-tailed) = .009) and subjects in the UI condition ( $t(27) = 2.418, p$  (one-tailed) = .011), but no differences were found between subjects in the UG and DG conditions. It was expected that both of these comparisons would reach statistical significance. Subjects in the UI condition also produced significantly more positive responses than subjects in the DI condition ( $t(26) = 2.360, p$  (one-tailed) = .013). Positive valence was not expected from subjects in the UI condition; this finding will be explored in the discussion section. Table 1 shows means for each of the five conditions.

Table 1

*Positive Responses to the “Twenty Statements Task” by Condition*

Condition	N	Mean %	SD
UG	17	.41	.26
DG	15	.48	.27
UI	12	.64	.23
DI	16	.42	.24
Control	16	.62	.21
Total	76	.51	.25

A one-way ANOVA was also conducted to test the percent of responses indicating group memberships. Assumptions for homogeneity of variance and normality were met for this variable, and testing revealed that differences reached statistical significance ( $F(4,71) = 2.616, p = .042$ ).

Examination of the boxplot (Chart 2) suggested that subjects in the UI condition may have produced more individual responses than subjects in the UG, DG, or DI conditions; these comparisons were used for one-tailed post-hoc tests using the Bonferroni correction ( $\alpha = .05/3 = .017$ ). Two of these comparisons yielded statistically significant results: subjects in the UI group produced more individual responses than subjects in the UG condition ( $t(27) = 2.987, p$  (one-tailed) = .003) and subjects in the DI condition ( $t(26) = 2.626, p$  (one-tailed) = .007). The first finding—that UI would produce more individualized responses than UG—was expected; however, significant differences between the UI and DI groups were not predicted. Significant differences were expected between UI and DG groups, but these predictions did not materialize. Descriptive statistics can be found in Table 2, and the implications of these findings will be considered in the discussion section.

Chart 2

*Boxplot of Group/Individual Activation by Condition*

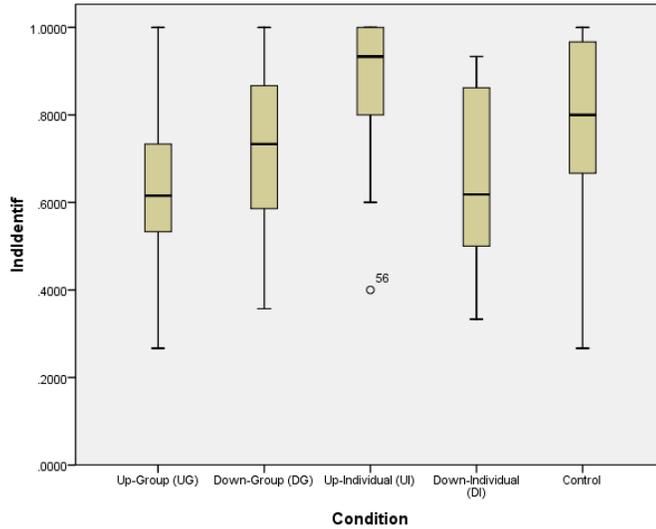


Table 2

*Group Identity Activation Responses to the “Twenty Statements Task” by Condition*

Condition	Mean %	N	SD
UG	0.37	17	0.21
DG	0.28	15	0.20
UI	0.14	12	0.19
DI	0.34	16	0.21
Control	0.23	16	0.23
Total	0.28	76	0.22

#### 4.2 Results of key hypothesis tests

The final pool subject pool included 76 subjects in the following conditions: UG ( $n = 17$ ), DG ( $n = 15$ ), UI ( $n = 12$ ), DI ( $n = 16$ ), Control ( $n = 16$ ). Of the subjects, 67.1% ( $n = 51$ ) were female and 32.9% ( $n = 25$ ) were male. Nearly all of the subjects (93.4%) were White; the remaining subjects included two who identified their race as Native

American, one as Black, one as Asian, and a non- respondent. The average age was 21.99 (Mdn = 19,  $SD = 6.06$ ), ranging in age from 18 to 52 years. Neither age nor sex varied significantly by condition, according to an Analysis of Variance (ANOVA) for age ( $F(4,71) = 1.594, p = .187$ ) and a chi square for sex ( $\chi^2(4, N=76) = 1.451, p = .835$ ).

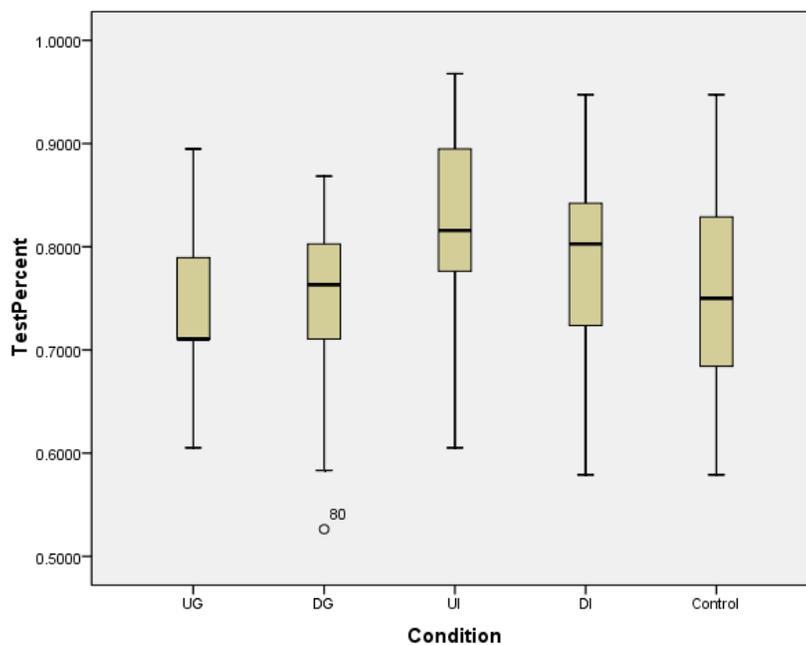
#### 4.2.1 Did test performance vary by condition?

Performance was measured through the number of items correct divided by the number attempted. Scores were normally distributed ( $M = .764, Mdn = .778, SD = .096$ ) with no notable outliers or extreme values in the dataset.

Data within subgroups were explored to determine whether they met conditions for analysis. Data in three of the conditions (UI, DI, and Control) appeared to be normally distributed with little skew. An extreme value was present in the DG condition, but there was no evidence of inconsistency or abnormalities in the response patterns of this individual, so the data were retained. Positive skew was evident in the DG condition and negative skew in the UG condition, so results were viewed with caution. (See Chart 3).

Chart 3

*Performance by Condition*



The following hypotheses were tested:

Hypotheses 1a: Performance for subjects in the upward group (UG) condition were expected to be no different from performance of subjects in the upward individual (UI) condition, allowing the two groups to be combined to form a single “Threat” group. An independent t-test did not support this prediction: subjects in the UI group performed differently from subjects in the UG group ( $t(27) = -2.463, p = .02$ ). Mean test scores by condition are presented in Table x. Cohen’s  $d$  with a pooled standard deviation was used to calculate effect size, revealing a large effect size of .90258.

Hypothesis 1b: Performance for subjects in the downward individual condition (DI) was expected to be no different from performance of subjects in the downward group (DG) condition, allowing subjects in these conditions to be combined to form a “Boost” group. This prediction was supported ( $t(29) = 1.071, p = .293$ ). See Table 3.

Table 3

*Test Performance by Condition*

Condition	Mean	N	SD
UG	.739*	17	.0755964
DG	.735	15	.0949184
UI	.820*	12	.1019450
DI	.774	16	.1045562
Control	.754	16	.1069558
Total	.761	76	.0987026

\* $p = .02$

Hypothesis 1c: Performance for subjects in a “Boost” group consisting of the two downward conditions (DI+DG) was expected to surpass performance of subjects in a “Threat” group consisting of the two upward conditions (UI+UG). Analysis revealed no significant differences between the “Threat” subjects who compared upward and “Boost”

subjects who compared downward ( $t(58) = .684, p(\text{one-tailed}) = .248$ ). Table 4 depicts performance scores for the threat, boost, and control conditions.

Hypothesis 1<sub>d</sub> and 1<sub>e</sub>: Performance of subjects in the control condition was expected to surpass the performance of subjects in the threat condition (H<sub>1d</sub>), and performance of subjects in the boost condition was expected to surpass the performance of subjects in the control condition (H<sub>1e</sub>). These hypotheses were not substantiated. The control group scored no differently than the threat group ( $t(43) = .578, p(\text{one-tailed}) = .289$ ), and no differently than the boost group ( $t(45) = .018, p(\text{one-tailed}) = .442$ ). See Table 4.

Table 4

*Performance for Subjects in Upward, Downward, and Control Conditions*

Condition	N	Mean	SD	SEM
Threat	29	.772	.095	.018
Boost	31	.755	.100	.018
Control	16	.754	.107	.027

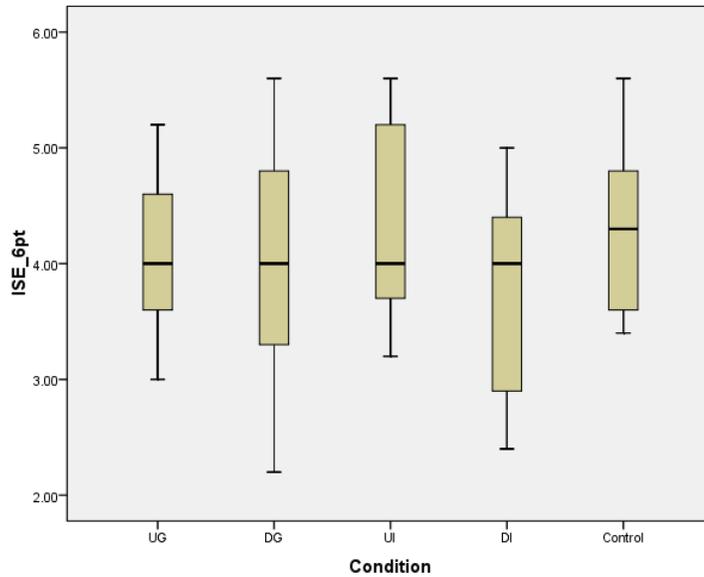
#### 4.2.2 Did individual explicit self-esteem vary by condition?

Explicit self-esteem (ESE) was measured through Likert-type responses taken immediately after the test. The data for ESE were normally distributed with very little skew and a mean of 4.095, median of 4.0, and standard deviation of .808.

ESE was also examined by condition to ensure that conditions for testing were met. In all conditions, data were normally distributed with little apparent skew. While variances did differ somewhat across conditions (See Chart 4), Levene's test of equality of variances did not reveal that these differences were significant.

Chart 4

*Explicit Self-Esteem (ESE) by Condition*



The following hypotheses were tested:

Hypothesis 2a: ESE for subjects in the DI condition was expected to surpass scores for subjects in the UI condition. This prediction was not substantiated, as can be seen in Table 5. Subjects in the DI group produced an average of 3.80, while subjects in the UI condition produced an average score of 4.30. This pattern was the opposite of what was predicted and did not reach statistical significance ( $t(26) = -1.561, p$  (one-tailed) = .935).

Table 5

*Explicit Self-Esteem (ESE) by Condition*

	Mean	N	SD
UG	4.08	17	0.64
DG	3.99	15	0.99
UI	4.30	12	0.84
DI	3.80	16	0.84

	Mean	N	SD
Control	4.35	16	0.69
Total	4.09	76	0.81

Hypotheses 2b and 2c: ESE for subjects in the DI condition was expected to surpass scores for subjects in the control condition and scores for subjects in the control condition were expected to surpass scores for subjects in the UI condition. Neither of these predictions materialized, and the patterns were the opposite of what was expected. (H<sub>2b</sub>: DI=Control;  $t(30) = -2.020$ ,  $p(\text{one-tailed - opposite direction}) = .948$ . H<sub>2c</sub>: UI = Control;  $t(26) = -.335$ ,  $p(\text{one-tailed - opposite direction}) = .259$ ).

Hypotheses 2d, 2e, and 2f: ESE for subjects in the UG condition, the DG condition, and the control condition were expected to have no significant differences. Each of these hypotheses was supported: subjects in the DG condition received ESE levels no different from levels for subjects in the UG condition ( $t(30) = .329$ ,  $p = .745$ ); subjects in the UG condition reported ESE levels no different from subjects in the control condition ( $t(31) = -1.155$ ,  $p = .257$ ), and subjects in the DG condition reported ESE levels no different from subjects in the control condition ( $t(29) = -1.187$ ,  $p = .245$ ).

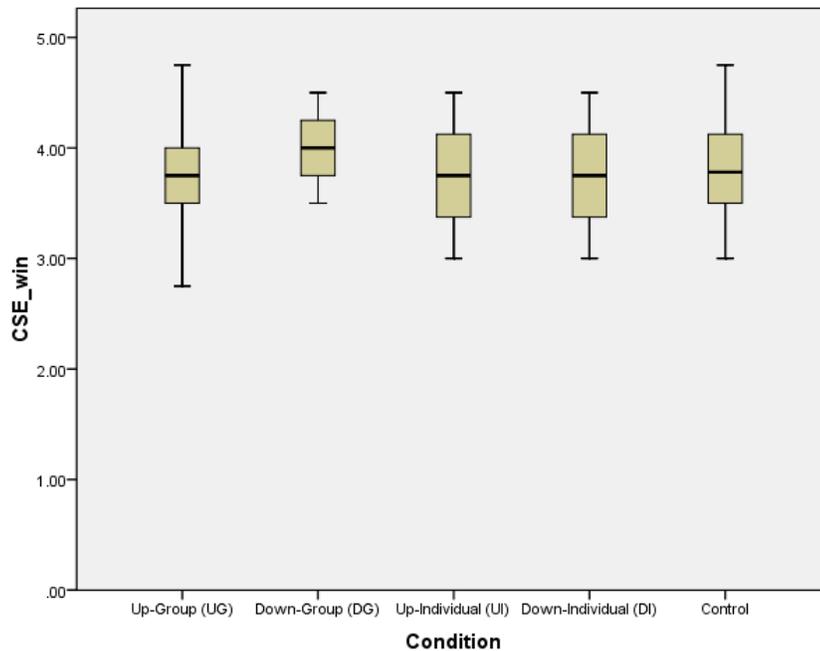
#### 4.3.3 Did collective self-esteem vary by condition?

Collective self-esteem (CSE) was measured through an average of Likert-type responses to four statements taken immediately after the test. The data for CSE were normally distributed with a mean of 3.71, median of 3.75, and standard deviation of .643. Data were skewed negatively and two outliers were present at the lower end of the continuum. In order to ensure that the means and standard deviations reflected the population most accurately, data were Winsorized so that the extreme values did not unduly influence them. Using a standard formula for identifying outliers ( $1.5 \times \text{IQR}$ ), two outliers were identified and were transformed to the lower bound value of 3.5636. This transformation produced a Winsorized mean of 3.79 ( $SD = 0.49$ ).

CSE was also examined by condition to ensure that conditions for testing were met. Using the Winsorized CSE responses (CSEwin), data were found to be normally distributed within each subgroup (See Chart 5).

Chart 5

*Collective self-esteem (CSE) by Condition*



The following hypotheses related to collective self-esteem were tested:

Hypothesis 3a: CSEwin for subjects in the DG condition was expected to surpass scores for subjects in the UG condition. This hypothesis was validated; subjects in the DG condition reported average CSE ratings of 4.02 (SD=.826), and subjects in the UG condition reported average CSE ratings of 3.54 (SD=.320). Using a one-tailed t-test, CSEwin for the DG condition was significantly greater than CSEwin for the UG condition ( $t(25.549) = -2.07$ ,  $p(\text{one-tailed})=.025$ ). The effect size was moderate, however: Cohen's  $d$  using a pooled standard deviation was equal to 0.721. See Table 6.

Table 6

*Collective Self-Esteem (Winsorized) by condition*

Condition	Mean	N	SD
UG	3.6802	17	.57700
DG	4.0167	15	.31997
UI	3.7500	12	.46466
DI	3.7227	16	.46263
Control	3.8048	16	.53752
Total	3.7928	76	.48648

\* $P < .04$ 

Hypotheses 3b and 3c: CSEwin for subjects in the DG condition was expected to surpass scores for subjects in the control condition, and CSEwin for subjects in the control condition was expected to surpass scores for subjects in the UG condition. Alpha in each test was set at .02. These predictions were not supported by the data. Subjects in the UG condition reported average CSEwin levels no different from those in the control condition ( $t(31) = -.478, p(\text{one-tailed}) = .318$ ). Likewise, subjects in the DG condition reported CSEwin scores no different from those in the control condition ( $t(24.701) = -1.322, p(\text{one-tailed}) = .091$ ).

Hypotheses 3d, 3e, and 3f: CSEwin for subjects in the UI condition was expected to be no different from CSEwin for subjects in the DI condition and or from subjects in the control condition. CSEwin for subjects in the control group was also not expected to differ from CSEwin for subjects in the DI condition.

All of these predictions were supported by the data. CSEwin for subjects in the UI condition was not found to be different from CSEwin for subjects in the DI condition ( $t(26) = -.174, p = .879$ ). CSEwin for subjects in the UI condition was no different from CSEwin for subjects in the control condition ( $t(26) = .283, p = .780$ ). Finally, CSEwin for

subjects in the DI condition found not to be different from CSEwin for subjects in the control condition ( $t(30) = -.643, p = .647$ ).

#### 4.3.4 Implicit individual self-esteem (ISE)

Paired-sample t-tests were used to compare signature area before and after the threat for subjects in each condition. Across conditions, signature sizes for each condition were selected, examined for normality and skew, and then tested for statistical differences using  $\alpha = .05$  in one-tailed test according to the direction of the hypotheses. The following hypotheses were then tested:

Hypothesis 4a: Signature size was expected to decrease significantly between Signature 1 (Sig1) and Signature 2 (Sig2) for subjects in the threat condition (UG+UI). No predictions were made for the differences between Sig2 and Signature 3 (Sig3). The increase in size between Signature 1 and Signature 2 did not materialize. Sig1 were no different from Sig2 for subjects in the threat conditions. Additionally, there were no differences between Sig2 and Sig3 for these subjects. Alpha was set to .05 for a one-tailed test. See *Table 7*.

Table 7

*Signature Sizes for Subjects in the Threat Conditions*

		Mean	SD	SEM	t-tests
Pair 1	Sig1	.5428	.32073	.05956	
	Sig2	.5623	.26358	.04895	$t(29) 0.429, p = .336$
Pair 2	Sig2	.5717	.26345	.04979	
	Sig3	.5803	.31681	.05987	$t(27) = 0.212, p = .834$

Hypothesis 4b: Signature size was expected to increase significantly between Sig1 and Sig2 for subjects in the Boost conditions (DI+DG). No predictions were made for the differences between Sig2 and Sig3; these differences will be examined to determine the duration of the effect. Alpha was set to .05 for a one-tailed test.

The hypothesis predicting an increase in signature size for subjects in boost conditions between the first and second measures was supported in the data analysis. Specifically, signature sizes increased from .47 to .51 units between the first and second measures ( $t(30) = -1.814$ ,  $p(\text{one-tailed}) = .038$ ; Table 8.). The effect size was small (Cohen's  $d = .17406$ ). Although no hypothesis was formed, signature sizes then decreased significantly after reading the test instructions, declining from .520 to .435 units between the second and third measure ( $t(29) = 3.338$ ,  $p=.001$ ) (The degrees of freedom were reduced for the second analysis because one subject's signature was unusable). Interpretation of this finding will be explored in the discussion section.

Table 8

*Signature Sizes for Subjects in the Boost Conditions*

		Mean	SD	SEM	t-tests
Pair 1	Sig1	.4666	.20589	.03698	
	Sig2	.5069	.25461	.04573	$t(30) = -1.814$ , $p(\text{one-tailed}) = .038$
Pair 2	Sig2	.5204	.24751	.04519	
	Sig3	.4353	.20976	.03830	$t(29) = 3.338$ , $p(\text{one-tailed})=.001$

Hypotheses 4c: Signature size was not expected to change significantly between Sig1 and Sig2 for subjects in the control group, and as before, no predictions were made for the differences between Sig2 and Sig3. These predictions were substantiated through the data, in which no significant differences were found between Sig1 and Sig2 nor between Sig2 and Sig3. See Table 9.

Table 9

*Signature Sizes for Sig1 vs. Sig2 and Sig2 vs. Sig3 for Subjects in the Control Condition*

	Mean	SD	SEM	t-tests
--	------	----	-----	---------

Pair 1	SigArea1	.5882	.40208	.10052	
	SigArea2	.5244	.27758	.06940	$t(16) = 1.092, p = .292$
Pair 2	SigArea2	.5244	.27758	.06940	
	SigArea3	.5729	.38944	.09736	$t(16) = -0.752, p = .463$

#### 4.3.5 Moderators

Three factors were presented for exploratory analysis: social comparison orientation, domain identification, and group identification. Each of these factors was expected to moderate the relationship between the threat condition and test performance.

*4.3.5.1 Social Comparison Orientation.* Subjects high in social comparison orientation (high-SCO) were expected to produce more pronounced differences in each of the dependent measures than subjects low in SCO (low-SCO). That is, subjects who tend to make more comparisons between themselves and others were expected to be more susceptible to the threat conditions invoked in this experiment. This susceptibility was expected, in turn, to produce greater differences in test performance across conditions than were produced for low-SCO subjects or for an aggregate of high- and low-SCO subjects.

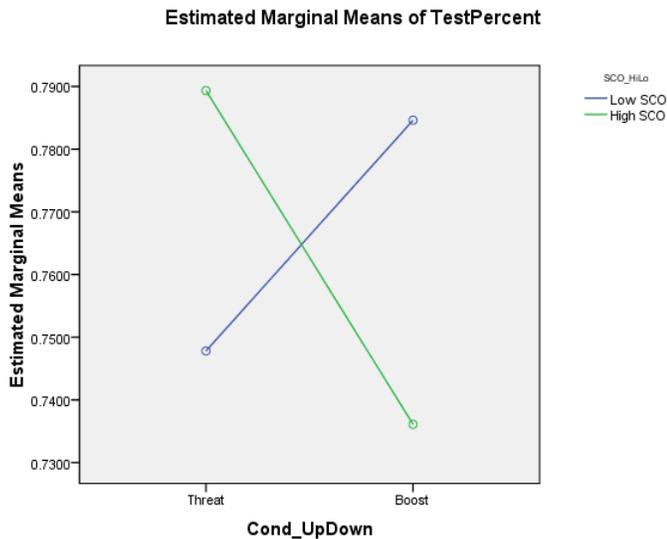
To test this hypothesis, subjects were divided into two groups: those whose SCO was equal to or greater than the group mean of 3.95 (high-SCO) and those whose average SCO scores were less than the group mean (low-SCO). In order to maintain cell sizes sufficient for analysis, the five conditions were divided into three “threat groups” in the same manner used for signature size: boost (DG+DI), threat (UG+UI). Because of small cell sizes, the control group was not included in the analysis. Following protocol used by Cadinu and colleagues (2006), a 2 (SCO: high vs. low) x 2 (Condition: threat vs. boost) ANOVA was conducted to determine whether there was an interaction effect between threat group and social comparison orientation. To correct for unequal cell sizes, Type III sums of squares were used. Tests were done only for the performance variable, as this was the only variable for which subjects could be grouped to provide adequately large

cell sizes; both CSE and ISE rested on comparisons among disaggregated conditions, producing small cell sizes.

Both visual inspection and statistical tests suggested that performance data for both SCO and condition met assumptions for normality and homogeneity of variance. Results of the ANOVA revealed that no main effects were found for performance for either SCO ( $F(1,59) = .019, p = .892$ ) or condition ( $F(1,59) = .104, p = .749$ ). Additionally, interaction effects did not prove to be statistically significant ( $F(1,59) = 3.115, p = .083$ ). See Chart 6.

Chart 6

*Test Percent as a Function of Threat Condition and SCO*



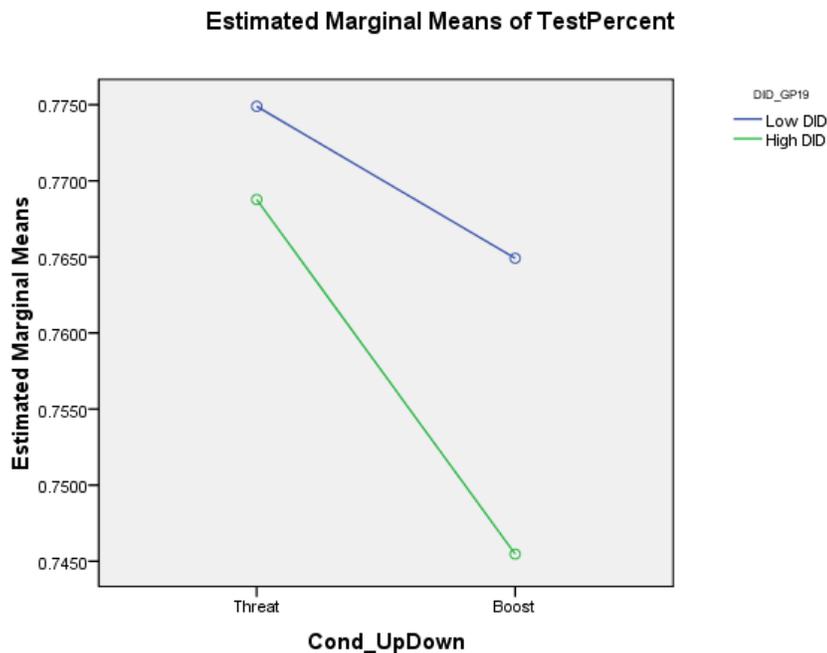
*4.3.5.2 Domain Identification.* Subjects high in student domain identification (DID) were expected to produce more pronounced differences in each of the dependent measures than subjects low in DID. That is, subjects whose identities were strongly immersed in their role as students (high-DID) were expected to experience greater responses to the threat and boost conditions than subjects who did not strongly identify with their student role (low-DID). In terms of this study, high-DID subjects were expected to produce better test scores than low-DID subjects in boost conditions (DI and

DG), and to perform more poorly than low-DID subjects in threat conditions (UI and UG).

To test this hypothesis, subjects were grouped into two groups: those whose average DID scores were equal to or greater than the group mean of 3.8 (high-DID) and those whose average DID scores were less than the group mean (low-DID). Performance for the DID groups met the assumptions for ANOVA analysis: data were normally distributed with homogeneity of variance. As with SCO, a 2 (DID: high vs low) x 2 (Condition: threat vs. boost) ANOVA was conducted to identify an interaction between DID and condition. Results revealed no significant main effects for either DID ( $F(1,59) = 2.48, p = .621$ ) or condition ( $F(1,59) = .420, p = .520$ ). No interaction between condition and DID was identified ( $F(1,59) = .067, p = .796$ ). See Chart 7.

Chart 7

*Test Performance as a Function of Condition and DID*

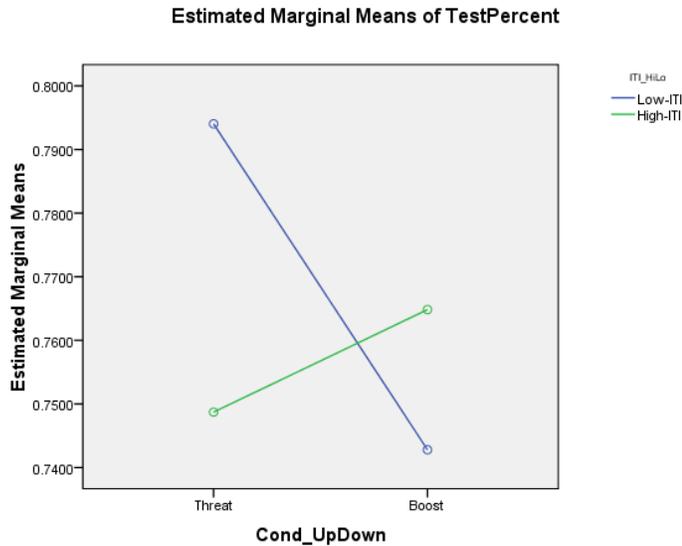


4.3.5.2 *Importance to Identity (ITI)*. ITI refers to the extent to which subjects include their role as community college students in their personal identities. Subjects who identified more strongly as community college students were expected to receive a greater threat or boost from conditions that activated their community college student identities. In terms of this study, subjects who identified strongly with their role as community college students (high-ITI) and were in the group comparison conditions (UG or DG) were expected to produce more extreme performance scores than subjects who did not identify strongly with their role as community college students (low-ITI). No performance differences between high- and low-ITI subjects were predicted in the individual or control conditions.

To test this hypothesis, subjects were grouped into two groups: those whose average ITI scores were equal to or greater than the group mean of 2.9 (high- ITI) and those whose average ITI scores were less than the group mean (low- ITI). Performance data within DID levels proved to be normally distributed with homogeneity of variance, thereby meeting assumptions for ANOVA. To determine whether there was an interaction between condition and DID, a 2 (condition: threat vs. boost) by 2 (DID: high vs. low) ANOVA was conducted. Results revealed no main effect for condition ( $F(1,59) = .481, p = .491$ ) or for DID ( $F(1,59) = .212, p = .647$ ). No interaction effects were identified ( $F(1,59) = 1.769, p = .189$ ). See Chart 8.

Chart 8

*Test Performance as a Function of Condition and ITI*



### 4.3 Ancillary Analyses

Several additional tests were conducted following the initial analyses to further understand unexpected findings. These tests addressed two dimensions that were ambiguous following the primary analyses: (a) unexpected findings for the UI group, and (b) nuances of the collective identity variables. Both explorations shed light on the overall results.

#### *4.3.1 Unexpected Findings for the Upward-Individual Group*

Contrary to expectations, subjects in the upward individual condition appeared to receive higher test scores on the dependent performance measure than subjects in the other conditions. An additional t-test was conducted, therefore, to determine whether this observation in the data proved to be significant: test scores for subjects in the UI group were compared with scores for subjects in all of the other conditions using a one-tailed t-

test. Because the additional test was not accounted for in the original Bonferroni correction, alpha for the additional tests was set at .02 for a total alpha level of .12 across all of the analyses using test percent as dependent variable. The t-test produced significant results, as shown in *Table 10*: subjects in the UI condition outperformed subjects in the other conditions ( $t(74) = -2.213$ ,  $p(\text{one-tailed}) = .013$ ). The effect size was large with Cohen's  $d$  producing a value of .713691. Thus, subjects who read the UI prime outperformed subjects in other conditions rather than experiencing a performance decrement due to the intended threat.

Table 10

*Test Percent for UI and All Other Conditions*

	N	Mean	SD	SEM
UI condition	12	.820*	.102	.029
All other conditions	64	.750*	.094	.019

\* $p < .02$

Though this unexpected significant result may be explained by a contradictory response to the prime, further examination suggests that that the difference may also be partially attributed to a variation in general aptitude. Specifically, self-reported ACT scores, which were provided by about 67% of the subjects, were also significantly higher for subjects in the UI condition than in the other conditions ( $t(53) = -1.962$ ,  $p(\text{one-tailed}) = .027$ ). A bivariate correlation revealed a moderate relationship between self-reported ACT scores and scores on the dependent test score measure (See Table 11;  $r(53) = .48$ ,  $p < .001$ ), therein providing additional support to the suggestion that a more general difference in aptitude might have produced the anomalous finding.

Table 11

*ACT scores and Test Scores for UI condition and all other conditions*

	N	Mean	SD	SEM
UI condition	10	24.00*	3.651	1.155
Other conditions	45	21.69*	3.31	.493

\* $p < .025$

*4.3.2 Nuances of the collective identity variables*

The extent to which subjects identified with their roles as community college students and as students in general served as predictor variables in this study, and the extent to which they esteemed this role served as a dependent variable. Disparities within and among these measures provoked the need for additional exploration.

The original research design utilized a single measure of collective self-esteem as a dependent variable. Over the course of analysis, it became evident that differences might exist between the factors identified by Luhtanen and Crocker (1992) as “public collective self-esteem” and “private collective self-esteem.” With marked regularity, subjects gave apparently lower ratings to the following public collective self-esteem items: “#4-Most people consider community college students to be more effective than students from other schools” and “#17-In general, others think that community college students are unworthy”\* (Combined  $M = 2.05$ ,  $SD = .533$ ). than to the private self-esteem items “#5-I often regret that I am a community college student”\* and “#13-In general, I’m glad to be a community college student” ( $M = 4.54$ ,  $SD = 1.06$ ).

In addition to differences within the CSE variable, differences among private and public CSE, Importance to Identity (ITI), and domain identification (DID) warranted

---

\*Reverse coded

exploration. The ITI variable was based on averaged responses to two Likert-scaled items collected at the same time as the other self-esteem items: “Overall, the fact that I am a community college student has very little to do with how I feel about myself” and “Being a community college student is an important reflection of who I am” (Luhtanen & Crocker, year). The ITI variable measured the extent to which subjects’ identities incorporated their role as community college students. The DID variable consisted of responses to three Likert-scaled statements also collected at the same time: “Being a student is important to me,” “My academic performance means a lot to me,” and “I don’t really care that much about how well I do in college (reverse-coded).” These items were used by Johnson and Stapel (2005) to determine the extent to which subjects valued their roles as students. Mean responses for each factor are presented in Table 12.

Table 12

*Mean ratings for Public CSE, Private CSE, ITI, and DID*

	Mean	N	SD
Public CSE	2.05	76	0.53
Private CSE	4.54	76	1.06
Domain Identification (DID)	4.64	76	0.63
Importance to Identity (ITI)	2.92	76	1.07

Paired t-tests were conducted to determine whether subjects’ responses differed across these four variables. Because each variable was used for three different paired t-tests, the Bonferroni correction was used to set alpha to  $.05 / 3 = .017$ . Tests revealed significant differences between each of the pairs in question with only one exception, as shown in Table 13. The non-significant comparison was ITI versus private CSE, which showed no significant difference ( $t(75) = .785, p = .437$ ). Thus, students appeared to

view their own role as community college students favorably, while at the same time recognizing that community college students in general were not as favorably perceived by others. The implications of these different findings to the overall conclusions of the study will be addressed at detail in the discussion section.

Table 13

*Paired t-test comparisons of Public CSE, Private CSE, ITI, and DID*

Pairs	t	df	p
Private CSE - Public CSE	19.738	75	.000
Private CSE - ITI	10.932	75	.000
Private CSE - DID	-.782	75	.437
Public CSE - ITI	-6.439	75	.000
Public CSE - DID	-26.987	75	.000
DID – ITI	-14.722	75	.000

\*\* $p < .001$

## CHAPTER 5. DISCUSSION

This study sought to demonstrate the parallel nature of stereotype threat and threat derived from individual comparisons by analyzing performance and several types of self-esteem across randomly assigned conditions. The stereotype that was utilized in the study related to academic underperformance of community college students. Both the stereotype and individual conditions were activated by assigning subjects to read a fictitious newspaper article about a superior/inferior group or individual for the ostensible purpose of determining whether it was Internet- or paper-based. It was hypothesized that subjects who compared upward, regardless of whether the comparison was at the group or individual level, would receive lower scores on a performance measure than subjects who compared downward or control subjects. It was further hypothesized that subjects would experience contrasting differences in self-esteem at the collective or individual level depending on whether they read a news article about community college students or individual comparison targets. Further, it was predicted that implicit self-esteem would differ in both the group and individual conditions in an opposite direction to the original prime. Finally, it was hypothesized that subjects high in social comparison orientation would experience these effects more markedly than subjects with low social comparison orientation.

Some of these hypotheses were supported with evidence from the study; these will be detailed and explored in subsequent paragraphs. More of the findings, however, did not support the hypotheses. Test performance did not vary by condition in the expected direction, and the pattern that emerged in the test score data contradicted predictions in some instances. Individual self-esteem did not vary by condition, and collective self-esteem did not reveal the expected contrast effects.

The findings that were significant lend insight to the processes involved, and will be explored in more detail in the first section of this discussion. The absence of the full spectrum of expected findings, however, warrants equal consideration and explanation as elucidators of the process being explored. This will be the focus of the second section of this discussion. Finally, future directions in exploring the original research question will be considered.

## 5.1 Significant and Non-Significant Findings

A number of statistically significant comparisons were found during analysis related to test scores, collective self-esteem, implicit self-esteem, and identity activation. Specifically, the following analyses proved to be significant:

1. Performance: Subjects in the UI condition significantly outperformed the subjects in the UG condition and both group conditions on the test used as the dependent measure. No other findings related to this test were significant. The findings may be related to self-reported ACT scores, which correlated moderately with scores on the performance measure and were significantly higher for subjects in the UI condition than for those not in the UI condition. Only about 67% of subjects provided these data, however, so these results are only suggestive of a pattern.
2. Collective self-esteem (CSE): As predicted, the downward comparison group (DG) reported higher levels of collective self-esteem than subjects in the UG condition.
3. Self-Esteem: Self-reports of explicit self esteem (ESE): did not differ across groups. Implicit self-esteem, measured through signature size at different points during the study, increased for subjects in Boost conditions immediately after the prime, and decreased for the same subjects at the start of the test. No decrease in signature size was found for the Threat group.
4. Twenty Statements Task: In response to the question “Who am I?” subjects in the UI condition produced a significantly higher proportion of positively valenced self-descriptors than did subjects in the UG or DI conditions. Subjects in the UG conditions produced significantly fewer of these responses than subjects in the control condition. Further, subjects in the UI condition provided significantly fewer group identifiers than did students in the UG and DI groups.
5. Collective identity: Subjects’ public collective self-esteem (CSE; ratings of others’ positive perceptions of community college students) was significantly lower than (a) private CSE (their ratings of their own role as community college students), (b) importance to identity (ITI; the extent to which subjects

identified with their role as community college students), and (c) domain identification (DID; the extent to which subjects valued their role as students in general). Additionally, ITI responses were significantly lower than private CSE and DID.

These significant findings will be explored in the following paragraphs, beginning with the unexpected findings related to the UI group, the absence of expected stereotype threat effects, and the implications of the signature size measures.

#### *5.1.2 Unexpected Findings in the Upward-Individual (UI) condition*

Intertwined hypotheses mark the expected key findings in this study. First, it was hypothesized that subjects in the UI and UG group would not score differently on a performance measure because they were both influenced by an upward comparison (albeit at different levels of identity), and that subjects in the DI and DG group would not perform differently because both would experience a performance boost due to a downward comparison. Resting on the assumption of non-significant findings in these first hypotheses, data from the UI and UG groups were combined to form a more general “Threat” group. Likewise, data from the DG and DI groups were combined to form a more general “Boost” group. By combining these groups, the key hypotheses could be tested: subjects in the boost condition (DI+DG) were expected to outperform subjects in the Threat condition (UI+UG), and both were expected to differ significantly from subjects in the control group (Boost > Control > Threat).

Unfortunately, one of the foundational premises—the expectation of similar test scores between the UG and UI group—was not supported. In fact, subjects in the UI group outperformed subjects in the UG group, even at the relatively conservative alpha level stipulated for this comparison. Further analysis revealed that, in fact, subjects in the UI group significantly outperformed subjects in both of the group conditions. The high-performance of the UI group was not predicted; in fact, subjects in this group were expected to suffer a performance decrement because of the threat of an upward interpersonal comparison.

This “overperformance” of the UI group led to several unintended consequences. Most importantly, it made the combination of UI and UG groups into a single “Threat” group, as indicated in the study design, an inappropriate decision. By combining the UI

and UG groups with their significantly different test scores, the performance of one group (e.g., UI) essentially nullified the performance of the second group (e.g., UG) and relegated any tests using this combined group non-significant. This did indeed prove to be the case: the Threat and Boost groups revealed no significant differences in test scores. Further exploration revealed, in fact, no differences between any of the other conditions.

Thus an important question remains: What could explain the unexpected findings for the UI group? One plausible explanation involves chance differences in the intellectual capacity of subjects in the UI group. Subjects in this condition reported significantly higher ACT scores in a self-report measure taken at the onset of the study than did subjects in the “group” comparison conditions. The relationship between self-reported ACT scores and test scores was established through a moderate correlation, so a connection between ACT and test percent must not be ignored. Thus, a fluke of random assignment may have produced a higher-performing group in the UI condition, resulting in higher test scores as well. The absence of data for one-third of the subjects prevents formal inclusion of the ACT variable, but nevertheless hints at the possibility of a pre-existing group difference.

Other differences between the UI group and other groups, however, suggest that the test score differences may have been related to the prime itself rather than solely to chance intergroup differences. A key component of this alternate explanation can be found in the newspaper article designed to activate the comparison. Subjects in the UI condition were asked to read a news story about a brilliant local student, Casey, who had been accepted into Harvard University but chose to go attend community college to save money. The article then continued to describe the student’s outstanding performance, leadership role, and popularity at the community college. The intention of the article was to cause readers to compare themselves unfavorably with Casey and suffer temporary reduction in performance.

In reality, the article may have produced the opposite effect. By describing a premiere student who intentionally chose the same community college as did all of the study participants, the article may have had iatrogenic effects. First, the decision of the outstanding student to attend community college many have validated their own identical decision and eliminated any potential self-doubt related to their role as community

college students. Rather than comparing upward with the outstanding student and experiencing a decrement, the subjects may have identified with the outstanding student and experienced a boost. This effect may have been heightened if students feel that they themselves might have been able to attend college elsewhere were it not for extenuating circumstances, which is often the case in community colleges. Similar boost effects from exposure to superior representatives of a poorly-perceived social group have been found in other studies. Black students in a threatening situation who were exposed to a high-performing Black classmate, for example, outperformed Black subjects who were exposed to a low-performing Black classmate and those who were threatened by exposure to a high-performing White classmate (Blanton, Crocker, & Miller, 2000).

This explanation, which suggests that the UI prime prompted differences in test scores because subjects identified rather than compared with the subject of the news article, has some other supporting evidence. In responding to statements in the Twenty Statements Task (TST), for example, subjects in the UI condition used significantly more positively valenced self-descriptors than subjects in the group conditions. Further, subjects in the UI condition produced significantly more individual identifiers than subjects in the UG and DI conditions. Although the TST is a projective measure and was intended as only a manipulation check, these results items suggest that the subjects in this condition were feeling particularly positive as individuals at the time the measures were taken. Whether the positive individuality was a consequence of the newspaper primes or of a good test performance cannot be determined, but the difference in the group seems fairly clear.

Though the UI group appears to have experienced increased individualization and positive feelings, this did not translate into higher scores on the Likert-scaled self-esteem statements taken immediately after the test. Though it was hypothesized that the UI group would report higher individual self-esteem than subjects in the DI group, these results did not materialize. Likewise, if the prime did in fact boost subjects' perceptions of the group rather than diminish perceptions of the individual, higher collective self-esteem might have been observed. This also did not prove to be the case. The ESE and CSE results for the UI group were not significantly different from results for any of the other groups.

It might therefore be hypothesized that the UI condition, rather than prompting a deficit in esteem as a consequence of upward comparison, instead produced a boost in implicit self-esteem. Both the inadvertent boost in implicit self-esteem and differences in innate ability (as reflected in self-reported ACT scores) may have contributed to the enhanced test scores for subjects in this condition. While the accuracy of this hypothesis cannot be known, evidence explored in the preceding paragraphs suggests the possibility that the boost in collective self-esteem may have occurred as a consequence of self-validation. This boost may then have produced test score gains, a more positive individual self-concept, and less likelihood of identifying with a group that might otherwise hold them down.

#### *5.1.2 Absence of stereotype threat effects.*

Performance did not differ between subjects in the UG and DG conditions, indicating that a stereotype threat was not produced in this study. This absence may have occurred for several reasons. First, the stereotype itself may have been inappropriate for prompting a threat effect. Second, subjects may not have been aware of the stereotype. Finally, subjects may have distanced themselves from the stereotype or considered it to be inapplicable to them so that it did not influence their performance. Each of these possibilities will be discussed in the succeeding paragraphs.

Throughout the literature, stereotypes that are merged with identity have proven to be the most powerful in eliciting stereotype threat. According to Haslam and Turner (1992), groups that are impermeable (cannot easily be entered or exited) and whose status is stable (unlikely to improve in others' views) bear the greatest capacity for producing threat effects on its members. Thus, ascribed groups that cannot be changed or hidden (e.g., race or gender) provide the greatest capacity for threat. Achieved groups (self-chosen, such as skateboarders, or rednecks) and temporary groups (e.g., teens, pregnant women) can experience threat from others' expectations of them, the threat is minimized because the individual can choose or anticipate an exit from the group.

The stereotype used in this study revolved around community college students. Although such stereotypes were shown to exist in the study by Caporrimo (2008), the mere existence of the stereotype does not automatically produce the conditions for stereotype threat. First, community college students represent a highly permeable group.

The very nature of a two-year college is temporary; the phrase “See how your credits transfer” highlighted the campus homepage, and cheaper tuition with guaranteed transfer of credits serves as the primary reason students attend the college. Because nearly all of the students intended to transfer to a four-year university, it is not likely that they have their identity enmeshed with the stereotypes about community college students. Thus, despite the fact that the stereotype exists and negatively portrays such students, it appears not to wield sufficient power to impact the performance of the students. The absence of significant results, therefore, supports the suggestions of Steele & Aronson (1995) regarding the characteristics of stereotypes that lend themselves to threat.

Further evidence of the inappropriateness of the stereotype derives from the study itself, which provided mixed results regarding student identification with the stereotype. Data from the public collective self-esteem variable (in which subjects responded to statements about the regard others held for community college students) reached an average of only 2.05 on a 6-point scale. In fact, subjects in the control and DI groups (who were not stereotype-activated) produced responses that averaged 1.96—no different from subjects in the UG condition who had read negative information about community college students. Subjects in the DG and UI conditions, who read articles that may have created a positive impression of community college students, still produced responses averaging only 2.27 and 2.21, respectively. Thus, students across conditions were aware of the stereotype, and even reading a favorable article about the stereotyped group did not create a positive view of others’ perceptions of the group.

Students’ ratings of their own role as community college students were considerably higher, as reflected in the private CSE variable. Responding to statements about whether they regretted or were glad to be community college students, subjects gave average ratings of 4.54—more than double their ratings for public CSE. This suggests that there is some degree of separation between subjects’ understanding of the stereotype and relation of the stereotype to themselves. If they strongly identified with the group that was poorly perceived by others (i.e., community college students), then they would be most likely to disidentify from the group and express disdain at being a member. Given that many students attending community college are “second chance” students (students who have overcome past struggles) or students constrained by location,

however, they may feel glad to have this opportunity to attend college even if the college is not well-regarded. Further, they may see themselves as the exceptions, or as students who could have done well elsewhere if not for certain interfering events. Either way, responses from these students suggested that the power of the stereotype over their own self-perceptions were minimal.

The Importance to Identity (ITI) measure further supports this suggestion. ITI is a variable consisting of two Likert-scaled responses to statements about the extent to which subjects felt that their role as community college students was a “reflection of the self.” While the average rating for private CSE was 4.54, the average rating for ITI was only 2.92. Thus, while many students were glad to be community college students, it appeared that their sense of identity did not incorporate their membership as a community college student. The temporary nature of community college attendance may explain this relatively low identification as community college students.

Finally, the importance subjects placed on their roles as student (domain identification, or DID) was significantly higher than the extent to which they identified with their role as community college students (ITI). Again, students valued their role as student and were glad to be attending community college, but did not identify specifically with their role as community college students.

The disparities among private CSE, public CSE, ITI, and DID further help to explain the significant difference in CSE (overall) found between subjects in the DG and UG conditions. This difference was expected—but was expected in conjunction with a parallel difference in test scores. The fact that the CSE difference occurred in the absence of test score differences would threaten the veracity of the overall hypothesis, were it not for the disparities in the other indicators of identity. If subjects in the group conditions recognized that other people perceived community college students poorly, (as indicated in the public CSE variable), but did not respond emotionally to this knowledge because they did not relate the general perception to themselves (as suggested in the private CSE and ITI variables), then no decrement in test performance could occur.

### *5.1.3 Implicit Self-Esteem*

Finally, findings related to implicit self-esteem are examined. Measured through signature size, this measure is not commonly used but nevertheless provides some

interesting direction. It was expected, as reported earlier, that subjects in the boost groups would produce larger signature sizes after reading about an underperforming group or individual. This was demonstrated in analysis: The signature taken after reading the article primes was significantly larger than the initial signature for subjects in the boost conditions. This suggests that, if the test is a proxy for implicit self-esteem, the subjects in the boost conditions experienced an increase in self-esteem after the article. The signature size then decreased after subjects read the test instructions, suggesting that the inflated self-esteem was short lived.

Subjects in the threat conditions did not experience a similar decrease in signature size following the reading of the article. At least part of the effect may be due to the anomalies within the UI group described earlier. Although the planned comparisons outlined in the proposal demanded that the UI and UG conditions be grouped together, this grouping may have been inappropriate in light of the differences found across groups. It was also suggested by Zweigenhaft and colleagues that these signature size changes were more likely to respond to a boost in self-esteem (increased signature size) than to a decrement to self esteem through smaller signature size.

## 5.2 Implications of the findings

The notable differences between the hypotheses and the findings were explored in the preceding paragraphs. Two general explanations could account for these failings: (a) the hypotheses could have been incorrect, and/or (b) the methods may have been inappropriate. These possibilities will be explored in the following sections.

### *5.2.1 Were the hypotheses incorrect?*

This study hypothesized the existence of a phenomenon of identity relativity through which activation of a particular group or individual identity prompts comparisons at that level. Comparisons at any level will prompt shifts in self-esteem at that level, which may or may not conflict with other self-views. Such shifts may then create dissonance, particularly if the comparison domain is valued by the comparer; accordingly, the physiological arousal accompanying dissonance may prove detrimental to a performance of any sort that takes place before consonance is reached.

In order to demonstrate the veracity of the hypotheses, several linkages must be established in existing research. First, the coexistence of collective and individual

identities within individuals must be established. Second, it must be demonstrated that different aspects of a single identity can be situationally activated. Third, activation of different levels of identity must be shown to elicit different behaviors and performance. Finally, individuals who have the greatest self-uncertainty must also have the greatest susceptibility to comparison effects. Abundant evidence supports each of these premises, as illustrated in the following paragraphs:

*Coexistence of Collective and Individual Identities.*

The work of Gordon Allport (1954), Henri Tajfel (1970), and John Turner (Turner & Tajfel, 1982) provides the framework through which identity is shown to be multi-leveled. Gordon Allport (1954) described the multi-faceted nature of identity in a powerful exploration of prejudice, depicting an individual's identity as concentric circles representing different levels of in-group/out-group activation. Following this lead, Henri Tajfel (1970) showed in a series of experiments that in-group identification can be prompted without conflict and eventually developed social identity theory, in which he showed that characteristics of a group can become linked with characteristics of the individual. Through self-categorization theory, Tajfel and Turner (1982) suggested that people will identify with the most positive of social identities, and that ambiguous social identities may prompt one to elevate the group identity through denigration of the out-group. Each of these theories was derived through synthesis of research; thus, we can conclude that such multiple levels of identity do indeed exist.

*Identities can be Situationally Activated.*

This phenomenon that identities can be situationally activated has been experimentally demonstrated through manipulation checks in social comparison experiments. Specifically, an experimenter might use a particular technique to activate an individual or group identity, and then verify through a prime check that the intended identity was indeed activated. Several studies illustrate this phenomenon. For example, Brewer and Gardner (1996) activated individual identities through a story in which "we" or "I" pronouns were used; those who were activated in the "we" condition used more collective descriptors in the Twenty Statements Task (completing the phrase "I am..."). Stapel and Blanton (2004) found that subjects who viewed pictures of happy or sad faces chose more personal pronouns on a manipulation check than did those who viewed

pictures of a tree, indicating individual activation. Further, they tended to report higher or lower ratings of their own happiness depending on whether they had viewed a happy or sad face. Other experiments have revealed similar conclusions, suggesting that different identities can indeed be activated through situational stimuli.

*Different Identities Produce Different Behaviors.*

The experimental support for these conceptual foundations is provided in myriad studies. First, Brewer and Gardner (1996) demonstrated that opinions changed in response to identity activation and comparison. Subjects who were activated in an individual identity rated their opinions as less similar to a series of statements than did subjects activated in a group identity. Second, expectations for success changed. Finally, self-perceptions changed. Onorato (2004) found that women who were activated in an individual identity endorsed more words related to independence (a less feminine trait) than did women who were activated in a collective female identity.

In addition to self-perceptions and expectations, however, performance itself changed in response to identity activations and comparisons. In some studies, identity activation was directly manipulated by the researcher. For example, Gordijn and Stapel (2006) found that subjects who were deliberately activated in an individual condition before reading about a successful peer performed more poorly on a test than did subjects who were activated in a group identity before reading about the same peer. The reverse was also true: subjects who were activated individually before reading about an unsuccessful peer outperformed subjects who were activated in a group identity. Other studies used indirect identity manipulation, such as those exploring stereotype threat. Steele and Aronson (1995), for example, activated group identities of Black subjects by using the word “intelligence” in a test description. Because intelligence is linked to a stereotype about Black people, both the identity and the stereotypes became salient, and performance of Black subjects suffered relative to those who were not reminded of a stereotype. Finally, direct activation of an individual identity was found to offset the potentially debilitating effects of stereotype activation: stereotype-threatened subjects who self-affirmed by writing a positive statements about themselves performed better on a test than did threatened subjects who were not asked to self-affirm (Martens, Johns,

Greenberg, & Schimel, 2006). Thus, identity activation is shown to influence self-perceptions, expectations, and performance.

#### *5.2.1.5 Some Individuals More Susceptible.*

Substantial evidence suggests that different individuals are more susceptible to comparison effects than others. First, studies in social comparison and stereotype threat indicate that subjects who regard the domain in question to be part of their identities are particularly susceptible to threat (e.g., Steele, 1997; Lockwood & Kunda, 1997). Second, studies have shown that subjects with unstable academic efficacy (Aronson, Steele, Elliot, & Dweck, 2005), an external locus of control (Cadinu, Maass, Lombardo, & Frigerio, 2006), and strong group identification (Schmader, 2002) are more susceptible to stereotype threat effects than are individuals with the opposite characteristics. Fewer moderators have been explored with respect to social comparison vulnerability, although Gibbons & Buunk (1999) were able to show that some people are simply more likely than other people to make comparisons. Festinger (1954) suggested that uncertainty is a key motivator for social comparison; if one is uncertain about ability or opinion, comparisons with similar peers provide a basis for gathering such information about the self. This characteristic of uncertainty, therefore, very likely has a strong linkage with one's tendency to compare—and may in turn be extended to the group or individual level. Thus, the extent to which individuals experience or are placed in situations which prompt uncertainty, the more likely they are to experience identity fluctuations as a consequence of comparisons.

The preceding paragraphs have shown comparisons under different identity activations produce different perceptions and behaviors, and that uncertain individuals are more susceptible to the effects of threat. It is this very uncertainty, in fact, which may make the dissonance experienced as a result of identity conflicts more powerful. Because of the substantial research base supporting similar conclusions, it is reasonable to persist in believing that the hypotheses were in fact not incorrect, but rather, that the methods used were not optimal and were therefore ineffective.

#### *5.2.2 Were the Methods Inappropriate?*

Several pieces of evidence, some already explained in earlier paragraphs, point to a methodological failure in explaining the absence of expected patterns.

First, the stereotype may have been inappropriate for a study of this sort. In order to be threatened by a stereotype, an individual must find the domain to which the stereotype pertains to be personally important, and must have some aspect of their identity merged with the stereotyped group (citations). In this case, the academic domain was reportedly important to the subjects, but the students' identities did not appear to be merged with the stereotyped group. Data presented in the previous section substantiated the notion that, while stereotypes about community college students were known to exist, they held relatively little power over subjects because of low group identification. The very temporary nature of a community college fails in one of the basic conditions under which stereotype threat can occur—that the individual neither chooses nor can escape from the stereotyped group. Thus, the first methodological failure results from the selection of a permeable group.

Second, one of the key primes had an iatrogenic effect. The prime for the Upward-Individual condition appeared to inadvertently prompt identification rather than comparison, therein boosting performance and diminishing further the power of the group over the individual. If students had been inclined to question their own role as a community college student, then the identical decision made by the outstanding Harvard-admitted student who served as comparison target would have validated their own decision, highlighted their own capacity, and cemented their distance from a group identity that might otherwise be disadvantageous. This failing of the UI prime rendered the individual comparison conditions impotent and irrelevant, and resulted in a comparison of entirely different activation schemas.

Third, the use of the artificial newspaper styles survey may have been too deeply buried in subterfuge and subtleties. The prime was embedded within an ostensible test of newspaper styles in which subjects were asked to determine whether several articles were Internet- or paper based. This very subtlety, however, may in fact have minimized their power to influence test results. As subjects sought to complete a task for which no solution was possible—determining whether the articles were Internet-based or newspaper based--a cognitive dimension may have been activated when an emotional reaction was desired. Subjects may have read the article appraisingly, seeking evidence for or against a particular news source and ignoring the content of the article. The

emotional response that might otherwise have been present would have been muted by the deliberate cognitive attempt to determine the source. If the articles failed to activate an implicit emotional response beneath the active cognitive processes (assuming, that is, that an emotional response might have been felt for so distant a stereotype), then the experiment could not have been successful.

Fourth, the study was conducted over a single day in a number of one-hour sessions. Although they were directed not to discuss the study, it is possible or likely that some subjects who participated in the afternoon became aware of the true purpose of the study by talking with friends on this relatively small campus with only 600 full time students. With such awareness, any potential effects of the primes would surely have been eviscerated for all but the earliest groups.

Fifth, the number of subjects was too low. The desired number of subjects ( $n=84$ ) were not obtained, and this planned number was too low to detect differences once subgroups like High-SCO were identified. After data cleaning, only 76 subjects remained in the pool for analysis. This low  $n$  diminished the power of the statistical techniques to detect differences if they were present, and resulted in a particularly low number of subjects ( $n=12$ ) in one of the conditions. After the High-SCO subgroup was identified, this power was further diminished.

Sixth, the planned comparisons used for analysis did not anticipate significant differences between the two upward groups—but differences between the UI and UG groups did in fact exist. Since the key comparisons relied on the combinations of these groups, much of the alpha power in this study was squandered on inappropriate comparisons. Any additional comparisons that may have been more appropriate were unlikely to result in significant findings because additional tests would have produced an inappropriately high familywise error rate.

Finally, the moderators related to social comparison orientation, domain identification, and group identification were for logistical reasons gathered after the test was taken. While the data still provided information of interest, they may have been tainted by the experiment itself. None of the moderators proved to be particularly useful—but it cannot be known whether they would have been predictive if gathered prior to the experiment.

Thus it is certainly plausible that the key hypotheses remain sound, but the methods with which to test them must be revamped in a number of ways to ensure that the desired effect is obtained. The next section will highlight the future directions in the illustration of this concept.

### 5.3 Future Explorations

If the conclusion can be drawn that the theory is indeed accurate but the methods were flawed, then changes to the design may produce the expected results. A future study should implement changes in a number of areas. For example, it should rely on an ascribed group about which stereotypes exist—women, perhaps, or members of a racial group. Stereotypes should be strong and well-known, and members of the stereotyped group should not be able to escape membership.

The individual comparison conditions must resemble the group and control comparison conditions as closely as possible, and must ensure that subjects compare with rather than identify with the member in question. This can likely be accomplished by (a) activating an individual identity rather than a group identity before reading the article, and (b) ensuring that subjects are not attending to the group but the individual characteristics of the comparison target.

A self-esteem measure of some sort must be embedded shortly after the prime and before the test, such that the effects of the prime can be assessed before the test is begun. The effects of the prime may have been diluted following the test and minimized the likelihood that significant results would be found. The measure should be short and concise, serving as a quick litmus test that still allows the effects to persist for the test itself.

Moderator data must be gathered before the experiment so that they are not tainted by the experiment itself. When they are gathered after the experiment, they are subject to the effects of the experiment and would probably serve better as consequences than predictors of an effect. Ideally, they should be gathered at an earlier time unrelated to the study in question; if that could not be achieved, then the gathering of such data before the study will improve their veracity.

Finally, a sufficiently large number of subjects must be obtained to balance power and Type I error through the range of group comparisons needed. If, for instance, SCO is

still predicted as a moderator, then a sufficient number of subjects is needed to allow the number to be halved and still have sufficient power to minimize Type II error.

#### 5.4 Concluding Thoughts

The power of stereotypes to threaten performance has generated increasing public interest over its 14-year exploration. With the recent national election in which members of stereotyped groups played key roles, the power of such psychological threats to diminish performance takes on global importance. If this theory can be validated, and the nature of the threat experience can be understood as a direct consequence of comparisons made in times of self-doubt, then its power can perhaps be minimized through shifts in the environment in which one learns to compare and through shifts in individual perceptions of others which lead to comparison decrements. Although this study was not able to illustrate the phenomenon as it was conceptualized, future studies should be conducted to more closely isolate and understand the nuances of this effect.

## REFERENCES

- Alicke, M. D., LoSchiavo, F. M., Zerbst, J., & Shaobo, Z. (1997). The Person Who Outperforms Me Is a Genius: Maintaining Perceived Competence in Upward Social Comparison. *Journal of Personality & Social Psychology*, 73(4), 781-789.
- Allport, G. W. (1954). *The nature of prejudice*. Oxford, England: Addison-Wesley.
- Ambady, N., Paik, S. K., Steele, J., Owen-Smith, A., & Mitchell, J. P. (2004). Deflecting negative self-relevant stereotype activation: The effects of individuation. *Journal of Experimental Social Psychology*, 40(3), 401.
- Aronson, J., & Inzlicht, M. (2004). The Ups and Downs of Attributional Ambiguity. *Psychological Science*, 15(12), 829-836.
- Aronson, J., & Lustina, M. J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, 35(1), 29.
- Aronson, J., Steele, C. M., Elliot, A. J., & Dweck, C. S. (2005). *Stereotypes and the Fragility of Academic Competence, Motivation, and Self-Concept*. New York, NY, US: Guilford Publications.
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, 54(7), 462-479.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71(2), 230-244.
- Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, 66(1), 5-20.
- Biernat, M., Manis, M., Stapel, D. A., & Suls, J. (2007). Stereotypes and Shifting Standards: Assimilation and Contrast in Social Judgment. In *Assimilation and contrast in social psychology*. (pp. 75-97). New York, NY, US: Psychology Press.
- Blanton, H., Crocker, J., & Miller, D. T. (2000). The effects of in-group versus out-group social comparison on self-esteem in the context of a negative stereotype. *Journal of Experimental Social Psychology*, 36(5), 519-530.
- Brewer, M. B., & Gardner, W. (1996). Who is This "We"? Levels of Collective Identity and Self Representations. *Journal of Personality & Social Psychology*, 71(1), 83-93.
- Bronfenbrenner, U. (1979). *The Ecology of Human Development*.
- Brown, I., & Inouye, D. K. (1978). Learned helplessness through modeling: The role of perceived similarity in competence. *Journal of Personality and Social Psychology*, 36(8), 900-908.
- Brown, R. P., & Pinel, E. C. (2003). Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology*, 39(6), 626.
- Bui, K. V. T., & Pelham, B. W. (1999). Cognitive and Affective Reactions to Social Comparison. *Journal of Social Behavior & Personality*, 14(4), 569-583.
- Butzer, B., & Kuiper, N. A. (2006). Relationships between the frequency of social comparisons and self-concept clarity, intolerance of uncertainty, anxiety, and depression. *Personality and Individual Differences*, 41(1), 167-176.

- Buunk, B. P., Collins, R. L., Taylor, S. E., & VanYperen, N. W. (1990). The affective consequences of social comparison: Either direction has its ups and downs. *Journal of Personality and Social Psychology*, 59(6), 1238-1249.
- Cadinu, M., Maass, A., Lombardo, M., & Frigerio, S. (2006). Stereotype threat: the moderating role of Locus of Control beliefs. *European Journal of Social Psychology*, 36(2), 183-197.
- Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal of Personality and Social Psychology*, 59(3), 538-549.
- Caporrimo, R. (2008). Community College Students: Perceptions and Paradoxes. *Community College Journal of Research & Practice*, 32(1), 25-37.
- Cheryan, S., & Bodenhausen, G. V. (2000). When Positive Stereotypes Threaten Intellectual Performance: The Psychological Hazards of 'Model Minority' Status. *Psychological Science*, 11(5), 399.
- Cohen, G. L., & Garcia, J. (2005). "I Am Us": Negative Stereotypes as Collective Threats. *Journal of Personality & Social Psychology*, 89(4), 566-582.
- Cooley, C. H. (1902). *Human Nature and the Social Order*. New York: Scribner & Sons.
- Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, 96(4), 608-630.
- Croizet, J.-C., Dutrevis, M., & Desert, M. (2002). Why do students holding non-prestigious high school degrees underachieve at the university? *Swiss Journal of Psychology*, 61(3), 167-175.
- Davis, C., III, Aronson, J., & Salinas, M. (2006). Shades of Threat: Racial Identity as a Moderator of Stereotype Threat. *Journal of Black Psychology*, 32(4), 399-417.
- DeGenaro, W. (2006). Community Colleges, the Media, and the Rhetoric of Inevitability. *Community College Journal of Research & Practice*, 30(7), 529-545.
- Dijksterhuis, A., Van Knippenberg, A., Spears, R., Stapel, D. A., Koomen, W., & Postmes, T. (1998). Seeing One Thing and Doing Another: Contrast Effects in Automatic Behavior. *Journal of Personality & Social Psychology*, 75(4), 862-871.
- Erikson, E. H. (1980). *Identity and the Life Cycle*. New York: W.W. Norton & Company, Inc.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-140.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.
- Gibbons, F. X. (1986). Social comparison and depression: Company's effect on misery. *Journal of Personality and Social Psychology*, 51(1), 140-148.
- Gibbons, F. X., Blanton, H., Gerrard, M., Buunk, B., & Eggleston, T. (2000). Does social comparison make a difference? Optimism as a moderator of the relation between comparison level and academic performance. *Personality and Social Psychology Bulletin*, 26(5), 637-648.
- Gibbons, F. X., & Buunk, B. P. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology*, 76(1), 129-142.
- Gibbons, F. X., & Gerrard, M. (1991). Downward comparison and coping with threat. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and*

- research. (pp. 317-345). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.
- Gibbons, F. X., & McCoy, S. B. (1991). Self-esteem, similarity, and reactions to active versus passive downward comparison. *Journal of Personality and Social Psychology*, 60(3), 414-424.
- Gilbert, D. T., Giesler, R. B., & Morris, K. A. (1995). When comparisons arise. *Journal of Personality and Social Psychology*, 69(2), 227-236.
- Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin*, 28(5), 659-670.
- Gordijn, E. H., & Stapel, D. A. (2006). Behavioural effects of automatic interpersonal versus intergroup social comparison. *British Journal of Social Psychology*, 45(4), 717-729.
- Haslam, S. A., & Turner, J. C. (1992). Context-dependent variation in social stereotyping 2: The relationship between frame of reference, self-categorization and accentuation. *European Journal of Social Psychology*, 22(3), 251-277.
- Heatherton, T. F., & Polivy, J. (1991). Development and Validation of a Scale for Measuring State Self-Esteem. *Journal of Personality and Social Psychology*, 60(6), 895-910.
- Higgins, E. T., Bargh, J. A., & Lombardi, W. J. (1985). Nature of priming effects on categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(1), 59-69.
- Hogg, M. A., & Turner, J. C. (1987). Intergroup behaviour, self-stereotyping and the salience of social categories. *British Journal of Social Psychology*, 26(4), 325-340.
- Inzlicht, M., & Ben-Zeev, T. (2000). A Threatening Intellectual Environment: Why Females Are Susceptible to Experiencing Problem-Solving Deficits in the Presence of Males. *Psychological Science*, 11(5), 365.
- Johnson, C. S., & Stapel, D. A. (2007a). No Pain, No Gain: The Conditions Under Which Upward Comparisons Lead to Better Performance. *Journal of Personality & Social Psychology*, 92(6), 1051-1067.
- Johnson, C. S., & Stapel, D. A. (2007b). When different is better: performance following upward comparison. *European Journal of Social Psychology*, 37(2), 258-275.
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93(2), 136-153.
- Katz, I., Epps, E. G., & Axelson, L. J. (1964). Effect upon negro digit-symbol performance of anticipated comparison with whites and with other negroes. *The Journal of Abnormal and Social Psychology*, 69(1), 77-83.
- Kiesler, C. A., & Pallak, M. S. (1976). Arousal properties of dissonance manipulations. *Psychological Bulletin*, 83(6), 1014-1025.
- Kuhn, N., & McPartland, T. S. (1954). An empirical investigation of self-attitudes. *American Sociological Review*, 19, 68-76.
- Levy, B. (1996). Improving Memory in Old Age Through Implicit Self-Stereotyping. *Journal of Personality & Social Psychology*, 71(6), 1092-1107.
- Lockwood, P., & Kunda, Z. (1997). Superstars and Me: Predicting the Impact of Role Models on the Self. *Journal of Personality & Social Psychology*, 73(1), 91-104.

- Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, *18*(3), 302-318.
- Major, B., Testa, M., & Blysm, W. H. (1991). Responses to upward and downward social comparisons: The impact of esteem-relevance and perceived control. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research*. (pp. 237-260). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.
- Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology*, *42*(2), 236-243.
- Marx, D. M., & Stapel, D. A. (2006a). Distinguishing Stereotype Threat From Priming Effects: On the Role of the Social Self and Threat-Based Concerns. *Journal of Personality & Social Psychology*, *91*(2), 243-254.
- Marx, D. M., & Stapel, D. A. (2006b). It depends on your perspective: The role of self-relevance in stereotype-based underperformance. *Journal of Experimental Social Psychology*, *42*(6), 768-775.
- Marx, D. M., & Stapel, D. A. (2006c). UNDERSTANDING STEREOTYPE LIFT: ON THE ROLE OF THE SOCIAL SELF. *Social Cognition*, *24*(6), 776-791.
- Marx, D. M., Stapel, D. A., & Muller, D. (2005). We Can Do It: The Interplay of Construal Orientation and Social Comparisons Under Threat. *Journal of Personality & Social Psychology*, *88*(3), 432-446.
- McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology*, *27*(5), 486-493.
- Meichenbaum, D. H. (1971). Examination of model characteristics in reducing avoidance behavior. *Journal of Personality and Social Psychology*, *17*(3), 298-307.
- Mussweiler, T. (2007). Assimilation and Contrast as Comparison Effects: A Selective Accessibility Model. In D. A. Stapel & J. Suls (Eds.), *Assimilation and contrast in social psychology*. (pp. 165-185). New York, NY US: Psychology Press.
- Mussweiler, T., Gabriel, S., & Bodenhausen, G. V. (2000). Shifting Social Identities as a Strategy for Deflecting Threatening Social Comparisons. *Journal of Personality & Social Psychology*, *79*(3), 398-409.
- Mussweiler, T., Ruter, K., & Epstude, K. (2004). The man who wasn't there: Subliminal social comparison standards influence self-evaluation. *Journal of Experimental Social Psychology*, *40*(5), 689-696.
- Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, *3*(1), 1-18.
- Onorato, R. S., & Turner, J. C. (2002). Challenging the primacy of the personal self: The case for depersonalized self-conception. In M. F. Yoshihisa Kashima, Michael Platow (Ed.), *Self and identity: personal, social, and symbolic*. Mahweh, NJ: Lawrence Erlbaum & Associates.
- Onorato, R. S., & Turner, J. C. (2004). Fluidity in the self-concept: the shift from personal to social identity. *European Journal of Social Psychology*, *34*(3), 257-278.
- Pleban, R., & Tesser, A. (1981). The effects of relevance and quality of another's performance on interpersonal closeness. *Social Psychology Quarterly*, *44*(3), 278-285.

- Pronin, E., Steele, C. M., & Ross, L. (2004). Identity bifurcation in response to stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology*, 40(2), 152.
- Rosenberg, N. (1979). *Conceiving Self*. New York: Basic Books.
- Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. New York, NY, England: McGraw-Hill Book Company.
- Sackett, P. R., Hardison, C. M., & Cullen, M. J. (2004). On the Value of Correcting Mischaracterizations of Stereotype Threat Research. *American Psychologist*, 59(1), 48-49.
- Salovey, P., & Rodin, J. (1984). Some antecedents and consequences of social-comparison jealousy. *Journal of Personality and Social Psychology*, 47(4), 780-792.
- Schmader, T. (2002). Gender Identification Moderates Stereotype Threat Effects on Women's Math Performance. *Journal of Experimental Social Psychology*, 38(2), 194.
- Schwinghammer, S. A., Stapel, D. A., & Blanton, H. (2006). Different Selves Have Different Effects: Self-Activation and Defensive Social Comparison. *Personality and Social Psychology Bulletin*, 32(1), 27-39.
- Shepperd, J. A., & Taylor, K. M. (1999). Ascribing advantages to social comparison targets. *Basic and Applied Social Psychology*, 21(2), 103-117.
- Shih, M., Ambady, N., Richeson, J. A., Fujita, K., & Gray, H. M. (2002). Stereotype performance boosts: The impact of self-relevance and the manner of stereotype activation. *Journal of Personality and Social Psychology*, 83(3), 638-647.
- Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, 10(1), 80.
- Simon, B., & Hamilton, D. L. (1994). Self-stereotyping and social context: The effects of relative in-group size and in-group status. *Journal of Personality and Social Psychology*, 66(4), 699-711.
- Spencer, S. J., & Steele, C. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35(1), 4.
- Stapel, D. A., & Blanton, H. (2004). From Seeing to Being: Subliminal Social Comparisons Affect Implicit and Explicit Self-Evaluations. *Journal of Personality & Social Psychology*, 87(4), 468-481.
- Stapel, D. A., & Koomen, W. (2000). Distinctness of Others, Mutability of Selves: Their Impact on Self-Evaluations. *Journal of Personality & Social Psychology*, 79(6), 1068-1085.
- Stapel, D. A., & Koomen, W. (2001). I, We, and the Effects of Others on Me: How Self-Construal Level Moderates Social Comparison Effects. *Journal of Personality & Social Psychology*, 80(5), 766-781.
- Stapel, D. A., & Tesser, A. (2001). Self-Activation Increases Social Comparison. *Journal of Personality & Social Psychology*, 81(4), 742-750.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613-629.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality & Social Psychology*, 69(5), 797-811.

- Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology, 77*(6), 1213-1227.
- Stone, J., Perry, Z. W., & Darley, J. M. (1997). 'White men can't jump': Evidence for the perceptual confirmation of racial stereotypes following a basketball game. *Basic and Applied Social Psychology, 19*(3), 291-306.
- Tajfel, H. (1982). *Social identity and intergroup relations*. Cambridge: Cambridge University Press.
- Tajfel, H., Turner, J. C., Jost, J. T., & Sidanius, J. (2004). The Social Identity Theory of Intergroup Behavior. In *Political psychology: Key readings*. (pp. 276-293). New York, NY, US: Psychology Press.
- Tesser, A. (1980). Self-esteem maintenance in family dynamics. *Journal of Personality and Social Psychology, 39*(1), 77-91.
- Tesser, A. (1991). Emotion in social comparison and reflection processes. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research*. (pp. 115-145). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.
- Tesser, A. (2000). On the Confluence of Self-Esteem Maintenance Mechanisms. *Personality & Social Psychology Review, 4*(4), 290-299.
- Tesser, A., Millar, M., & Moore, J. (1988). Some affective consequences of social comparison and reflection processes: The pain and pleasure of being close. *Journal of Personality and Social Psychology, 54*(1), 49-61.
- Thompson, T., & Perry, Z. (2005). Is the Poor Performance of Self-Worth Protective Students Linked with Social Comparison Goals? *Educational Psychology, 25*(5), 471-490.
- Turner, J. C. (1982). *Towards a cognitive redefinition of the social group*. Cambridge: Cambridge University Press.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Cambridge, MA, US: Basil Blackwell.
- Walton, G. M., & Cohen, G. L. (2003). Stereotype Lift. *Journal of Experimental Social Psychology, 39*(5), 456.
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences*. Oxford, England: Rand McNally.
- Wheeler, S. C., Jarvis, W. B. G., & Petty, R. E. (2001). Think Unto Others: The Self-Destructive Impact of Negative Racial Stereotypes. *Journal of Experimental Social Psychology, 37*(2), 173.
- Wheeler, S. C., & Petty, R. E. (2001). The Effects of Stereotype Activation on Behavior: A Review of Possible Mechanisms. *Psychological Bulletin, 127*(6), 797.
- Wills, T. A. (1991). Similarity and self-esteem in downward comparison. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research*. (pp. 51-78). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.
- Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin, 106*(2), 231-248.
- Yeung, K.-T., & Martin, J. L. (2003). The Looking Glass Self: An Empirical Test and Elaboration. *Social Forces, 81*(3), 843-879.

Yopyk, D. J. A., & Prentice, D. A. (2005). Am I an athlete or a student? Identity salience and stereotype threat in student-athletes. *Basic and Applied Social Psychology*, 27(4), 329-336.

## APPENDICES

## Appendix A. Newspaper Article Primes by Condition

### Newspaper Styles Survey Department of Journalism

The College of Journalism is conducting a study to determine whether readers can distinguish news stories published on the Internet from those published in print news sources. You will read two news stories, randomly assigned from a collection of stories. Both of the stories are taken from legitimate news sources, but one, neither, or both of the stories may have come from an Internet news source. Please read both stories, and answer the questions following the stories. You will have 4 minutes for each story, so need to work quickly. It may be difficult to determine the source of the story, but please do your best.

#### *All conditions:*

#### **Bridge to Reopen Under Limited Conditions**

The Michigan Street Bridge in Sturgeon Bay reopened to limited motor vehicle traffic Wednesday morning. The decision came after Sturgeon Bay officials committed to strictly enforce the bridge's 5-ton weight limit. "We simply cannot place too much strain on this structure," commented city official. "This is the last connection from the east to the west side of the city now that the Maple-Oregon Bridge is closed," he added.

The Wisconsin Department of Transportation (DOT) closed the Michigan Street Bridge to all motorists on Monday, July 21 as a safety precaution. Unfortunately, despite warnings from DOT for stricter enforcement of the 5-ton weight limit, heavily loaded vehicles continued to use the bridge. That left the department with few options but to close the structure to all motor vehicle traffic.

Since the closure, the DOT has been working closely with the town of Sturgeon Bay officials to reach an agreement to reopen the bridge, while guaranteeing the 5-ton weight limit is followed.

The new guidelines include:

- \* On-site enforcement of the 5-ton weight limit by uniformed local law enforcement
- \* Bridge open to traffic only between 6am and 10pm Monday through Sunday
- \* Law enforcement checkpoints at intersections to detour overweight vehicles
- \* Maximum vehicle type allowed to cross  $\frac{3}{4}$  ton truck (empty) and automobiles
- \* County and city law enforcement personnel will install and remove all barricades
- \* A bridge tender will be on-duty around the clock
- \* Signs and message boards to advise motorists on 5-ton weight limit

These arrangements will remain in place until the Maple-Oregon Bridge opens to regular traffic, expected sometime in late August. Town residents inconvenienced by the new guidelines are asked to submit their comments to the city commissioner before the next public forum.

1. Which of the following do you believe is the source for this news story?
  - a. Internet
  - b. Print newspaper
  - c. Other source (Describe: \_\_\_\_\_)
2. What aspects of the article led you to believe that it was from the source you indicated above?
  - a. Language use
  - b. Topic
  - c. Grammar

- d. Title \_\_\_\_\_)  
 e. Length \_\_\_\_\_  
 f. Other: (please list: \_\_\_\_\_)

3. Write down one phrase from the article that led you to believe it was from the source you indicated above: \_\_\_\_\_

**Appendix A (cont'd)**

***1. Condition Downward Group (DG)***

**New Initiative Helps Technical School Students Transition to Universities**

Four-year comprehensive universities in the UW system are gearing up to address the challenges faced by students who transfer from technical schools.

More than 1,100 technical school students transfer to a University of Wisconsin campus each year, but their prospects once they reach the comprehensive campus are poor. Compared to students attending who began college at 2-year community colleges, students from technical schools face formidable challenges.

“It’s been tough to get them [technical school students] caught up,” commented Dr. John Markus, dean of admissions at UW-Eau Claire. He explained that they simply had not had the necessary experiences to prepare them for academic challenges on a comprehensive campus.

Students who begin their degrees at technical schools earn GPAs that are nearly half of a grade point lower than their peers who begin college in a 2-year community college. An average student who transfers from a technical school to a university will earn an average GPA of 2.6, compared with 3.0 for students who transfer from community colleges.

But the problems don’t end with grade point averages. Nearly half of students who transfer from technical schools drop out within a year of transferring to the 4-year campus, and many face academic probation for maintaining inadequate GPAs.

“The two-year campuses simply prepare their students better for academic success. We are never worried about students who come from a 2-year campus,” commented Dr. Markus.

The new initiative, called “Staying the Course,” will bring together representatives from both institutions who will work together to ensure greater success for technical school students. They hope to support technical school transfer students through increased integration into the university community and access to tutoring and child care.

1. Which of the following do you believe is the source for this news story?
  - a. Internet
  - b. Print newspaper
  - c. Other source (describe: \_\_\_\_\_)
2. What aspects of the article led you to believe that it was from the source you indicated above?
  - a. Language use
  - b. Topic
  - c. Grammar
  - e. Title
  - e. Length
  - f. Other: (please list: \_\_\_\_\_)
3. Write down one phrase from the article that led you to believe it was from the source you indicated above: \_\_\_\_\_

**Appendix A (cont'd)**

**2. Condition Upward Group (UG)**

**New Initiative Helps Community College Students Transition to Universities**

Four-year comprehensive universities in the UW system are gearing up to address the challenges faced by students who transfer from 2-year community colleges.

More than 1,100 community college students transfer to 4-year comprehensive universities in Wisconsin each year, but their prospects once they reach the comprehensive campus are poor. Compared to students attending four-year universities from the start, students at community colleges face formidable challenges.

“It’s been tough to get them [community college students] caught up,” commented Dr. John Markus, dean of admissions at UW-Eau Claire. He explained that they simply had not had the necessary experiences to prepare them for academic challenges on a four-year campus.

Students who begin their degrees at community colleges earn GPAs that are significantly lower than their peers who begin college at a 4-year university. An average student who transfers from a community college to a comprehensive university will earn a GPA of 2.85, compared with 3.0 GPA for students who began at a comprehensive institution.

But the problems don’t end with grade point averages. Nearly half of students who transfer from community colleges to four-year universities drop out within a year of transferring to the 4-year campus, and many face academic probation for maintaining inadequate GPAs.

“The two-year campuses simply don’t prepare their students for academic success. We rarely worry about students who begin at a 4-year campus,” commented Dr. Markus.

The new initiative, called “Staying the Course,” will bring together representatives from both institutions for collaborative efforts to ensure greater success for transfer students. They hope to provide support to community college transfers through increased integration into the 4-year campus and access to tutoring and child care.

1. Which of the following do you believe is the source for this news story?
  - a. Internet
  - b. Print newspaper
  - c. Other source (describe: \_\_\_\_\_)
  
2. What aspects of the article led you to believe that it was from the source you indicated above?
  - a. Language use
  - b. Topic
  - c. Grammar
  - f. Title
  - e. Length
  - f. Other: (please list: \_\_\_\_\_)
  
3. Write down one phrase from the article that led you to believe it was from the source you indicated above: \_\_\_\_\_

**Appendix A (cont'd)**

**3. Condition Upward Individual (UI)**

**College to Provide Travel and Scholarships for Top Students**

The advising department on community college campuses is directing its focus to its most successful students, adding travel opportunities, recognition, and scholarships to enrich their academic experiences.

The number of top students on two year campuses has been on the rise in recent years, due in part to changes in tuition and financial aid across top campuses.

“We are now finding that students with ACT scores as high as 32 or 33 and outstanding academic and leadership records are staying at home to attend community colleges,” commented Dr. Lynn Markus, Dean of Admissions. “The gives the two-year campuses an unprecedented opportunity and obligation to create rich academic experiences.”

One such student, “Casey” (not the student’s real name), is on the list of students to receive enrichment. “I had been accepted into both Harvard University and the University of Michigan, but the costs were just impossible for my family and I couldn’t get financial aid,” Casey said. “Now I am doing all I can to make my two-year campus the best place it can be. Eventually, I’ll have enough money to attend a larger university.”

In addition to maintaining a 4.0 grade point average with 21-credit loads, Casey had led the student senate, organized a range of environmentally-friendly campus practices, organized the “Broad Thinkers,” (a student group that organizes campus debates), and activated the social life on campus through a range of films, bands, and gatherings.

“Students like Casey have improved the entire culture of our campus exponentially. It’s time we offer something back to them,” commented Dr. Markus.

Students who have reached such notably high levels of achievement will be notified by letter of the opportunities available to them.

1. Which of the following do you believe is the source for this news story?
  - a. Internet
  - b. Print newspaper
  - c. Other source (describe: \_\_\_\_\_)
  
2. What aspects of the article led you to believe that it was from the source you indicated above?
  - a. Language use
  - b. Topic
  - c. Grammar
  - g. Title
  - e. Length
  - f. Other: (please list: \_\_\_\_\_)
  
3. Write down one phrase from the article that led you to believe it was from the source you indicated above: \_\_\_\_\_

**Appendix A (cont'd)**

**4. Condition Downward Individual (DI)**

**College to Provide Tutoring and Counseling for Struggling Students**

The advising department on community college campuses is directing its focus to struggling students, adding tutors and academic counselors to guide them through difficult academic periods.

The number of students who struggle academically has been on the rise in recent years, due in part to changes in admission practices across campuses. "We are now admitting students who received scores as low as 10 or 11 on the ACT and low C and D grade point averages in high school," commented Dr. Lynn Markus, Dean of Admissions. "Students who enter college at this level cannot be expected to succeed without a lot of support."

One such student, "Casey" (not the student's real name), is on the list of students to be assisted. "I never really expected to even get into college," Casey said. "My grades have always been pretty bad, and I knew I didn't do good on the ACT and stuff like that."

Casey's GPA slipped to a 1.1 this semester, down from 1.4 in fall.

"Mostly, it's because I have trouble with tests," Casey said, explaining that studying for tests was difficult and that the required reading was difficult. "I also skip quite a few classes for work or because I'm too tired," Casey admitted, adding that several late assignments had made things worse.

Missed classes, late assignments, and failed exams account for most of the students on probation, agreed Dr. Markus. "Students like Casey just don't have an easy time with academics," she added.

Students who are in academic difficulties will receive a letter from the advising department inviting them to seek assistance from several counselors and tutors who have been hired specifically to address such problems.

1. Which of the following do you believe is the source for this news story?
  - a. Internet
  - b. Print newspaper
  - c. Other source (describe: \_\_\_\_\_)
  
2. What aspects of the article led you to believe that it was from the source you indicated above?
  - a. Language use
  - b. Topic
  - c. Grammar
  - h. Title
  - e. Length
  - f. Other: (please list: \_\_\_\_\_)
  
3. Write down one phrase from the article that led you to believe it was from the source you indicated above: \_\_\_\_\_

**Appendix A (cont'd)**

**5. Control Condition**

**Residents Remember Original Ashwood Bridge as New Bridge Opens**

Winston resident Daniel DeForest was 14 years old when the James Ashwood Bridge was dedicated in 1933. Since his father was a city alderman and prominent politician, the young boy got a prime seat at the ceremony.

Though DeForest's memories of actual events during the dedication might be a little fuzzy, he clearly recalls that the U.S. senator Tim Bradford was in attendance. Bradford, DeForest remembered, looked just like U.S. President Franklin Roosevelt. Bradford served as U.S. senator from 1939 to 1963. He ran unsuccessfully for governor in 1936.

"You'd see pictures of him all over the place, and he was a dead-ringer for FDR," DeForest said. "The stage was set up on the east end of the bridge, on a slope heading toward the river. There were a lot of people there, but my dad got a good seat."

Margaret Alvery, another longtime resident of the area, recalled her first trip over the old Ashwood Bridge. "It was the first time I'd crossed such a big bridge, and I can still remember my heart racing as I looked down so far to the river," she recalled. Her uncle and cousin had been in a boat below, waving as they passed over.

Like many in the area who grew up in or near Winston, DeForest and Alvery have fond childhood memories of crossing the James Ashwood Bridge on foot or bicycle. And while those contacted by the newspaper cherish the memories, a few said it's also time to say goodbye.

The structure is scheduled to be torn down in 2008 when the new James Ashwood Bridge, now under construction just south of the current span, is open to traffic.

1. Which of the following do you believe is the source for this news story?
  - a. Internet
  - b. Print newspaper
  - c. Other source (describe: \_\_\_\_\_)
  
2. What aspects of the article led you to believe that it was from the source you indicated above?
  - a. Language use
  - b. Topic
  - c. Grammar
  - i. Title
  - e. Length
  - f. Other: (please list: \_\_\_\_\_)
  
3. Write down one phrase from the article that led you to believe it was from the source you indicated above: \_\_\_\_\_

## Appendix B. Demographic Items by Condition

---

### 1. Condition Upward Group (UG)

#### Measure of Intellectual Potential for College Success

You are about to take a test consisting of items to measure intellectual potential. Before taking the test, it is necessary that you provide some general information so that we can better interpret your results. Please respond to the following questions. You will have 3 minutes.

1. What is your current year in college?      Fresh.    Soph.    Jr.      Sr.      Other
2. What type of school do you attend?
  - 2-year UW college (e.g., Manitowoc, Marshfield, Wausau)
  - 4-year UW school (e.g., UW-Stevens Point, UW-Eau Claire, etc.)
  - Other: (please name: \_\_\_\_\_)
3. For how many months have you been a student at this college?    \_\_\_ months
4. What degree (if any) do you plan to obtain? \_\_\_\_\_
5. How many miles do you travel to get to college?    \_\_\_ miles each direction
6. What is your usual mode of transportation? \_\_\_\_\_

#### ABOUT THE TEST

GENERAL INFORMATION: The following items are being tested to see how well they predict success for students in the later years of college. Thus far, the questions are shown to be effective predictors of success in junior and senior years of college and good indicators of intelligence.

---

### 2. Condition Downward Group (DG)

#### Measure of Intellectual Potential for College Success

You are about to take a test consisting of items to measure intellectual potential. Before taking the test, it is necessary that you provide some general information so that we can better interpret your results. Please respond to the following questions. You will have 3 minutes.

1. What is your current year in college?      Fresh.    Soph.    Jr.      Sr.      Other
2. What type of school do you attend?
  - 2-year UW college (e.g., Manitowoc, Marshfield, Wausau)
  - Technical School (e.g., Midwest Technical College, Fox Valley Tech)
  - Other: (please name: \_\_\_\_\_)
3. For how many months have you been a student at this college?    \_\_\_ months
4. What degree (if any) do you plan to obtain? \_\_\_\_\_
5. How many miles do you travel to get to college?    \_\_\_ miles each direction
6. What is your usual mode of transportation? \_\_\_\_\_

#### ABOUT THE TEST

GENERAL INFORMATION: The following items are being tested to see how well they predict success for students in the later years of college. Thus far, the questions are shown to be effective predictors of success in junior and senior years of college and good indicators of intelligence.

## Appendix B (Cont'd)

### 3. Condition Upward Individual (UI)

#### Measure of Intellectual Potential for College Success

As part of the testing protocol, we are interested in knowing what qualities students believe lead to academic difficulties. To help us in understanding this, we ask you to think of someone in your grade in college who you think might perform worse than you on a test like the ACT or SAT. If you are female, think of another female. If you are male, think of another male.

1. Write this person's first initial: \_\_\_\_\_
2. About how many years have you known this person? \_\_\_\_\_ years
3. Has this person previously performed very poorly on some academic measure?  
\_\_\_ Yes      \_\_\_ No      \_\_\_ Unsure
4. If you answered yes to number 2, what measure was it? \_\_\_\_\_
5. Write one characteristic of this person that contributes to his or her difficulty:  
\_\_\_\_\_

#### ABOUT THE TEST

GENERAL INFORMATION: The following items are being tested to see how well they predict success for students in the later years of college. Thus far, the questions are shown to be effective predictors of success in junior and senior years of college and good indicators of intelligence.

-----

### 4. Condition Downward Individual (DI)

#### Measure of Intellectual Potential for College Success

As part of the testing protocol, we are interested in knowing what qualities students believe lead to academic success. To help us in understanding this, we ask you to think of someone in your grade in college who you think might perform better than you on a test like the ACT or SAT. If you are female, think of another female. If you are male, think of another male. When you are finished, you may try out the sample test items on the next page.

1. Write this person's first initial: \_\_\_\_\_
2. About how many years have you known this person? \_\_\_\_\_ years
3. Has this person previously performed very well on some academic measure?  
\_\_\_ Yes      \_\_\_ No      \_\_\_ Unsure
4. If you answered yes to number 2, what measure was it? \_\_\_\_\_
5. Write one characteristic of this person that contributes to his or her success:  
\_\_\_\_\_

#### ABOUT THE TEST

GENERAL INFORMATION: The following items are being tested to see how well they predict success for students in the later years of college. Thus far, the questions are shown to be effective predictors of success in junior and senior years of college and good indicators of intelligence.

### 5. Control Condition

## Measure of Intellectual Potential for College Success

Please read the instructions before answering the questions.

You are about to take a test consisting of items to measure intellectual potential. Before taking the test, it is necessary that you provide some general information so that we can better interpret your results. Please respond to the following questions. When you are finished, you may look at the sample test items on the following page.

1. What is your current year in college?    Fresh.    Soph. Jr.    Sr.    Other
2. For how many months have you been a student at this college?    \_\_\_\_\_  
months
3. About how many miles do you travel to reach your college?    \_\_\_\_\_ miles
4. What is your usual mode of transportation? \_\_\_\_\_
5. About how many people live in the town you consider to be your hometown? \_\_\_\_\_
6. About how many miles did you live from your high school?    \_\_\_\_\_ miles

### ABOUT THE TEST

GENERAL INFORMATION: The following items are being tested to see how well they predict success for students in the later years of college. Thus far, the questions are shown to be effective predictors of success in junior and senior years of college and good indicators of intelligence.

## Appendix C. Test Instructions and Items

### TEST INSTRUCTIONS

1. You must work **mentally**. Do **not** use pencil or paper during this test except to mark your answers.
2. For every question, indicate whether the statement is true or false.
3. If you average more than **twenty seconds** to respond to each of the questions, your score will be lower. During the test, you must read and respond to a total of 37 statements in thirteen minutes or less.  
**So, be ready to concentrate and think fast!**

1. T F The word "asterisk" can be spelled using only the letters found in the word "Strike east!"
2. T F This sequence of four words, "tricycle, hand, ruler, eyes," corresponds to this sequence of numbers "3, 5, 12, 2."
3. T F 36 minutes before 9 o'clock is 24 minutes past 7 o'clock.
4. T F The word, "stickers," is spelled by using the first letters of the words in the following sentence: "Several tame ibex can kiss each relevant salad."
5. T F If written backwards, the number, "one thousand, two hundred forty-seven," would be written "seven thousand, four hundred twelve."
6. T F Jessie has only sixty-eight dollars, but he can buy a bicycle that costs one hundred forty dollars, (disregarding tax) if he borrows seventy-seven dollars from Irma and fifteen dollars from Irene.
7. T F A round wall clock that has been rotated until it is hanging upside down will have a minute hand that points to your left when it is three fifteen.
8. T F If Sally looks into a mirror and touches her right ear with her left hand, Sally's image seems to touch its left ear with its right hand.
9. T F If you leave the letters in the same order, but rearrange the spaces in the phrase, "The rest ones," it can be read as, "There stones."
10. T F The words, "education, auctioned, and cautioned," all use the exact same letters.
11. T F Jean weighs 87 pounds. Jane weighs 123 pounds. Josie weighs 103 pounds. Two of them standing together on the same scale could weigh 200 pounds.
12. T F The seventh vowel that appears in this sentence is the letter "e."
13. T F Seven chickens, three dogs, and three cats have a total of forty legs.
14. T F Forty minutes are to one hour as sixteen hours are to one day.
15. T F In the English alphabet, there are exactly four letters between the letter "P" and the letter "J."
16. T F If the word, "BEG," is written under the word, "TOY," and the word, "VAN," is written

under "BEG," then the word, "TEN," is formed diagonally.

17. T F By removing seven letters from the word, "fatherhood," the word, "hoof," can be formed.
18. T F Three of the following numbers add up to the number 51: 27, 3, 2, 29, 5.
19. T F The words, "any, when, foot, way," can form common compound words using, respectively, "body, ever, toe, free."
20. T F If Tuesday is the first day of the month, the very next Sunday is the fifth day of the month.
21. T F If cats have five claws in their front paws and five in their back claws, then three front paws and four back paws normally hold thirty-five claws.
22. T F Fannie will be four blocks from her starting place if she travels three blocks north, then two blocks east, and then four blocks south.
23. T F The following words are the opposites of words that begin with the letter T: hot, false, short, permanent, give.
24. T F The following, disregarding punctuation, is spelled the same forwards as it is backwards: "Are we not drawn onward, we few, drawn onward to new era?"
25. T F The letters of the word, "stroked," appear in reverse alphabetical order.
26. T F The numbers, 4-8-3-5-9-2-6, are read backwards as 6-2-9-5-3-8-4.
27. T F The odd numbers in this group add up to an even number: 17, 46, 9, 13, 42, 5, 3.
28. T F Without breaking or bending a toothpick, you can spell the word, "AIM," with exactly seven toothpicks, with no letter sharing a toothpick used by another letter.
29. T F The number of letters in this sentence is forty-five.
30. T F A square whose sides each measure fifteen centimeters can completely fit inside of a regular octagon whose sides each measure fifteen centimeters.
31. T F Eight identical triangles can be formed by drawing two straight lines through an hexagon's center point.
32. T F The number 80 is the next logical number in the following sequence of numbers: 3, 8, 18, 38...
33. T F Jane is shorter than Sam. Rhonda is shorter than Jane. Therefore, Sam is the tallest boy.
34. T F The sum of all the odd numbers from two to 18 is an odd number.
35. T F If each of eight cats in a rubs noses with each of another seven cats, then a total of fifty-six noserubs occurs.
36. T F Three congruent regular octagons can be drawn in such a way that all of them overlap each other and create exactly nine distinct areas or compartments.
37. T F If you begin outside a doughnut shaped house that has three doors to the outside and four doors to the inner courtyard, then it's possible to end up back outside by walking through all seven doors of the house without ever walking through the same door twice.

## Appendix D. Self-esteem measures and moderators

### PERSONAL VIEWS

The following items are being used to determine whether students today feel differently about themselves and others than students from their parents' generation. Please circle the number indicating your agreement with each of the following statements.

	Strongly Disagree			Strongly Agree		
1. I am confident in my abilities.	1	2	3	4	5	6
2. Being a student is important to me.	1	2	3	4	5	6
3. I often compare myself with others with respect to what I have accomplished in life	1	2	3	4	5	6
4. Most people consider community college students to be more effective than students from other schools.	1	2	3	4	5	6
5. I feel that others respect me.	1	2	3	4	5	6
6. I often regret that I am a community college student.	1	2	3	4	5	6
7. If I want to find out how well I have done something, I compare what I have done with how others have done.	1	2	3	4	5	6
8. My academic performance means a lot to me.	1	2	3	4	5	6
9. I am worried about what other people think of me.	1	2	3	4	5	6
10. Overall, the fact that I am a community college student has very little to do with how I feel about myself.	1	2	3	4	5	6
11. I always pay a lot of attention to how I do things compared with how others do things	1	2	3	4	5	6
12. I feel concerned about the impression I am making	1	2	3	4	5	6
13. In general, I'm glad to be a community college student.	1	2	3	4	5	6
14. I am not the type of person who compares often with others.	1	2	3	4	5	6
15. I don't really care that much about how well I do in college.	1	2	3	4	5	6
16. I feel smart.	1	2	3	4	5	6
17. In general, others think that community college students are unworthy.	1	2	3	4	5	6
18. I often compare how I am doing academically with other people.	1	2	3	4	5	6
19. Being a community college student is an important reflection of who I am.	1	2	3	4	5	6

## Appendix E. Manipulation Check 1 (Twenty Statements Task)

### WHO ARE YOU?

Please tell who you are to someone who does not know you by completing the following fifteen statements. You will have three minutes, so you will have to work quickly.

1. I am \_\_\_\_\_.
2. I am \_\_\_\_\_.
3. I am \_\_\_\_\_.
4. I am \_\_\_\_\_.
5. I am \_\_\_\_\_.
6. I am \_\_\_\_\_.
7. I am \_\_\_\_\_.
8. I am \_\_\_\_\_.
9. I am \_\_\_\_\_.
10. I am \_\_\_\_\_.
11. I am \_\_\_\_\_.
12. I am \_\_\_\_\_.
13. I am \_\_\_\_\_.
14. I am \_\_\_\_\_.
15. I am \_\_\_\_\_.

## Appendix F. Manipulation Check 2

### *Conditions UG, DG, and Control*

#### **General Understanding of the Tasks**

Please complete the following questions so that the researchers can be certain that you fully understood why you were completing the various tasks.

1. According to the news article you read at the start of the study, how successful were students from community colleges?

Not successful--1      2      3      5      6--Very successful      Not sure

2. What was the purpose of the Newspaper Styles Survey?

*[Answer space provided]*

3. Why was the purpose of the test items?

*[Answer space provided]*

4. Were the various studies you completed connected in any way? If yes, explain.

*[Answer space provided]*

---

### *Conditions UI and DI*

#### **General Understanding of the Tasks**

Please complete the following questions so that the researchers can be certain that you fully understood why you were completing the various tasks.

1. According to the news article you read at the start of the study, how successful was the student Casey Halverson?

Not successful--1      2      3      5      6--Very successful      Not sure

2. What was the purpose of the Newspaper Styles Survey?

*[Answer space provided]*

3. Why was the purpose of the test items?

*[Answer space provided]*

4. Were the various studies you completed connected in any way? If yes, explain.

*[Answer space provided]*

## **Appendix G. Recruiting script**

You will all have the opportunity to earn [x] points of extra credit this year by participating in three studies. The studies will involve approximately 50 minutes of your time, and will be offered at these times and locations: [List times and locations]. For the first of the four studies, the journalism study, you would be reading two newspaper articles and indicating the sources of the articles. For the second of the four studies, the education study, you would be taking a test to test the usability of items being developed for a new intellectual measure. For the third of the four studies, the personality study, you would be describing yourself and indicating the extent to which you agree with approximately sixteen statements about how you feel about yourself. Finally, you will complete a series of items exploring your identity. You may choose to participate or not to participate, and may leave at any time during the study. To receive full bonus points, you should participate in all four studies.

I will now hand out information about the studies. If you have questions, you may ask me in the next class period or contact the lead researcher by email or telephone. If you agree to participate, you may return this signed form to me before [insert date]. At the time of the study, you will sign permission slips for each of the individual departments as well.

After taking part in the study, you will be given an extra credit slip to give to your instructor. If you agree to participate, all of your information will be completely confidential, and your name would not be associated with the data you provided. Your participation in this study will be very helpful to researchers exploring a variety of topics, and may eventually benefit many more people as the knowledge gained by researchers is applied to more people. Thank you for considering using your time to support this research."

## **Appendix H. Informed Consent**

### **STUDY INFORMATION AND CONSENT FORM**

You are invited to participate in several studies through the University of Minnesota. You were selected because you are currently enrolled in a college in a non-urbanized area, and information from many different types of students are necessary for good research. Please read this form and ask any questions you may have before agreeing to be in the study.

#### **What is being studied, and what will I have to do?**

You will be asked to participate in four separate studies for different colleges at the University of Minnesota. You can read about each of the studies below.

Study 1. This study is interested in how college students like you view news sources. The researchers want to know if you can tell whether a news story is from the Internet or from a newspaper. This is an important topic, as your generation is one of the first in which online news sources are used more often than text-based sources. For this study, you will be asked to read a news story and indicate whether you believe it is Internet or newspaper-based, and why you think this is the case. This study will take less than 10 minutes

Study 2. Some education researchers are working to develop new test items for a test of intellectual ability. To tell if the test items are valid, they need many college students like you to try to answer the questions. This study will take about 15 minutes. In this test, you will be asked to try and figure out the correct answers for the new test items.

Study 3. This study is interested in knowing how students like you see yourselves. This is an important topic, because some people believe that changes in student self-views contribute to the increases in school violence. This study will take about 10 minutes. In this study, you will indicate be asked to how much you agree with a 16 statements, and to provide brief descriptions of yourself.

#### **How do I participate?**

If you decide to participate in the study, you simply need to come to [insert room] at any of the following times on [insert date]: 9:00, 10:00, 11:00, 12:00, 1:00, 2:00, 3:00, or 4:00. One or more researchers will be in the classroom to conduct the study and answer the questions. You would then do your best to answer the questions that are given to you. When you are done, you will receive a bonus coupon that you can turn in to your instructor for bonus points.

#### **What if I don't want to participate?**

It's fine if you don't want to participate. If you don't want to participate, don't come to the study! You can earn extra credit in another way, which will be described by your professor.

## **Appendix H (cont'd)**

### **What if I come to participate, but then change my mind and want to stop?**

You can stop taking the study at any time just by putting down your pencil. Although we want you to participate in all of the studies, you can stop at any time for any reason. Nobody will ask you to explain why you have stopped your work, and nobody will criticize you if you want to stop your work. If you prefer to leave the room instead of staying in the room, your teacher will call the librarian and you will be given a pass to go to the library. It's perfectly all right. If you decide to quit the study, you will still receive part of your bonus points.

### **Will anybody know what answers I give?**

No. Any answers you give will be completely private. Nobody, not even the researcher, will know what answers you gave. You will be asked to sign your name on permission forms for the different departments, but your name will not be written down anywhere and the signed forms will not be kept with your answers to the test. Your responses are completely confidential and private.

### **What risks or benefits are there?**

We don't expect you to experience any risks by participating in the study. If you felt anxious for any reason after participating in the study, several different people will be available to talk with you.

You also will not have any direct benefits by participating in the study, except for the bonus coupon you will receive. However, your work will provide future benefits to other high school students as the researchers make use of the information you provided in their work.

You will be asked not to talk to your peers about the study until all of the studies are complete, at 5:00 p.m. on the day of the study. This is to preserve the integrity of the study; if you tell somebody else what you did, then they might prepare for the study in a way that would bias their responses.

### **Who is conducting the study?**

Suzanne Russ will be conducting the studies in the Fox Cities area. Suzanne Russ is currently completing a doctoral degree in educational psychology at the University of Minnesota-Twin Cities, and is advised by Dr. David W. Johnson.

### **What if I have more questions?**

You may ask any questions you have now. If you have questions later, **you are encouraged** to contact Suzanne Russ or Dr. David Johnson at the University of Minnesota, or Dr. Linda Tollefsrud at UW Barron County. Suzanne can be contacted at 920-606-4664, or by email at [russ0227@umn.edu](mailto:russ0227@umn.edu). David can be reached at [insert David's preferred phone], or by email [insert David's preferred email]. Dr. Linda Tollefsrud can be contacted at 715-234-8176 or by email at [linda.tollefsrud@uwc.edu](mailto:linda.tollefsrud@uwc.edu).

If you have any questions or concerns regarding this study and would like to talk to someone other than the researcher(s), **you are encouraged** to contact the Research Subjects' Advocate Line, D528 Mayo, 420 Delaware St. Southeast, Minneapolis, Minnesota 55455; (612) 625-1650.

**Appendix H (cont'd)**

**Statement of Consent**

*You will be given a copy of this information to keep for your records on the day of the study.*

1. What if I don't want to participate?

---

---

2. What if I come to participate, but then change my mind and want to stop?

---

---

3. Will anybody know what answers I give?

---

---

4. What risks or benefits are there?

---

---

5. What if I have more questions?

---

---

I have read the above information. I have asked questions and have received answers. I consent to participate in the study.

Student signature: \_\_\_\_\_ Date: \_\_\_\_\_

Signature of Investigator: \_\_\_\_\_ Date: \_\_\_\_\_