
Theses and Dissertations

Summer 2010

Controlling personality tendencies: predicting observer-rated personality from the interaction between general mental ability and self-rated personality

Jonathan Andrew Shaffer
University of Iowa

Copyright 2010 Jonathan Andrew Shaffer

This dissertation is available at Iowa Research Online: <http://ir.uiowa.edu/etd/737>

Recommended Citation

Shaffer, Jonathan Andrew. "Controlling personality tendencies: predicting observer-rated personality from the interaction between general mental ability and self-rated personality." PhD (Doctor of Philosophy) thesis, University of Iowa, 2010.
<http://ir.uiowa.edu/etd/737>.

Follow this and additional works at: <http://ir.uiowa.edu/etd>



Part of the [Business Administration, Management, and Operations Commons](#)

CONTROLLING PERSONALITY TENDENCIES: PREDICTING OBSERVER-
RATED PERSONALITY FROM THE INTERACTION BETWEEN GENERAL
MENTAL ABILITY AND SELF-RATED PERSONALITY

by

Jonathan Andrew Shaffer

An Abstract

Of a thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Business Administration
in the Graduate College of
The University of Iowa

July 2010

Thesis Supervisor: Professor Frank L. Schmidt

ABSTRACT

Research has determined that measures of general mental ability (GMA) and personality are valid predictors of a wide range of work outcomes. Two of the most well established findings in the field of organizational psychology are that GMA and two of the Big Five personality traits, conscientiousness and emotional stability, predict overall job performance and training performance across all jobs. Though both GMA and personality are valid predictors of job performance, the validities of personality measures are much weaker than those observed for measures of GMA.

Some argue that personality may play a larger role in predicting work outcomes than currently believed, but that current measures of personality do not capture the construct fully. Several researchers have attempted to increase the validity of personality measures by altering the items in the measures so that they refer specifically to work contexts, and others have examined the validity of observer ratings of personality. This study draws on the theory of cognitive buffering to test the possibility is that GMA itself that causes the impact of personality traits on real life performances to be limited. That is, that people may use their GMA to control the expression of their personality tendencies in their behavior.

The results showed that GMA and personality interacted to predict peer ratings of personality, but not as initially hypothesized. Self-monitoring and personality also interacted to predict peer ratings of personality, but, again, not as hypothesized.

Several possible explanations for the results of this study are discussed, including the notion that that individuals may make efforts to manage only those personality traits that are most relevant in given situations. Moreover, it may be the case that dispositions

are less subject to the process of cognitive buffering than are emotions and affect.

Limitations of this study and opportunities for future research are also discussed.

Abstract Approved: _____

Thesis Supervisor

Title and Department

Date

CONTROLLING PERSONALITY TENDENCIES: PREDICTING OBSERVER-
RATED PERSONALITY FROM THE INTERACTION BETWEEN GENERAL
MENTAL ABILITY AND SELF-RATED PERSONALITY

by

Jonathan Andrew Shaffer

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Business Administration
in the Graduate College of
The University of Iowa

July 2010

Thesis Supervisor: Professor Frank L. Schmidt

Copyright by

JONATHAN ANDREW SHAFFER

2010

All Rights Reserved

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Jonathan Andrew Shaffer

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Business Administration at the July 2010 graduation

Thesis Committee:

Frank L. Schmidt, Thesis Supervisor

Kenneth G. Brown

Amy E. Colbert

Steven B. Robbins

Thomas E. Vaughn

For Melodie

ABSTRACT

Research has determined that measures of general mental ability (GMA) and personality are valid predictors of a wide range of work outcomes. Two of the most well established findings in the field of organizational psychology are that GMA and two of the Big Five personality traits, conscientiousness and emotional stability, predict overall job performance and training performance across all jobs. Though both GMA and personality are valid predictors of job performance, the validities of personality measures are much weaker than those observed for measures of GMA.

Some argue that personality may play a larger role in predicting work outcomes than currently believed, but that current measures of personality do not capture the construct fully. Several researchers have attempted to increase the validity of personality measures by altering the items in the measures so that they refer specifically to work contexts, and others have examined the validity of observer ratings of personality. This study draws on the theory of cognitive buffering to test the possibility is that GMA itself that causes the impact of personality traits on real life performances to be limited. That is, that people may use their GMA to control the expression of their personality tendencies in their behavior.

The results showed that GMA and personality interacted to predict peer ratings of personality, but not as initially hypothesized. Self-monitoring and personality also interacted to predict peer ratings of personality, but, again, not as hypothesized.

Several possible explanations for the results of this study are discussed, including the notion that that individuals may make efforts to manage only those personality traits that are most relevant in given situations. Moreover, it may be the case that dispositions

are less subject to the process of cognitive buffering than are emotions and affect.
Limitations of this study and opportunities for future research are also discussed.

TABLE OF CONTENTS

LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: LITERATURE REVIEW	6
General Mental Ability	6
Cognitive Buffering	8
Conceptualizing Expressive Control	14
Personality and Self-Observer Agreement	17
Trait Factors	17
Situational Factors	22
Agreeableness	24
Emotional Stability	27
The Moderating Effects of Time	28
Self-monitoring	32
GMA and Self-monitoring	38
CHAPTER 3: METHOD	43
Participants	43
Context	43
Measures	45
Independent Variables	45
Control Variables	48
Dependent Variables	50
Interaction Terms	52
Corrections for Range Restriction	52
Computing Confidence Intervals	55
Procedures	55
Analysis	57
CHAPTER 4: RESULTS	59
Agreeableness	59
Results for Time 1	59
Results for Time 2	61
Comparing the Interaction Terms	61
Emotional Stability	62
Results for Time 1	62
Results for Time 2	64
Comparing the Interaction Terms	64
Post-hoc Analyses: Conscientiousness	65
CHAPTER 5: DISCUSSION	86
Contextual Issues	87
Levels of Acquaintanceship	87
The Task-oriented Nature of the Project Teams	89

Theoretical Implications	90
Cognitive Buffering	90
Self-monitoring	93
Limitations	96
Future Research.....	99
Summary.....	101
APPENDIX: EMOTIONAL INTELLIGENCE.....	103
NOTES.....	107
REFERENCES.....	109

LIST OF TABLES

Table 1. Interrater reliabilities for observer ratings of personality.....	58
Table 2. Means, Standard Deviations, and Correlations for Study Variables.....	67
Table 3: Self-rated agreeableness, GMA, and self-monitoring on observer ratings of agreeableness.....	71
Table 4: Self-rated emotional stability, GMA, and self-monitoring on observer ratings of emotional stability.....	72
Table 5: Self-rated conscientiousness, GMA, and self-monitoring on observer ratings of conscientiousness.....	73

LIST OF FIGURES

Figure 1. Interaction between GMA and agreeableness at Time 1.	74
Figure 2. Interaction between self-monitoring and agreeableness at Time 1.	75
Figure 3. Interaction between GMA and agreeableness at Time 2.	76
Figure 4. Interaction between self-monitoring and agreeableness at Time 2.	77
Figure 5. Interaction between GMA and emotional stability at Time 1.	78
Figure 6. Interaction between self-monitoring and emotional stability at Time 1.	79
Figure 7. Interaction between GMA and emotional stability at Time 2.	80
Figure 8. Interaction between self-monitoring and emotional stability at Time 2.	81
Figure 9. Interaction between GMA and conscientiousness at Time 1.	82
Figure 10. Interaction between GMA and conscientiousness at Time 2.	83
Figure 11. Interaction between self-monitoring and conscientiousness at Time 1.	84
Figure 12. Interaction between self-monitoring and conscientiousness at Time 2.	85

CHAPTER 1: INTRODUCTION

Research has determined that measures of general mental ability (GMA) and personality are valid predictors of a wide range of work outcomes. Two of the most well established findings in the field of organizational psychology are that GMA and two of the Big Five personality traits, conscientiousness and emotional stability, predict overall job performance and training performance across all jobs (Barrick & Mount, 1991; Hunter & Hunter, 1984; Hunter, 1986). Though both GMA and personality are valid predictors of job performance, the validities of personality measures are much weaker than those observed for measures of GMA. In fact, while GMA and personality measures are each useful individual predictors of job performance, the validity of GMA for predicting job performance is .65, while the validities of conscientiousness and emotional stability, the two personality traits that best predict performance, are only .22 and .12, respectively. In addition, path analyses based on true score correlations between GMA, personality, and job performance reveal that no personality trait provides substantial incremental validity above and beyond GMA (Schmidt, Shaffer, & Oh, 2008). The research findings related to GMA, personality, and performance are extremely robust. So robust, in fact, that among many industrial and organizational psychologists it is taken for granted that personality plays only a small role relative to GMA in predicting job performance.

However, not all agree that personality plays such a small role in the workplace. Some argue that personality may play a larger role than currently believed, but that current measures of personality do not capture the construct fully (Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, 2007). In keeping with this viewpoint, it

should be noted that the findings outlined above apply only to the validity of global, self-report personality measures, which some have argued are not ideal for use in predictive validity studies because they measure broadly defined personality as opposed to a more narrow, “workplace” personality. In keeping with this concern, several researchers have attempted to increase the validity of personality measures by altering the items in the measures so that they refer to specific contexts. Schmit, Ryan, Stierwalt, and Powell (1995) added “...at school” to each of the 48 items in the conscientiousness scale of the NEO-PI-R (Costa & McCrae, 1992) and used the revised items to predict the GPA of a large sample of college students. The authors found that the school-specific scale predicted GPA ($r = .41$) better than did the original scale ($r = .25$). Similarly, Gill and Hodgkinson (2007) designed a Big Five Personality measure that included scale anchors that were work-specific and tested its validity in three separate samples of salespersons, customer service workers, and warehouse staff. The results showed that in 14 of the 15 validity analyses the work-specific scale predicted job performance better than did global measures of personality.

At the same time, a few studies have examined the validity of observer ratings of personality. Socioanalytic theory (Hogan, 2007) distinguishes between an individual’s beliefs about their own disposition—their personality—and an outside observer’s view of that same individual—their reputation. The theory proposes that because a person’s reputation reflects their actual behaviors, and because past behavior is the best predictor or future behavior, reputation may be a better predictor than is personality. In a rigorous investigation of the validity of observer ratings of personality, Mount, Barrick, and Strauss (1994) examined this possibility in a sample of salespersons and found that for all

of the Big Five traits, observer reports of personality were a stronger predictor of supervisory job performance ratings than were self reports of personality. In fact, several of the Big Five personality traits that, when self-rated, are generally considered not to predict job performance across broad categories of jobs (e.g., agreeableness and extraversion) were strong predictors of supervisory performance ratings when reported by coworkers and clients (Oh, Wang, & Mount, 2010)

Another possible reason for the difference in the validities of GMA and personality is that linear models are not the best way to conceptualize the relationship between GMA and personality. More than five decades ago, a multiplicative model of job performance was proposed in which performance was hypothesized to be a function of ability and motivation (Maier, 1955). Maier's hypothesis and performance equation— $P = f(A \times M)$ —quickly drew interest, was tested in a variety of settings, and met with mixed results (Fleishman, 1958; French, 1958; Kipnis, 1962; Locke, 1965; Vroom, 1962). More recent research expanded Maier's original hypothesis to include a wide range of broad and narrow personality traits. This research, too, is replete with mixed results (Hollenbeck, Brief, Whitener, & Pauli, 1988; Mount, Barrick, & Strauss, 1999; O'Reilly & Chatman, 1994; Sackett, Gruys, & Ellingson, 1998; Wright, Kacmar, McMahan, & Deleeuw, 1995).¹

Although the literature currently provides mixed support for an ability-motivation or ability-conscientiousness interaction, some evidence still suggests that ability and personality may interact in way that that is meaningful to workplace outcomes. Borrowing from the conceptualization of personality that is proposed by socioanalytic theory—specifically, that a very real difference exists between an individual's personality

and their reputation—an unexplored theoretical possibility is that GMA itself that causes the impact of personality traits on real life performances to be limited. That is, people may use their GMA to control the expression of their personality tendencies in their behavior. For example, consider a dispositionally introverted employee in a job that requires meeting and dealing with a wide variety of people; that is, a job in which extroverted behavior is conducive to success. Such a person could use his or her GMA to learn to display extroverted behaviors, even though such behavior is not his or her natural or spontaneous inclination. This follows from the idea that GMA is an “all purpose tool” that can be used to solve a wide variety of problems (Jensen, 1998), including the problem of how to meet the social requirements of a job. The possibility that individuals can use their cognitive resources to control the expression of their personality has been raised by several researchers (Allik & Realo, 1997; Hogan, 2007; Schmidt, Shaffer, & Oh, 2008), but little, if any, empirical work has examined the extent to which people can and do use their GMA to override their personality tendencies in the interests of enhancing their reputation and gaining the rewards associated with that reputation.

The purpose of my dissertation is to draw on the theory of cognitive buffering as a foundation for testing the hypothesis that interactions between GMA and personality may manifest through a process in which individuals with higher levels of cognitive resources are able to consciously inhibit or suppress the expression of undesired personality tendencies and instead express behaviors that are desired or acceptable in a particular context. Specifically, I plan to examine the theory of cognitive buffering in relation to agreeableness and emotional stability in a team context in which individual performance evaluations are, in part, dependent on peer evaluations of performance. As explained in

the following sections, I expect ability and agreeableness and emotional stability to interact with GMA in such a way that individuals high in ability but low in agreeableness or emotional stability will control the expression of their natural personality tendencies in such a way as to elicit favorable personality judgments from their peers.

CHAPTER 2: LITERATURE REVIEW

General Mental Ability

The theoretical foundation of research claiming a general factor of mental ability can be found in Spearman's (1904) seminal work on the nature of intelligence. After administering various rudimentary measures of intelligence to several samples of school children and a sample of adults, Spearman applied an early form of factor analysis to his data with the result being that each of the measures loaded onto a single factor. Spearman concluded that this single factor could "provisionally be called... 'General Intelligence'" (p. 284). Thurstone (1931) disagreed with Spearman's theory of general intelligence and presented an alternative hypothesis in which he proposed that there were multiple intelligences that were independent of one another. However, Eysenck (1939), using more sophisticated factor analysis methods than had been previously available, concluded that Spearman was, in fact, correct. Since then, the general factor theory of intelligence has been overwhelmingly accepted among psychometricians. And, although some researchers have attempted to revive theories of multiple intelligences (e.g. Gardner, 1983; Sternberg, 1986), such theories have met with little support from serious academics, largely because they fail to hold up under empirical scrutiny as, again and again, tests that measure specific facets of intelligence tend to correlate with a single, general factor. As Jensen (1998) puts it, "Provided the number of tests in the analyzed battery is sufficiently large to yield reliable factors and the tests are sufficiently diverse in items types and information content to reflect more than a single narrow ability, a g factor always emerges" (p. 73).

The g factor, or GMA, has been shown to correlate with a number of physiological processes. Biological studies have provided evidence that GMA is correlated with physiological processes in the brain. Jensen (1998) reviews studies of brain functioning that show that individuals high in GMA have more efficient and active brainwave patterns. Individuals high in GMA also metabolize blood glucose as fuel for brain activity more efficiently than do individuals low in GMA.

GMA has also been found to have some degree of impact on virtually every important life outcome that has been measured. GMA is correlated to tangible life outcomes. Individuals high in GMA are less likely to experience long-term unemployment or poverty, be injured on the job, or be convicted of a crime. They are more likely to complete advanced degrees and to have fewer health problems (Herrnstein & Murray, 1994), and to earn higher salaries and achieve higher levels of occupational status (Judge, Ilies, & Dimotakis, 2010; Judge, Klinger, & Simon, 2010; Judge, Higgins, Thoresen, and Barrick, 1999). Individuals high in GMA are also more likely to be capable of engaging in positive interpersonal interaction (Gottfredson, 1997).

Although empirical research has gone far in illuminating the correlates of GMA, "...data and analysis can take us only so far in saying what intelligence is. At some point, it becomes a matter of definition" (Herrnstein, 1971; p. 49). The definition that I adopt for the purposes of the present study is one that was agreed upon by 52 leading researchers in the field of intelligence and psychology and was published by the Wall Street Journal in 1994. The article defines GMA as "a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely

book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings—‘catching on,’ ‘making sense’ of things, or ‘figuring out’ what to do” (p. A18). This definition of GMA is important because, as Hunter (1986) explains “...performance of any kind is primarily dependent on learning. Since data shows that all complex learning is predicted by general cognitive ability (which they call intelligence), performance in all complex tasks will be closely predicted by general cognitive ability” (p. 346).

Cognitive Buffering

Certainly, the process of controlling the expression of one’s trait-driven behavioral tendencies in an effort to display behaviors that are appropriate for various social situations can be very complex. Exhibiting appropriate behaviors in a given situation calls for an understanding of the demands of the situation, the social cues given by others, one’s own feelings and natural urges, what actions or reactions are most appropriate in the situation, and which behaviors will successfully communicate one’s intentions. Although a seemingly daunting task, Snyder (1974) notes that “some social observers have proposed that the ability to manage and control expressive presentation is a prerequisite to effective social and interpersonal functioning” (p. 526), and the ability of some individuals to control the outward expression of their personality, affect, or emotion has long been a subject of interest in experimental and clinical psychology. Psychologists have theorized that a process of cognitive buffering may allow individuals to suppress temporarily their natural instincts and reactions in order to consider reasonably all possible responses to a given stimulus and the potential consequence of those actions (Bjorklund & Harnishfeger, 1995). Diamond, Blavin, and Diamond (1963) explain that

during inhibition, a person's "unconscious (facilitative) associate processes throw up suggestions, and that conscious (inhibitory) processes accept or reject them" (p. 173). Perhaps because the inhibition of natural emotional or behavioral expression has been conceptualized as a complex, cognitive process, individual differences in GMA have been targeted as the main mechanism of inhibition.

For a simple example of the cognitive buffering process, consider a football quarterback. A quarterback must ignore the threat of physical injury in order to be successful. When, on a given play, a quarterback consciously chooses not to run from or try to avoid a defensive player in order to attempt to throw the football, the quarterback is engaging in the process of controlling his or her instinctive reactions to flee from danger. Indeed, when a quarterback waits until the last second to throw the ball and is then immediately hit by a defensive player, he or she is lauded by sports analysts for such behavior. Likewise, it is considered heroic when, in wartime, soldiers perform their duties regardless of any feelings of fear that they might experience. In both of these examples, in order to perform their jobs, individuals have to first ignore their emotions (the fear of pain or death) and the behaviors associated with those emotions (running away or refusing to follow orders). Second, they must choose to display behaviors that are desirable based on the requirements of their occupation.

The examples used above are relatively uncomplicated but are relevant, as much of the research in the area of cognitive buffering has largely focused on the suppression or inhibition of behaviors associated with negative affective traits (such as aggression and anxiety) or cognitive processes (such as undesired, irrelevant thoughts). In a series of five experiments, Gernsbacher and Faust (1991) found that subjects higher in general

comprehension skill—defined as the propensity to understand and recall information—were better able to suppress or inhibit cognitive interference in a series of verbal and nonverbal problem sets. Rosen and Engle (1998) performed two experiments in which subjects participated in several speeded word pairing tasks in which they had to filter out irrelevant information while completing word pairs. Using working memory capacity (WMC), a construct highly correlated with GMA (Conway, Kane, & Engle, 2003), as a predictor they found that individuals higher in WMC not only learned word lists faster, but were more able to suppress irrelevant information that threatened to interfere with their ability to correctly match word pairs. Brewin and Beaton (2002) instructed subjects not to think about a white bear and to tell the researchers each time the thought of a white bear occurred. The researchers measured WMC and fluid intelligence (gF) and found that WMC and gF significantly predicted thought suppression such that subjects higher in WMC or gF suppressed thoughts of a white bear more effectively than did low WMC or gF subjects. In a related experiment, Brewin and Smart (2005) directed participants to recall an undesired, unpleasant thought after which each participant engaged in three five-minute sessions of continuous verbalization of their thoughts. The results yielded further evidence that WMC is positively related to thought suppression.

The studies reviewed above deal with the suppression of superfluous information and specific thoughts in isolated experimental tasks, but there also exists a sizable stream of research dealing with GMA and more salient emotional processes, such as post traumatic stress disorder (PTSD). Findings from the field of clinical psychology consistently show that individuals high in GMA are less at risk for developing PTSD. Breslau, Lucia, and Alvarado (2006) reported that in a longitudinal sample of children

and adolescents, higher GMA was correlated with a lower risk of PTSD. However, the authors also found that subjects higher in GMA were also at a lower risk for experiencing traumatic events, which was not controlled for in this study. McNally and Shin (1995) improved on this design by controlling for exposure to traumatic events in their study of 105 Vietnam combat veterans. In this study, reports of PTSD symptoms were correlated positively with combat exposure ($r = .41$) and negatively with GMA ($r = -.33$), and after the authors controlled for combat exposure, the correlation between GMA and PTSD was slightly stronger ($r = -.35$). One limitation of this study is that the authors did not measure GMA until after combat exposure had occurred. Thus, they state that their study does not allow them “to determine whether current cognitive performance is an antecedent or consequence of PTSD” (p. 937). Macklin and colleagues (1998) overcame the aforementioned design flaws and extended the findings of McNally and Shin by including measures of pre- and post-combat GMA in a similar study of 90 Vietnam veterans. They reported that after controlling for combat exposure, the partial correlation between pre-combat GMA and PTSD was $-.33$. One interpretation of these results may be that individuals higher in GMA are more able to inhibit or suppress feelings of stress and anxiety and thus report fewer symptoms of PTSD. Likewise, such individuals may also be better able to choose constructive coping behaviors when they feel stress or anxiety that help relieve their symptoms. Another possibility is that individuals higher in GMA are more likely to recognize and seek professional help for their condition, which may result in an actual reduction in symptom occurrence. In any case, the findings from these studies suggest that individuals high in GMA are better equipped to deal with negative emotional events.

In two recent field studies, Perkins and Corr (2005, 2006) examined the relationship between GMA and emotional stability in the workplace, drawing on the cognitive buffering literature to hypothesize that aspects of low emotional stability may be beneficial under certain conditions, especially if the individual in question has the mental ability necessary to choose constructive behavioral expressions that are appropriate for a work setting. This proposition makes sense if one considers some of the potential positive benefits of worry and anxiety, mental states commonly associated with low levels of emotional stability (McCrae & Costa, 1987). Luu, Tucker, and Derryberry (1998) state that “normal levels of anxiety are required to infuse the appropriate degree of caution, anticipation, and self-discipline in many behavior situations” (p. 577-578) and further explain that anticipatory anxiety may be a beneficial stressor that helps people assess dangers and risks and produces a sense of motivation to avoid such dangers. If low levels of emotional stability or its facets can produce a general sense of caution and motivation, then it may be the case that low levels of emotional stability are not necessarily harmful, but that unmanageable, extremely low levels of emotional stability are. Even if an individual is highly neurotic, if they also possess the ability to effectively control or redirect their neurotic impulses, they may be able to perform their job at acceptable, and even exceptional, levels.

Perkins and Corr (2005) tested their hypothesis in a sample of 63 financial sector managers in which they measured GMA and trait anxiety (but not emotional stability, *per se*) and found that managers with the best performance were those high in both trait anxiety and GMA, while those high in trait anxiety and low in GMA had the worst performance. Perkins and Corr (2006) retested this hypothesis in two independent

samples of 607 British Naval officer and 62 British Army officers. In this study the authors used a Big Five measure of emotional stability and replicated their previous results. In this larger sample, Perkins and Corr found that the performance of officers with high GMA was unaffected by their self-reported level of emotional stability, while those officers low in emotional stability and low in GMA had the worst performance. The results reported by Perkins and Corr suggest that even when an individual is extremely low in emotional stability, if they also possess the ability to effectively control or redirect their neurotic impulses, they may be able to perform their job at acceptable, and even exceptional, levels. Exactly how this interaction between GMA and emotional stability occurs is not examined in these two studies. While it could be the case that those high in GMA and low in emotional stability simply suppress their natural tendencies, similar to the suppression and inhibition of thoughts that has been observed in experimental settings, it may instead be the case that individuals high in GMA and low in emotional stability choose behaviors that help them overcome their neurotic urges. Perhaps these individuals, because they become overly worried about missing deadlines, are better able to prioritize their projects or organize their schedules. And of course, both the suppression of negative behaviors and the forced expression of positive behaviors may occur simultaneously.

The empirical findings reviewed above seem to support the notion that some individuals are better able to suppress certain thought patterns or emotions than are other individuals, and that such suppression may be achieved through a process of cognitive buffering, the use of which is highly dependent on GMA. At the same time, theories of general intelligence propose that individuals high in GMA are better able to understand

and adapt to new information and situations. Gottfredson (1997) points out that this conceptualization of GMA extends to social situations that require a high degree of interaction with other people and reports data that indicate that interpersonal job requirements are very cognitively demanding. Thus, overall, the evidence reviewed in this section seems to suggest two things. First, individuals high in GMA may be better at concealing negative behavioral tendencies than are those low in GMA. Second, individuals high in GMA may be better able to assess what types of behavior are appropriate in a given context. If individuals high in GMA can both suppress negative traits and enhance their expression of positive traits, then this would cause the impact of GMA on performance to be much stronger than the impact of personality.

Conceptualizing Expressive Control

At this point, it is necessary to explain how I will operationalize the ability of some individuals to control the expression of their personality tendencies. Personality is most commonly operationalized in research through self-report measures. In studies of the validity of personality, the scores from self-report measures are typically correlated with some external criterion. This correlation is, in turn, interpreted as the validity of the self-reported personality scale or trait for predicting the criterion of interest.

The same procedure can be applied to observers' ratings of an individual's personality. For example, Mount, Barrick, and Strauss (1994) asked the coworkers, supervisors, and customers of 105 salespersons to complete personality measures for each of the salespersons and then correlated those scores with supervisory ratings of job performance. The authors found that the scores from the observer-ratings of personality were much more valid than the self-reported personality ratings of the salespersons for

predicting sales performance. For some personality traits, such as extraversion and agreeableness, self-reports measures did not meaningfully predict supervisory ratings of performance at all ($\rho = .09$ and $.07$, respectively), while customer ratings of extraversion and agreeableness were strong predictors of performance ($\rho = .38$ and $.42$, respectively). In this same study, coworker ratings of extraversion and agreeableness did not predict performance very well. Mount and colleagues conclude that one possible explanation for their findings is that salespersons, whose performance is more highly dependent on their relationship with their customers than on their relationship with their coworkers, may strive to display behaviors that present a favorable persona to the people whose impressions matter the most. If this is the case, then a salesperson who rates him or herself as low in agreeableness and is also rated low in agreeableness by coworkers may selectively behave in an agreeable manner when interacting with a customer, causing the customer to rate the salesperson as high in agreeableness.

The results of Mount et al. (1994) raise an important question: Which method of measuring personality is more accurate? That is, do self-report measures reflect a person's true personality better or worse than do other-reports of personality? The issue has been of interest since the 1920s, and two approaches have generally been used in conceptualizing the accuracy of personality judgments (Funder (1999) provides an extensive review of this literature). The first approach is rooted in behaviorism and defines the accuracy of personality judgments as the extent to which "...a personality judgment can predict the independently measured behavior of the person who is judged..." (Funder, Kolar, & Blackman, 1995). This view treats personality as a phenotypic phenomenon which must be observable to be considered real. The behaviorist

view might interpret the results of Mount et al. (1994) as indicating that the salespersons personality ratings provided by supervisors and customers were the most accurate measures of personality as they were the ratings that best predicted actual behavior.

The behaviorist approach has been largely abandoned over the last two decades, in part due to advances in the field of behavior genetics, and has given way to a second perspective. This second approach views personality as a fundamental individual difference that originates from a deeper level than simple, observable behavior. This approach sees personality as a genotypic, psychological construct that is stable over time (McCrae, 1993) and influences individual behavioral tendencies but does not wholly determine them. This paradigm views self-report measures as indicators of individuals' personality traits and uses correlational estimates of interjudge agreement as its main index of accuracy (Funder et al., 1995). According to the genotypic approach to accuracy, to the extent that self- and observer-ratings of personality are highly correlated, observer-ratings are considered accurate. Over time, the use of the correlation between self- and observer-rated personality has become the most accepted and widely used method for studying the accuracy of observer ratings of personality (e.g. Beehr & Watson, 2008; Funder & Dobroth, 1987; Watson, 1989).

For the purposes of the current study, because I am interested less in the actual accuracy of observer rating of personality and more in the absolute ratings themselves, I intend simply to use the observer ratings of personality as an indicator of behavioral control. Because scores on self-reported personality assessments are assumed to be based on a target's full range of affective and behavioral tendencies, it is generally assumed that self-report measures should reflect the target's true, genotypic self. It is precisely that

assumption under which I work in the remainder of this study. Observers of the target, however, do not have at their disposal the target's full range of feelings and behavioral choices. Observers are able to judge personality based only on the phenotypic information that the target makes available to them. Thus, if an individual is normally anxious and nervous but consciously conceals this fact by purposely engaging in behaviors not normally associated with anxiety or nervousness, the individual may rate him or herself high on anxiety, while observers of the individual would not rate him or her high on anxiety. In terms of broader personality traits, if the self-ratings of emotional stability, for example, are low for a given individual and observer-ratings of emotional stability for that same individual are high, I will consider that to be a potential indicator that the individual in question has concealed their low levels of emotional stability from the observer. The term commonly used to describe the similarities between self-ratings and observer-ratings of personality is self-observer agreement, and that is the term that I will use for the remainder of this paper.

Personality and Self-Observer Agreement

The main purpose of this study is to examine the extent to which GMA influences the expression of personality traits. The influence of GMA on that expression will supposedly be captured by the level of self-observer agreement in personality ratings. However, GMA is not the only factor to be considered in examinations of self-observer agreement. Below I discuss several other factors that can affect self-observer agreement.

Trait Factors

Theoretical work suggests that low self-observer agreement in personality ratings may be attributed to the fact that some personality traits are easier to observe and assess

than others. Funder's (1995) Realistic Accuracy Model (RAM) of personality judgment proposes that in order for accurate judgment of a personality traits to occur, the trait must first generate a behavior that is "relevant to that trait" (p. 658). That behavior, in turn, must be observable to the judge. Certain personality traits, because their associated behaviors are less observable in everyday behaviors, are associated with vast differences in self-observer personality ratings agreement. The behaviors associated with extraversion, for example, are more likely to emerge in a social setting than are the behaviors associated with low levels of emotional stability. Funder (1995) states, "...a trait like sociability, which is revealed by frequent positive social interaction, is easier to judge than a trait like ruminates and daydreams, which must be inferred from verbal statements or, even more ambiguously, from dreamy looks, distracted responses, and the like, any of which could have other meanings as well or instead" (p.662). Researchers have coined the term visibility to describe the degree to which the behaviors associated with a given personality trait are available for observation (Funder & Dobroth, 1987; Watson, Hubbard, & Wiese, 2000). A trait that has high visibility has several characteristics. First, it is easy for observers to conceive of behaviors that are representative of the trait. Second, the context of the interactions between the target and the observer gives rise to opportunities for the trait behaviors to be displayed. Third, the observer needs to witness relatively few instances of a trait behavior to infer the underlying trait.

Empirical findings support the idea that the visibility of a trait affects self-observer agreement (Ready, Clark, Watson, & Westerhouse, 2000). Funder and Dobroth (1987) and Funder and Colvin (1988) obtained ratings of perceptions of visibility of

personality traits from a sample of undergraduates and correlated the ratings with self-observer agreement for those same traits from a second sample. They found that the traits that were perceived to be more easily observable were also the traits for which self-observer agreement was highest. For extraversion, which was rated as the most visible of the Big Five traits, the correlation between self- and observer-ratings was .29 in both studies. In contrast, emotional stability was rated as the least visible Big Five trait. Self-observer agreement for emotional stability was -.53 (Funder & Dobroth, 1987) and -.32 (Funder & Colvin, 1988), indicating that to some extent, certain personality traits are difficult for even close friends to identify in each other. Beehr and Watson (2008), Borkenau and Liebler (1992), and Funder, Kolar, and Blackman (1995) reported similar findings. Across these studies extraversion consistently achieves the highest level of self-observer agreement among the Big Five traits (though it is still low), while emotional stability achieves the lowest. Overall the consistency of the results across studies gives strong support to the notion that some traits are more easily judged than are others.

In addition to the visibility of traits, a second factor that may explain the discrepancies between self- and observer-ratings of personality is that some personality traits are perceived as socially undesirable and thus kept hidden from observers (Funder, 1995). For example, an individual who is disagreeable by nature may wish to conceal this fact from his or her peers and will therefore refrain from displaying disagreeable behaviors. In this case, self-observer agreement in ratings of agreeableness may be low, but not because the behaviors associated with agreeableness are more difficult to recognize. Instead, the differences between self- and observer-ratings of personality may be caused by the conscious effort of the disagreeable individual to engage in more

socially acceptable behaviors. Thus, observers may accurately judge the behavioral cues that the target of judgment makes available to them, but the behaviors are not truly representative of the target's personality.

There is evidence from several studies that suggest the above may be more than just a possibility. Independent ratings of the desirability of personality traits are related to self-observer agreement such that the correlation between self- and observer-ratings is highest for traits that are perceived as desirable (Funder & Colvin, 1988; Funder & Dobroth, 1987; Ready et al., 2000). This line of research finds that low emotional stability, which is labeled as the most undesirable of the Big Five traits, consistently emerges as the trait on which there is the least agreement in self- and observer-ratings. In fact, the perception of low emotional stability seems to be so unfavorable that people may have stronger negative reactions to individuals that display low levels of emotional stability than they have positive reactions to individuals that display high levels of emotional stability, conscientiousness, or other favorable aspects of the Big Five traits (Funder & Dobroth, 1987). At the same time, self-observer agreement for extraversion and conscientiousness, two socially favorable traits, tends to be positive. In addition, self-observer agreement in ratings of agreeableness tends to be lower than for all other traits except emotional stability (e.g. Funder et al., 1995; Norman & Goldberg, 1966; Watson, 1989).

It stands to reason that individuals may make efforts to conceal their personality tendencies if such tendencies are perceived as socially undesirable. But what about traits that are perceived as desirable? That is, why would anyone conceal the fact that they are highly agreeable or emotionally stable? John and Robins (1993) addressed this issue and

found a curvilinear relationship between ratings of trait desirability and self-observer agreement such that agreement was lowest when ratings of traits were perceived as either desirable or undesirable and was highest when the desirability of a trait was perceived as neutral. In fact, when traits—such as agreeableness and conscientiousness—were perceived as either extremely desirable or undesirable, the correlation between self- and observer-ratings approached .00, indicating virtually no self-observer agreement. In contrast, when traits—such as extraversion—were perceived as neutral, the correlation of self- and observer-ratings reached between .30 and .40. The authors note that their findings suggest that self-reports of personality are influenced by social desirability. That is, if a respondent reports him or herself to be high on a favorable traits, such as agreeableness (as part of a self-deceiving process) while at the same time observers (based on observations of the respondent's actual behaviors) report that the respondent is low in agreeableness, self-observer agreement for this trait would be low. They state that this proposition “implies that under conditions of ego involvement self-judgments may be less accurate than the judgments of a well-informed other” (p. 548). While their data is insufficient to support such a conclusion (which the authors acknowledge) it does raise the question of the effects of social desirability and faking on personality assessments.

Social desirability and faking certainly may play a role in the agreement of self- and observer-ratings of personality, but probably not as John and Robins (1993) proposed. Although I do not expand on the topic in this proposal, research has largely determined that faking on personality assessments does not seem to affect their validity, even in the field of employment testing, a context in which the motivation to fake can be quite high (Ones & Viswesvaran, 1998). If faking on a personality assessment is not a

concern in high stakes settings, it is unlikely that faking is an issue in experimental settings. Instead of faking on the assessment itself, it may be the case, as Funder (1995) theorized, that individuals “fake” inasmuch as they choose not to display behaviors that are associated with social undesirable personality traits or they actively attempt to engage in behaviors that associated with positively viewed traits.

Overall, there is support for the notion that individuals will consciously make efforts to manage the influence of their personality traits in such a way as to display visible, observable behaviors that lead observers to form positive perceptions about them, and that the greatest efforts will be associated with personality traits that are considered the least favorable to observers. For example, because agreeableness is perceived as a highly desirable trait, individuals who are low in agreeableness should make efforts to display agreeable behaviors to their peers. In addition, because low agreeableness is difficult for people to perceive without the aid of specific behavioral cues (Funder & Dobroth, 1987), individuals who are low in agreeableness may be able to mask their true personality by engaging in behaviors that are associated with high levels of agreeableness.

Situational Factors

In addition to the nature of personality traits themselves, situational factors also affect the extent of the efforts that individuals make to conceal their true personality. Controlling one’s natural, trait-based behavioral inclinations may be more or less appropriate depending on the demands of the situation. For example, while disagreeable behaviors may be perceived as undesirable and may have a negative impact on the performance ratings of individuals with jobs that have high interpersonal demands, such

as salespeople or administrators, the same behaviors may be completely irrelevant to the performance ratings of those in jobs that require little contact with other people, such as scientific researchers or truck drivers. Personality researchers are aware of the effects that the situation may have on the validity of personality traits. Barrick and Mount (1991) found that conscientiousness was the best of the Big Five personality traits at predicting individual level performance across all types of jobs, but Stewart (2003) noted that research has shown that “teams create a setting requiring increased cooperation, thereby altering the impact of certain traits on individual performance” (p. 184).

Two examples of traits that are relatively unimportant to individual performance but are important to performance in a team are agreeableness and emotional stability. The meta-analytic correlation between agreeableness and individual job performance is quite low, reaching just above zero (Barrick & Mount, 1991; $\rho = .07$). Almost the same correlation was reported between emotional stability and job performance ($\rho = .08$). In contrast, the meta-analytic correlation between agreeableness and individual performance in a team setting was much higher ($\rho = .33$), as was that between emotional stability and performance ($\rho = .27$), suggesting that personality traits do indeed affect individual performance in teams in a unique way (Mount, Barrick, & Stewart, 1998). In fact, Mount and colleagues found that agreeableness and emotional stability predicted performance in teams better than did any other Big Five trait. A subsequent meta-analysis that included additional primary studies reported that agreeableness is the most important personality trait for predicting team performance ($\rho = .34$) (Barrick, Mount, and Judge, 2001). Furthermore, although agreeableness does not show any substantial incremental validity above and beyond GMA for predicting individual performance, in a team setting this is

not the case. Neuman and Wright (1999) found incremental validity for agreeableness and conscientiousness over and above GMA in a sample of 316 full-time employees (the authors reported the incremental validity for only those two traits).

Individual motivation to act in a manner that is perceived as promoting healthy team processes may be increased even further if multirater, or 360-degree ratings, are used to assess individual performance. In multi-source performance ratings, the performance ratings that an individual receives from his or her peers or team members factor into that individual's final performance evaluation. These types of performance evaluations have become more commonly used over the last 10 years (Ployhart, Schneider, & Schmitt, 2006) and offer a high degree of incentive for individuals to manage the impressions of not only their supervisor, but also their coworkers and work group members. Thus, when individual performance ratings are based on the evaluations that people receive from their peers, they should be motivated to display behaviors that elicit favorable performance evaluations from their peers.

Agreeableness

The current study will be conducted in a team context in which team members, working together over the course of sixteen weeks, rate one another's performance. In such a setting, agreeableness is an appealing personality trait to study in relation to cognitive buffering for several reasons. First, certain personality traits have characteristics that cause them to be perceived as more or less desirable in interpersonal contexts. At the same time, the behaviors associated with certain traits are more open to observation. Agreeableness not only is perceived as a desirable trait in social settings, but also is associated with easily recognizable behaviors that people can choose to exhibit.

However a social setting alone may not be enough to motivate individuals to attempt to control the expression of their personality traits. In order for cognitive buffering to occur there needs to be some incentive for individuals to present a favorable image to their peers. A project team setting in which peer evaluations factor into individual performance evaluations may help to generate just such an incentive, as individuals may attempt to manage the interactions that they have with the members of their work team in an effort to influence the performance evaluations that they receive from their team members. Below, I review why situational and trait factors make agreeableness a personality trait that is central to this study.

Digman (1990) explains that agreeableness “appears to involve the more humane aspects of humanity—characteristics such as altruism, nurturance, caring, and emotional support at the end of one dimension, and hostility indifference to others self-centeredness, spitefulness, and jealousy at the other” (p. 422-424). McCrae and Costa (1987) proposed that agreeableness, at its core, is most easily understood by its negative side-- disagreeableness. Of disagreeable people, they state, “Cognitively they are mistrustful and skeptical, affectively they are callous and unsympathetic, behaviorally they are uncooperative, stubborn, and rude” (p. 88). After such descriptions of agreeableness and disagreeableness, it is plain to see why agreeableness is the best of the Big Five traits for predicting performance in teams. Work done in teams requires higher levels of interpersonal interaction and cooperation than does work done individually. Even if the work tasks that team members complete is done on an individual level, interpersonal interaction between team members is required for planning and organizing workloads. The nature of the situation, therefore, demands a certain type of behavior. Team members

are expected to behave courteously and professionally toward one another, and are expected to engage in helping behaviors that benefit the team. “Thus, employees who are more argumentative, inflexible, uncooperative, uncaring, intolerant...are likely to have lower ratings of teamwork” (Barrick, Mount, & Judge, 2001; p. 12).

In addition, there is some evidence that indicates that while agreeableness is a highly favorable trait—second only to emotional stability (Funder & Dobroth, 1987)—as mentioned above, self-observer agreement tends to be quite low for agreeableness. Keeping in mind that previous research has found that for traits that are either extremely desirable or extremely undesirable self-observer agreement is close to zero, it is reasonable to believe that agreeableness is perceived to be so highly desirable that, in general, individuals low in agreeableness attempt to conceal, to some extent, that trait from others.

Taken as whole, individuals in team settings may attempt to display behaviors associated with agreeableness and eliminate behaviors associated with disagreeableness based in part on demands of the situation and in part based on the general social acceptability of agreeable behaviors regardless of the team context in which this study occurs. Because being high in agreeableness is not perceived as unfavorable, I expect that the cognitive buffering process will be the most salient in those individuals that are low in agreeableness. Therefore, I expect that individuals high in GMA and low in agreeableness will control the expression of their personality by engaging in behaviors that lead their peers to perceive them as agreeable.

Hypothesis 1: Individuals high in GMA are more able than are individuals low in GMA to control the expression of their personality. Specifically, when self-

reported levels of agreeableness are low, observer ratings of agreeableness will be higher for individuals high in GMA than for those low in GMA.

Emotional Stability

Emotional stability is also a trait of interest in this study. As is the case with agreeableness, emotional stability is perceived as a desirable trait that is associated with behaviors that, while they are perhaps at times more subtle than are behaviors associated with agreeableness, are generally recognizable behaviors that people can choose to exhibit. Below, I review why situational and trait factors make emotional stability a personality trait that I examine in this study.

First, as mentioned previously, emotional stability is a personality trait that is associated with strong social judgments. Specifically, the negative side of emotional stability—neuroticism—is perceived to be the most undesirable personality trait that a person can have (Funder & Dobroth, 1987). McCrae and Costa (1987) described individuals high in neuroticism as “...worrying, insecure, self-conscious, and temperamental” (p. 86). As I argued in the case of agreeableness, it may be the case that low emotional stability is perceived to be so highly undesirable that individuals low in emotional stability attempt to hide this trait from others. Such attempts at concealment of emotional stability may help explain why emotional stability consistently emerges as the Big Five personality trait for which self-observer agreement tends to be lowest. Simply put, the social stigma attached to being labeled neurotic may motivate people to hide their neuroticism, making it difficult for observers to accurately assess their true personality. In terms of the situational constraints that working in a team places on the extent to which individuals display behaviors associated with emotional stability, Barrick, Mount, and

Judge (2001) state that "...being calm, secure and not depressed or hostile (high in emotional stability), should result in more effective interactions with co-workers or customers" (p. 12). Results reported by Barrick and colleagues (Barrick, Stewart, Neubert, & Mount, 1998) support this notion, and show that emotional stability at the team level correlates with team performance ($r = .24$), and has even stronger correlations with social cohesion ($r = .53$), team conflict ($r = -.42$), team flexibility ($r = .41$), and team communication ($r = .48$) than did any other Big Five trait. The authors conclude that emotional stability leads to more positive team interactions, social cohesion, and long-term viability.

Just as I expect that individuals in team settings may attempt to conceal behaviors associated with disagreeableness, I also expect individuals low in emotional stability to conceal their personality from their team members. Because being low in emotional stability is perceived as unfavorable, the cognitive buffering process should be most relevant to those individuals that are low in emotional stability. Therefore, I expect that individuals high in GMA and low in emotional stability will control the expression of their personality by engaging in behaviors that lead their peers to perceive them as emotionally stable.

Hypothesis 2: When self-reported levels of emotional stability are low, observer ratings of emotional stability will be higher for individuals high in GMA than for those low in GMA.

The Moderating Effects of Time

As argued above, there is much evidence that self-observer agreement varies based on the nature of the personality trait in question and the nature of the situation in

which the behaviors of the target are being observed. In addition to these considerations, there is one factor that has shown strong moderating effects on the relationship between self-observer agreement in virtually every study in which it has been measured. That moderator is extremely simple and straightforward: The observer's amount of exposure to the target of the personality judgments. Put succinctly, as an observer's exposure to a given target person increases, self-observer agreement increases. In layman's terms this makes perfect sense; the more time people spend together, the better they know each other.

This also makes sense in light of trait based theories on self-observer agreement. Certain traits are more difficult for observers to judge for two main reasons. First, because the visibility of some traits is low, it may take many interactions between a target and an observer for the observer to be exposed to the behaviors associated with a given trait. Second, because some traits are more socially undesirable than others, an observer may need to view a target in several different contexts over time in order to gather enough information to make an accurate assessment of the target's personality. As a result, self-observer agreement should become higher as the length of exposure between the two individuals increases. This should especially be the case for affective traits, such as low agreeableness and low emotional stability, for which the behaviors that are associated with a person's true personality are more easily hidden (Watson, Hubbard, & Wiese, 2000).

The moderating effects of the length of the exposure between the target and the observer, or, the acquaintanceship effect (Watson, Hubbard, & Wiese, 2000), have been shown in several studies to have a strong impact on self-observer agreement—especially

in the case of the self-observer agreement on emotional stability and agreeableness. Norman and Goldberg (1966) found that self- and observer-rating from a sample of undergraduate psychology students that completed rating of their classmates on the first day of class showed moderate correlations for extraversion ($r = .38$), openness to experience ($r = .32$) and conscientiousness (.34), while the correlation for agreeableness was lower (.15) and that for emotional stability was almost zero (.02). In an independent sample of Peace Core volunteers who had spent three months in close contact with one another, the correlations were all higher ($r = .54, .45, .47, .32, \text{ and } .27$, respectively).

Waston, Hubbard, and Wiese (2000) computed self-observer agreement from three sets of dyads: friends, dating couples, and married couples. In this study, each dyad consisted of two people that had considerable interaction with one another but with varying degrees of intimacy. Self-observer agreement was positive for all of the dyads, but was higher for dating couples than for friends, and higher for married couples than for dating couples. This pattern was more prominent in the case of emotional stability and agreeableness, which showed the greatest increase in agreement from friends to married couples. One finding of further interest from this study is that self-observer agreement on four negative affect scales—fear, sadness, guilt, and hostility—was similar for both friends and dating couples (mean $r = .24$ and $.26$, respectively), but was substantially higher for married couples (mean $r = .45$). In contrast, self-observer agreement on four positive affect scales was more similar in all three studies (mean $r = .33, .35, \text{ and } .43$). Overall, these findings not only suggest that the length and intimacy level of a relationship may affect self-observer agreement, but that individuals may indeed make efforts to conceal negative traits from others.

Finally, a recent meta-analysis by Connolly, Kavanagh, and Viswesvaran (2007) reported large differences between self-observer agreement levels for strangers, peers, and close relatives.² Their results show that self-observer agreement is lowest when the personality of the target individual is rated by strangers, higher when rated by peers, and highest for close relatives. Specifically, self-observer agreement between targets and strangers was lowest for agreeableness ($\rho = -.01$) and emotional stability ($\rho = .08$). Self-observer agreement for these two traits increased considerably when rated by peers ($\rho = .44$ for agreeableness and $.42$ for emotional stability) and close relatives ($\rho = .51$ and $.69$, respectively). Connelly and Ones (in press) report a similar pattern of results for agreeableness ($\rho = .23$, $.60$, and $.91$ for strangers, friends, and family, respectively) and emotional stability ($\rho = .22$, $.63$, and $.80$ for strangers, friends, and family, respectively). The reader should note that these correlations are higher than those previously reported because Connelly and Ones corrected for interrater reliability in observer ratings of personality, which are quite low.

In keeping with the current literature on self-observer agreement in personality ratings, I expect that when individuals rate one another's personality, self-observer agreement should increase as the amount of time the individuals spend together increases. Because high agreeableness is associated with behaviors that are easily identifiable and low agreeableness is associated with behaviors that can be concealed, it is possible that individuals low in agreeableness can lead others to believe that they are high in agreeableness by refraining from disagreeable behaviors and instead engaging in agreeable behaviors. Over time, however, it becomes more likely that individuals will display behaviors that are representative of their true personality, leading observers to

adjust any personality judgments that they had previously made. For the same reasons, I expect the same pattern for levels of self-observer agreement in ratings of emotional stability. In terms of the interaction of GMA with the personality traits in question, the process of using one's cognitive resources to maintain a desirable image should become more difficult as observers are exposed to a greater range and frequency of target behavior on which to base subsequent personality judgments. Thus, over time, the effects of cognitive buffering on observer ratings of the target's personality should weaken. Said another way, as observers gain more exposure to the target of their personality ratings, the interaction between personality and GMA should become a weaker predictor of self-observer agreement, and the self-rated personality of the target should become a stronger predictor of self-observer agreement.

Hypothesis 3: Self-observer agreement in ratings of agreeableness and emotional stability will increase as the observer's exposure to the target person increases.

Self-monitoring

As the above review indicates, cognitive ability has been a common independent variable in studies of inhibition and expression. Although this study certainly focuses on the importance of GMA as a predictor of personality expression, there is one other construct that is particularly relevant to a discussion about individual differences in the ability to control the expression of personality—self-monitoring.³ In the following section, I review self-monitoring and its relationship to cognitive buffering and personality expression and advance several hypotheses focusing on self-monitoring.

Although the main purpose of this study is to test the hypothesis that GMA is a cognitive tool through which people can control the outward expression of their

personality traits, the literature provides several theoretical constructs upon which competing hypotheses can be based. One of these is the construct of self-monitoring. Snyder (1974) first introduced the construct of self-monitoring, defining it as the extent to which people accurately assess the demands of social situations and respond appropriately to those demands. Essentially, self-monitoring is a process by which individuals “out of a concern for the situational appropriateness of their expressive self-presentation, have come to monitor their expressive behavior and accordingly regulate their self-presentation for the sake of desired public appearances” (Gangestad & Snyder, 2000; p. 530).

“The goals of self-monitoring may be (a) to communicate accurately one’s true emotional state by means of an intensified expressive presentation; (b) to communicate accurately an arbitrary emotional state which need not be congruent with actual emotional experience; (c) to conceal adaptively an inappropriate emotional state and appear unresponsive and unexpressive; (d) to conceal adaptively an inappropriate emotional state and appear to be experiencing an appropriate one; (e) to appear to be experiencing some emotion when one experiences nothing and nonresponse is inappropriate” (Snyder, 1974; p. 527).

High self-monitors, therefore, in order to manage the perceptions that others have of them, actively evaluate their environment and engage in a variety of behaviors in order to present themselves favorably to observers in a given social situation. Low self-monitors, on the other hand, choose their behaviors based less on the cues of their social environment and more on their true selves. For these low self-monitors, “expressive behaviors are not controlled by deliberate attempts to appear situationally appropriate; instead, their expressive behavior functionally reflects their own inner attitudes, emotions, and dispositions” (Gangestad & Snyder, 2000; p. 531).

Organizational psychologists have hypothesized that high self-monitors should receive higher job performance ratings due to their increased likelihood of managing impressions or modifying their behavior to suit job demands across various situations (Day & Schleicher, 2006; Miller & Cardy, 2000). However, meta-analytic results show that the direct effect of self-monitoring on job performance is low. Day, Schliecher, Unckless and Hiller (2002) reported a corrected correlation between self-monitoring and job performance of only .10. However, there are several reasons to believe that these findings are hardly definitive. First, the inclusion criteria for job performance were very broad. Included in the analysis were studies reporting objective sales performance, number of promotions, and subjective ratings of job performance. It could be the case the self-monitoring has a greater impact on subjective measures of job performance than on objective measures of performance, as high self-monitors may be able to influence their supervisors' impressions of them but not be able to influence their actual performance on the job. Second, Day et al., (2002) consider only the linear effects of self-monitoring on job performance. It may be the case that self-monitoring interacts with another work-related variable to predict job performance.

Neither of the two shortcomings of the Day et al. (2002) meta-analysis has gone unnoticed in the literature, as there has been interest in both the predictive validity of self-monitoring for more specific facets of job performance and the interaction between self-monitoring and other individual differences. Warech, Smither, Reilly, Millsap, and Reilly (1998) collected measures of self-monitoring and supervisor ratings of job performance from 191 managers from the telecommunication industry. Their results showed that while self-monitoring predicted supervisor ratings of interpersonal skills, it did not predict

ratings of business competence. Likewise, in a sample of 102 Executive MBA students, Barrick, Parks, and Mount (2005) reported that self-monitoring had a weak positive correlation with supervisor and peer ratings of interpersonal performance but a weak negative correlation with supervisor or peer ratings of actual task performance

While these studies shed light on some of the consequences of self-monitoring, they do not reveal empirically how self-monitoring has its effects. In the only study that I located that included an interaction between personality and self-monitoring, Barrick, Parks, and Mount (2005) found that self-monitoring and personality interacted to predict interpersonal performance such that interpersonal performance was highest for those individuals that were either high on self-monitoring and low on traits relevant to interpersonal performance—extraversion or agreeableness—or low on self-monitoring and high on extraversion or agreeableness. These results indicate that individuals high in extraversion or agreeableness may have high interpersonal performance when their level of self-monitoring is low, but that individuals high in self-monitoring can make up for the personality deficiencies in extraversion or agreeableness. Although these results do not allow for a definitive conclusion about this interactive process to be drawn, one possibility is that high self-monitors who were low in extraversion or agreeableness consciously chose to engage in the behaviors associated with these traits.

In a series of four studies, Flynn, Reagans, Amanatullah, and Ames (2006) reported that high self-monitors were more likely to be able to accurately judge the status of individuals in a social network and were also more likely to engage in helping behaviors directed at higher status peers. They also found that high self-monitors tended to be in positions of higher status than were low self-monitors. Again, although no

definitive conclusion can be drawn from these studies, the authors suggest that while high-self monitors seem to give of themselves to others “such generosity may not be entirely altruistic” (p.1134). It is, thus, conceivable that high self-monitors were more likely to help others based on the instrumentality of the helping behaviors in terms of their careers, and not based on either a high or low level of trait altruism.

It is relevant at this point to note the extent to which self-monitoring and agreeableness and emotional stability are related. While no comprehensive meta-analysis has been conducted on the relationship between self-monitoring and agreeableness or emotional stability, the data seem to suggest that self-monitoring is unrelated or is weakly related to either Big Five trait. The literature that I reviewed generally showed that the uncorrected correlation between agreeableness and self-monitoring ranged from $-.02$ to $-.27$. That between emotional stability and self-monitoring ranged from $-.10$ to $.16$ (Ashton & Lee, 2005; Barrick et al., 2005; Keller, 1999; Morrison, 1997; Tett, Fox, & Wang, 2005). In terms of the broader relationship of self-monitoring to the five-factor model of personality, self-monitoring has been linked theoretically to extraversion, as both constructs encompass the desire to achieve status through social interactions. Empirical work supports this notion, showing that self-monitoring and extraversion are at least moderately related. The uncorrected correlation between self-monitoring and extraversion ranged from $.18$ to $.55$ (Ashton, & Lee, 2005; Barrick et al., 2005; Flynn et al., 2006; Keller, 1999; Morrison, 1997; Tett et al., 2005; Warech et al., 1998).

Because self-monitoring is defined as the ability to recognize situational demands and to then express behaviors relevant to specific contexts, it follows that in certain situations, high self-monitors may override a natural inclination towards engaging in a

behavior that is inappropriate in a given situation and will instead choose to display a behavior that will be perceived in that situation as positive. This explanation can apply not only to the expression of emotions, but to moods and tendencies that stem from personality traits. For example, because emotional stability is perceived as a desired trait, a highly neurotic and self-monitoring individual is more likely to hide the emotions (anger), the moods (irritability), and the tendencies (hostility) that are normally associated with low levels of emotional stability. Efforts to control behavioral expressions should be especially elevated in work group contexts in which individual performance ratings are based on peer evaluations. Although the influence of personality traits on behaviors is strong, for high-self monitors the motivation to make a good impression on group members may supersede their natural behavioral tendencies. Thus, if high self-monitors know that they will be evaluated by their peers, they should be more likely to display to their group members behaviors that are associated with positive group interactions and performance and conceal any natural behavioral tendencies that stem from personality traits that are associated with negative group interactions.

Hypothesis 4: Individuals high in self-monitoring are more able than are individuals low in self-monitoring to control the expression of their personality. Specifically, when self-reported levels of agreeableness are low, observer ratings of agreeableness will be higher for individuals high in self-monitoring than for those low in self-monitoring.

Hypothesis 5: When self-reported levels of emotional stability are low, observer ratings of emotional stability will be higher for individuals high in self-monitoring than for those low in self-monitoring.

GMA and Self-monitoring

At the same time, it is important to recall that central to the definition of self-monitoring is the ability to recognize and respond to social cues. While this definition seems somewhat straightforward, social cues are not always readily interpretable. In addition, even if one assumes that social cues are interpretable, the process of generating a set of possible behavioral responses and then selecting the most appropriate response for the situation is not a simple task. There is potential for a misstep at any point in the process. One might perceive a situation correctly, but fail to generate a viable set of solutions from which to choose a behavioral response. Alternately, one might generate a set of response that includes the most appropriate response but fail to identify which, of all the potential response, is the most optimal in a given situation. Thus, a high self-monitor both understands his environment and chooses acceptable behaviors in response to that environment. A low self-monitor, on the other hand, may not comprehend his social environment and thus not be able to respond appropriately or may understand the environment but choose an inappropriate response regardless of the situation. (Gangestad & Snyder, 2000).

It should be clear at this point that the processes encompassed by the self-monitoring construct may be quite complex depending on the situation. In addition, parts of the process—generating a set of potential responses to a situation, for example—may be quite abstract processes. Taken, one step further, if self-monitoring essentially refers to the ability to deal with complex and abstract social and relational situations, then it may have much in common with GMA. Indeed, Gottfredson (1997) writes that GMA is “not the amount of information people know, but their ability to recognize, organize,

update, select, and apply it effectively” (p. 93). Gottfredson also reports results from decades of accumulation of data from the Position Analysis Questionnaire that show that interpersonal job tasks require a high degree of mental ability. In this study, both the requirements of dealing with people and the extent of the use of behavioral information in such dealings were strongly correlated with overall mental difficulty ($r = .59$).

Gottfredson concludes that “This should not be surprising, because other individuals are among the most complex, novel, changing, active, demanding, and unpredictable objects in our environment. Living and working with others is a complicated business” (p. 107).

Although self-monitoring and GMA may seem somewhat related on a theoretical level, only one meta-analysis has examined the correlation between the two constructs (Day et al., 2002). Their results seem to suggest that the two are only weakly related ($r = .07$) and differentially predict a number of work related outcomes. It should be noted, however, that results for the relationship between ability and self-monitoring should be interpreted with caution for a number of reasons. First, Day and colleagues used a data set that included 10 studies that reported correlations between GMA and self-monitoring for use in their meta-analysis. Of these 10 studies, only 6 included actual measures of GMA. Problem-solving performance, school performance, and skill level were also coded as ability measures. Second, they corrected only for unreliability in measures of self-monitoring, ignoring the unreliability of measures of GMA and the effects of range restriction. Thus, it is likely that the true score correlation between GMA and self-monitoring is higher than that reported by Day and colleagues.

Part of the reason for the low correlation between GMA and self-monitoring might be the result of the difference between how self-monitoring and GMA are

measured. GMA is generally measured with an assessment that captures an actual outcome, such as the number of correct answers on a test of verbal or mathematical ability. That is, GMA assessments yield a concrete indication of actual cognitive ability. Measures of self-monitoring, on the other hand, do not necessarily measure the extent to which an individual successfully exercises expressive control. Instead, they measure an individual's perception of their own ability to exercise expressive control. The distinction is important because research has shown that in many cases people do not evaluate their own abilities accurately. In general, meta-analyses have shown that individuals generally are able to make more accurate assessments of their abilities in areas in which they are able to compare easily their abilities to those of others. For example, the uncorrected correlation between an individual's athletic ability and their perception of their athletic ability has been reported to be .48 (Mabe & West, 1982). One's athleticism is easy to compare to that of others if one regularly participates in athletic activities. Likewise, because individuals who complete standardized tests such as the ACT, SAT, GMAT, GRE, or LSAT are almost immediately given access not only their raw scores but also their standing on the exam relative to other who have completed the assessment (given in the form of percentiles) they are quite accurate at assessing their own cognitive ability. Postlethwaite (2009) reported a correlation of .52 between actual cognitive ability test performance and estimated cognitive ability. Self-estimates of so-called soft skills, in contrast, were not as strongly correlated with actual measures of those abilities. For example, Mabe & West (1982) reported that the correlation between perceptions of interpersonal skills and actual interpersonal skills was only .17, and that between perceived and actual managerial skill was .08 (these correlations are all uncorrected for

measurement error). It stands to reason, then, that one factor that has an impact on the correlation between GMA and self-monitoring is that GMA is measured with an actual assessment of cognitive ability, while self-monitoring is measured by assessing the perception of self-monitoring.

In summary, on one hand at the conceptual level GMA and self-monitoring seem difficult to disentangle completely. On the other hand there seems to be only a small, but positive, relationship between GMA and self-monitoring. Even the difference in the methods in which the two constructs are measured would suggest that they are not the same. Add this to the fact that the validity of GMA for predicting work outcomes such as job performance (Schmidt, Shaffer, and Oh, 2008) and leadership (Judge, Colbert, & Ilies, 2004) is much higher than the validity of self-monitoring for the same outcomes (Day et al., 2002), further suggesting that the two constructs are somewhat distinct. In the current study, this translates into an expectation that GMA and self-monitoring will differentially predict self-observer agreement in personality ratings. Because the process of cognitive buffering has been conceptualized as being wholly cognitive in nature (as opposed to dispositional), I expect that the interaction between GMA and personality will be a better predictor of self-observer agreement than will be the interaction between self-monitoring and personality.

Hypothesis 6: The interaction between GMA and agreeableness will predict self-observer agreement in ratings of agreeableness more strongly than will the interaction between self-monitoring and agreeableness.

Hypothesis 7: The interaction between GMA and emotional stability will predict self-observer agreement in ratings of agreeableness more strongly than will the interaction between self-monitoring and emotional stability.

CHAPTER 3: METHOD

Participants

Participants were 310 undergraduate students at the University of Iowa who were enrolled in an introductory management course. I chose to use this sample for this study for a number of reasons. First, students in this particular course are assigned to work in teams which remain intact throughout the course of the semester and complete four assignments. The team assignments factor heavily into their final course grade, therefore, it was my estimation that students would be motivated to perform well in their teams. Second, individual performance on the team assignments is based, in part, on peer evaluations. This should have increased participants' motivation to engage in behaviors that were perceived favorably by their team members. Third, students were assigned to project teams randomly and the teams remained intact throughout the course of the semester, which allowed me to examine the moderating effects of time and acquaintanceship on the agreement between self- and observer-ratings of personality. Most of the teams consisted of four students, though some of the teams consisted of only two or three students.

Context

An important aspect of this study is the fact that while team members were all enrolled in the same discussion section throughout the course of the semester, after the initial team meeting teams met outside of class on their own time to work on their projects as needed. While teams were required to attend classes with each other throughout the course of the semester, data collected from past courses indicates that the amount of time that team members spend together outside of class varies greatly between

teams. In addition, some teams communicate mostly through email, while other teams meet face-to-face (Bradley, Postlethwaite, Burns, & Brown, 2008). Also, some team members may have found it more convenient to communicate via email with certain team members while communicating face-to-face with others. As a result, I expected considerable variance in the amount of time that team members spent with one another. At the same time, the course in which the study occurs was structured such that some class meetings were set aside for meetings between project teams. Thus, even those teams that communicated mostly through email or phone conversations had to attend class sessions in which they met together as a group.

Another important contextual feature is the fact that it was possible that some members within a team would spend more time with certain team members than with others. While teams were randomly constructed, it was possible that some team members may have known each other for quite some time before the construction of their project team. This was due to the fact that while students could not choose their team members, they could choose the discussion section in which they enrolled. Therefore, it is reasonable to assume that some participants chose to enroll in the same discussion section as that of someone they knew well. This would not have guaranteed that such friends would have placed in the same project team, but it certainly would have increased the probability of previously acquainted individuals being assigned to the same team.

One potentially beneficial feature of this study that was not designed as a part of this study but was instead a normal part of the course from which I collected data is this: Project teams had the option of firing poorly performing team members. Team members who were fired still had to complete the same assignments that they would have had to

complete if they were still a part of their original team, though any work completed by an individual who was fired from a team was worth a maximum of only 80% of its original raw score value in regards to the final semester grade. Said another way, individuals who are fired from a project team had to complete the same amount of work as an entire team but received fewer points for that work. Historically, about 1% of the individuals enrolled in the course are fired from their team each semester (about 4 to 6 people each semester). I estimate that this number was slightly higher (about 1.5%) during this semester in which I collected data for this study. I perceive this feature of the course to be beneficial because it creates an incentive for individuals to present a favorable impression of themselves to other team members. While the total number of individuals who usually were fired from their team was not high, I believe that it was sufficient to help create an atmosphere in which students were highly motivated to perform well on the team projects.

Measures

Independent Variables

GMA. The Wonderlic Personnel Test (WPT) was used. It is a 50-item assessment that measures verbal and mathematical ability. The assessment is timed and has a 12 minute time limit. WPT scores are strongly correlated with other measures of GMA, such as the WAIS and GATB (correlations range from .70 to .92).

In the domain of cognitive ability and aptitudes, item-specific measurement error and transient measurement error have generally been believed to be small, leaving random error as perhaps the largest source of measurement error in measures of GMA. However, Schmidt, Le, and Ilies (2003) showed that transient error may indeed be a

concern in the domain of cognitive ability. Their results showed that coefficient alpha may overestimate the reliability of measures of cognitive ability by as much as 6.7%. Thus, the coefficient of stability and equivalence is the most appropriate reliability estimate to use when estimating the reliability of the WPT, because it corrects for random, item-specific, and transient error.

According to the WPT test manual, the assessment has a test-retest reliability of between .82 and .94. However, the time periods between test administrations are not reported. Reliability estimates for test-retest with parallel forms are between .73 and .95, but again, the length of the time period between testing is not reported. I used the average of the estimates of test-retest reliability for parallel forms given in the WPT manual, .84. This estimate is similar to the meta-analytic estimates of .81 given by Hunter (1986) for the General Aptitude Test Battery and to the estimate of .85 given by Salgado and colleagues for a variety of GMA measures (see Table 1 in Salgado, Anderson, Moscoso, Bertua, & De Fruyt, 2003).

Personality. Saucier's (1994) 40-item scale was used to measure both self- and observer-ratings of personality. Because each participant had to complete a personality assessment twice for each of their team members and once for themselves (for a total of seven assessments), I chose to use a measure that was minimally onerous and complex. Saucier's measure uses single word adjectives as items instead of full sentences. Not only do I believe that single word items were faster for the participants to complete, but also it may have been easier for participants to think about their team members in descriptive rather than behavioral terms. Items are rated on a nine-point Likert scale. Example items include "bashful", "imaginative", and "kind".

While Saucier's (1994) measure is short, it has displayed reliabilities that are comparable to those of other Big Five measures. Meta-analyses report coefficient alphas of between .76 and .80 for agreeableness (Barrick & Mount, 1991; Hough, 1992; Hurtz and Donovan, 2000; Salgado, 1998; Salgado, 2003), which is similar to the reliability of .75 reported by Saucier (1994). While the coefficient of equivalence takes into account neither transient error—which is generally small in the personality domain (Schmidt, Le, & Ilies, 2003)—nor scale-specific factor error—which is generally quite large in the personality domain (Le, Schmidt, & Putka, 2009)—this is the estimate that is most commonly used in the literature, perhaps because it is the most widely available estimate.

Transient error, although small in the domain of personality, does affect the reliability of personality measures. In order to appropriately correct for unreliability in the agreeableness measure, I started with the reliability estimate of .76 for agreeableness that was reported by Hough (1992) as it is calculated from the most comprehensive dataset yet published on this topic ($k = 87$; $N = 22,060$). Schmidt, Le, and Ilies (2003) estimated that transient error further reduced the reliability of agreeableness scores by an average of .03. Thus, used a final reliability estimate of .73 (.76 - .03) in all corrections for unreliability in agreeableness scores. For the analyses that include emotional stability, I started with a reliability estimate of .76 (Hough, 1992) and combined this with findings from Schmidt, et al. (2003) to obtain a final reliability estimate of .67 (.76-.08).

Self-monitoring. Self-monitoring was assessed using Snyder's (1974) 25-item scale. Sample items include "I would probably make a good actor," and "I may deceive people by being friendly when I really dislike them." In order to correct for unreliability in the self-monitoring measure, I used the estimate of coefficient alpha of .71 that was

reported by Day et al. (2002) as it is calculated from the most comprehensive dataset currently available. No estimate of transient error is available for this measure, thus while using coefficient alpha will result in an undercorrection for measurement error, it is the best estimate currently available.

Control Variables

Previous acquaintance. As mentioned above, some participants may have spent more time with one another than did others, and some may have been more well acquainted than were others. Thus, it was important to determine the extent to which team members knew one another and to control for any relationships that team members might have had prior to the start of the current study. To do so, I used the Personal Acquaintance Measure (PAM) developed by Starzyk, Holden, Fabrigar, and MacDonald (2006) to measure the depth of relationships between individuals.

The full PAM consists of 18 items designed to assess the extent to which respondents know the target of the questionnaire and to which the respondent spends time with the target. 3 of the 18 items deal with the physical intimacy between the respondent and the target. In order to present the participants of this study with a measure that was as innocuous and unobtrusive as possible, these 3 items were omitted in this study. Test-retest reliability (with three weeks between administrations) of this measure has been reported to be quite high. In fact, Starzyk et al. (2006) report it to be .97. However, the sample size from which this reliability estimate was computed was only 99. Because the sample size of the current study was much higher and because I had dropped three items from the original measure, I used the current data to compute the reliability estimates for this measure. In this dataset, $\alpha = .83$ and $.86$ at time 1 and time 2, respectively. It is important

to note that to obtain final PAM score for a given individual, I averaged all of the within-group PAM scores for that individual. This method yielded a single score for each participant.

It is important to note that PAM scores may change over time, because the scale assesses the extent to which the rater is acquainted with the target of the ratings at the time at which the measure is administered. Three of the items on the PAM are designed to assess long-term acquaintanceship and, therefore, responses to these three items are unlikely to change over the duration of the current study. An example of these items is: "I have known [target] for many years." The other 12 items that were used in this study assess shorter-term acquaintanceship. It is possible that responses to these items can change over the course of just a few months. Examples of these items include: "Seeing [target] is a part of my weekly routine." and "[target] has told me about his/her interests."

Motivation. For reasons explained above, I expected that participants will be highly motivated to perform well in their project teams. However, it is possible that participants who have the ability to control the expression of their personality tendencies may not choose to do so if they are not motivated to present a favorable impression to the other members of their team. At the same time, it may be the case that a participant is motivated to perform well in some aspect of their academic work, but is not motivated to perform well in the class in which this study will take place and will thus not make efforts to obtain favorable performance evaluations from their team members. Therefore, while it is important to control for motivation, it is equally important to distinguish between participants' motivation to perform well in the particular class in which this study occurs from their more global sense of motivation. Thus, I controlled for

motivation in this study using two of three scales from a measure of motivation developed by Elliot and Church (1997). These scales measure the extent to which participants were concerned with a) performing poorly in the class (avoidance motivation) and b) performing well in the class (approach motivation). Examples of items from this questionnaire include “I worry about the possibility of getting a bad grade in this class.” and “I am striving to demonstrate my ability relative to my peers in this class.” Elliot and Church reported alpha reliabilities of .77 and .91 for avoidance and approach motivation scales, respectively. However, as with the PAM, the dataset from which these estimates were derived had a smaller sample size ($N = 178$) than does the current study. In the current dataset, $\alpha = .81$ and $.90$ for the avoidance and approach scales, respectively.

Dependent Variables

Observer ratings of personality. When a scale with multiple items is administered to a single rater (and whether that rater is rating themselves or another person), individual items from that scale are combined into a single score, usually by summing or averaging the items contained in the scale. As it relates to the current study, Viswesvaran, Ones, and Schmidt note: “In the context of ratings, a single rater is analogous to an item on a test. The use of several raters is analogous to use of a multiple-item test or scale” (2005: 109). Said another way, when several observers rate a target individual on a particular dimension (e.g., job performance, physical attractiveness, or personality) the ratings of each observer can be treated as an individual item and can be combined into a single rating for that target. With this in mind, I calculated the observer-rated agreeableness and emotional stability of each target individual in my dataset by treating the scores provided

by each observer for these traits as individual items. For each target, these scores were then averaged across all observers for that individual. The average agreeableness and emotional stability scores for each individual target were then used as the dependent variables in all further analyses. In some cases, a given target was only rated by a single observer. In such cases, I retained the scores provided by that rater for use in my analyses. The average number of observers that rated each target was 2.00 at Time 1 and was 1.90 at Time 2.

Because the dependent variables were measured in the form of ratings made by several observers of a single target, the appropriate method for correcting measurement error in these ratings was to use an estimate of the interrater reliability of the observers. I computed estimates of interrater reliability for the observer ratings of personality by first determining the correlation between scores on each trait for all rater pairs that rated the same target. That is, if raters A, B, and C all rated the same target I computed the correlation between pairs A-B, B-C, and A-C. I then averaged the correlations between each pair of raters within a given group and, finally, averaged the within-group results across all groups in my dataset. In computing these estimates, I used only the data from groups in which targets were rated by two or more raters (cases in which targets were rated by only one rater cannot be used to computing interrater reliability). This procedure yielded an estimate of the reliability of a single rater for each trait. I then used the Spearman-Brown formula (Nunnally, 1978) to adjust these initial estimates to account for the average number of raters for each target during each time period. The final interrater reliability estimates that were used to correct for measurement error in the criteria are shown in Table 1.

It worth noting that the above estimates of interrater reliability for a single rater are similar those given by Connelly and Ones (in press), who reported meta-analytically derived estimates of .20 and .22 for agreeableness and emotional stability, respectively, among “incidental acquaintances” (this category consisted of ratings made by classmates). I elected to use my own estimates of interrater reliability due to the fact that their analysis included a sample size ($K = 5$; $N = 338$) that was smaller than the sample size of the current study.

Interaction Terms

To obtain the interaction terms used in this study, I first standardized the independent variables (agreeableness, emotional stability, and self-monitoring) and the moderator variable (GMA) in order to minimize collinearity between these variables and the interaction terms. For each interaction, I computed the crossproduct of the independent variables and the moderator variable. I then computed the observed correlation between the crossproducts and each of my study variables. In order to perform analyses using true score correlations between all variables, it is necessary to correct the interaction terms between GMA and agreeableness (and GMA and self-monitoring) for measurement error. The formula for determining the reliability of a product term is given in Bohrnstedt and Marwell (1978) and is shown below. I used this formula to compute the reliability of all product terms in my analyses.

$$r_{xz} = [(r_{xx} \times r_{zz}) + r_{xz}^2] / (1 + r_{xz}^2)$$

Corrections for Range Restriction

Prior to performing any regression analyses, I also corrected all correlations for indirect range restriction in GMA. Relative to the normative data provided for the

Wonderlic Personnel Test ($M = 21.06$; $SD = 7.12$; $n = 118,549$) the sample in the current dataset showed considerable range restriction ($M = 25.24$; $SD = 4.71$). Range restriction is computed as $u_x = s_x / S_x$; that is, the ratio of the restricted to the unrestricted observed SDs of x . The initial u_x value that I obtained ($4.71 \div 7.12 = .63$) was subsequently used to compute u_t (which was used to make the final corrections for range restriction). Using the equation shown below (Hunter & Schmidt, 2006), where r_{xx_a} is equal to the reliability of the Wonderlic Personnel Test in the applicant population (in this case, $r_{xx_a} = .84$, as mentioned previously, I computed a value of .53 for u_t . This value was used to correct for range restriction in GMA.

$$u_t = \sqrt{\frac{u_x^2 - (1 - r_{xx_a})}{r_{xx_a}}}$$

I did not correct for range restriction self-monitoring. Thus was due to the fact that there was no evidence that range restriction existed in this dataset because no definitive set of normative data for self-monitoring exists. In fact, the SD of self-monitoring scores was not reported by Snyder (1974) when the scale was originally published. Subsequent studies of self-monitoring have used a variety of assessment lengths and scoring methods, making descriptive statistics from such studies difficult to compare to those of the current study.

I also performed no corrections for range restriction in personality. First, other researchers have noted that the range of personality scores is restricted very little, even among job applicants (Ones & Viswesvaran, 2003). Second, normative data for the measure of personality used in the current study was derived from a sample of college students that was only slightly larger than the sample contained in the current dataset

(Saucier, 1994; $n = 320$). Second, even if I assume that the original sample norms are representative of the population norms, I cannot conclude there was any range restriction in personality in the current sample. In fact, the use of Saucier's original data would lead to the conclusion that there was range enhancement in the current dataset. Saucier (1994) reported SDs of .69 and .84 for agreeableness and emotional stability, respectively. In the current dataset, the SD of agreeableness and emotional stability was .84 and .89, respectively. Again, I was hesitant to conclude that range enhancement existed in this dataset because I was not convinced that Saucier's dataset was anymore representative of the population than was the current dataset. Therefore, I made no corrections for range restriction in personality. Finally, for the reasons listed above, I also performed no corrections for range restrictions in the measures of approach or avoidance motivation. Using data from Eliot and Church (1997) as normative data would lead to the conclusions that there was a slight degree of range restriction in both approach ($u_x = 1.33/1.41 = .94$) and avoidance motivation ($u_x = 1.07/1.18 = .91$), however the current dataset is considerably larger than that of Eliot and Church ($n = 178$).

Finally, I did not correct for range restriction in the interaction terms. Though there was restriction of range in the interaction terms due to the fact that there was considerable restriction of range in GMA scores in my dataset, I was unable to determine the most appropriate method for applying range restriction corrections to the interactions terms. For informational purposes, I recomputed the reliability of the interaction terms as follows: I corrected the correlations between the independent variables and the moderator (GMA) for range restriction in GMA. This yielded slightly higher correlations between the independent variables and the moderator (r_{xz}) that I then used in the formula provided

by Bohrnstedt and Marwell (1978). Interestingly, the resulting reliability estimates were identical (within rounding error) to those that I had originally computed using the observed correlation between the independent variables and the moderator.

Computing Confidence Intervals

In order to construct confidence intervals around all the regression results for this study, it was necessary to compute the standard error for each of the variables on this study. Because I used the average PAM scores as a control variable and I also used the average of the observer-rated personality scores as the dependent variable of interest, I did not have equal sample sizes in all of the cells of the correlation matrix that was used as input in the regression analyses that I conducted. Therefore, in order to obtain a sample size with which to derive the standard error of my results, I computed the harmonic mean sample size for both Time 1 ($n = 285$) and Time 2 ($n = 236$). These means were used to construct confidence intervals around all of my results.

Procedures

Two weeks into the semester, the students were randomly assigned to groups and immediately completed an assignment that was designed to allow the new team members to get to know one another. The maximum amount of time that group members had to spend together at this time was approximately 30 minutes. Immediately following the initial group formation activity, participants had 24 hours to complete the first administration of observer reported personality measures for each of their team members. Participants also reported the extent to which they were previously acquainted with each of their team members.

Six weeks into the semester, students completed self-report measures of personality, self-monitoring, and goal orientation. Because personality traits are stable over time, the timing of the administration of the personality measures was chosen to minimize any impact that the administration of the measures will have on participant responses to subsequent study measures. Thus, students completed self-report measures at this time in order to separate the first and second collections of observer-ratings by enough time to attenuate any effects that completing the self-report measures will have on the participants' behavior in their groups.

Twelve weeks into the semester, after the third team project was completed, participants again completed observer reports of personality for each of their team members. At this point the teams had completed three of the four team assignments, and I expected that by this time students would have been able to view each other behaviors multiple times and would have potentially shown changes in their judgments of their peers' personality traits. At the same time, collecting observer-ratings of personality after the third team project was part of an attempt to keep the data free of the effects of any drastic behavioral changes that individuals might have displayed in an attempt to influence the fourth and final round of peer performance evaluations.

The timing of the administration of the measure of general mental ability was determined based on how well it fit with the topics being addressed in the management course. Therefore, fourteen weeks into the semester, I measured the GMA of the participants as part of a discussion session dealing with recruiting and selection.

Analysis

Once all data had been collected and standardized, I computed the interaction terms between GMA and personality and self-monitoring and personality.⁴ I then derived an observed correlation matrix that contained the observed correlations between all study variables and the interaction terms. All of the observed correlations were then corrected for measurement error and range restriction (the interaction terms were not corrected for range restriction, as discussed above). Finally, the true score correlation matrix was used as input in a series of hierarchical regression analyses in which the control variables (PAM, approach and avoidance orientations) were entered first, followed by the independent and moderator variables (personality, GMA and self-monitoring). The interaction terms were entered last. In order to interpret the results of the regression analyses, I plotted values for the dependent variables given high levels (+1.0 standard deviation) and low levels (-1.0 standard deviation) of the independent variables and the moderator variables (Aiken & West, 1991).

Table 1. Interrater reliabilities for observer ratings of personality

	<u>Time 1</u>		
	Reliability of a single rater	Average number of raters	Final reliability estimate
Agreeableness	.12	2.00	.21
Emotional Stability	.25	2.00	.40
	<u>Time 2</u>		
	Reliability of a single rater	Average number of raters	Final reliability estimate
Agreeableness	.23	1.90	.36
Emotional Stability	.21	1.90	.34

CHAPTER 4: RESULTS

Descriptive statistics and zero-order correlations for each study variable are shown in Table 2. Observed correlations are shown in the lower triangle of Table 2, while correlations corrected for measurement error are shown in the upper triangle. Though both observed and corrected correlations are reported, all further results and discussion provided in this paper are based on the corrected correlations.

Agreeableness

Results for Time 1

Hypothesis 1 predicted an interaction between GMA and agreeableness such that those high in GMA and low in agreeableness would be perceived by their team members as being high in agreeableness and that individuals low in GMA and low in agreeableness would be perceived by their team members as being low in that trait. Hypothesis 3 predicted that any interaction between GMA and agreeableness would weaken as observer exposure to the target increased. Hypothesis 4 predicted that those high in self-monitoring and low in agreeableness would be seen by their team members as being high in agreeableness and that individuals low in self-monitoring and low in agreeableness would be seen as being low in agreeableness. Finally, Hypothesis 6 predicted that the interaction between GMA and agreeableness would be a better predictor than would the interaction between self-monitoring and agreeableness. Results for each of these hypotheses are contained in Table 3. Model 1 in Table 3 shows the regression equation for each of the control variables and each of the main study variables. In Model 2, the interaction terms were added to the regression equation. Note that Models 1 and 2 give

only the results for Time 1. Models 1-T2 and 2-T2 correspond to Models 1 and 2, but show results for Time 2.

An initial examination of Model 2 shows that interactions between both GMA and agreeableness ($\beta = .25$; 95% CI = .15, .35) and self-monitoring and agreeableness ($\beta = -.13$; 95% CI = -.23, -.03) were detected at Time 1, which might lead to the conclusion that both Hypothesis 1 and 4 were supported. However, an inspection of the graphs for each of the interactions reveals that the interactions did not predict observer ratings of agreeableness as expected. Figure 1 shows both the predicted and actual results for the interaction between GMA and agreeableness. As predicted, individuals high in GMA were rated as being higher in agreeableness than were individuals low in GMA when their self-rated level of agreeableness was low. Individuals who were high in GMA and high in agreeableness received the highest observer-ratings of agreeableness. Unexpectedly, those low in GMA and high in agreeableness were perceived as being lowest in agreeableness.

Interestingly, a strong main effect for GMA was detected at Time 1 ($\beta = .75$; 95% CI = .63, .87), suggesting that, in general, individuals high in GMA were seen by their team members as being more agreeable. No main effect for self-rated agreeableness was found ($\beta = .06$; 95% CI = -.04, .16). Again, an interaction between GMA and self-rated agreeableness was detected, but the results were not as predicted. Thus, I concluded that Hypothesis 1 was not supported.

Similarly, Figure 2 shows the results for the interaction between self-monitoring and agreeableness. As mentioned above, an interaction was detected but the results were not as predicted. First, those high in self-monitoring were actually rated lower in

agreeableness than were those low in self-monitoring. Second, those low in self-monitoring and high in self-reported agreeableness were rated as being lower in agreeableness than were those low in self-monitoring and low in agreeableness. As shown below, individuals low in self-monitoring and low in agreeableness were perceived as being higher in agreeableness than were any other participants. Therefore, although an interaction was detected, I concluded that Hypothesis 4 was not supported.

Results for Time 2

Moving to the results from Time 2, Model 2-T2 indicates that at Time 2 there was a main effect for self-rated agreeableness ($\beta = .26$; 95% CI = .14, .38) and GMA ($\beta = .16$; 95% CI = .01, .31), but not for self-monitoring ($\beta = -.03$; 95% CI = -.15, .09). As shown in Figure 3, an interaction between GMA and agreeableness ($\beta = .14$; 95% CI = .01, .27) was detected. As seen in Figure 3, individuals high in GMA were rated higher in agreeableness than were those low in GMA, and individuals low in GMA were rated as such, regardless of their level of GMA. No interaction between self-monitoring and agreeableness ($\beta = .03$; 95% CI = -.09, .15) was detected at Time 2, as shown in Figure 4. Figure 4 shows that individuals low in agreeableness and individuals high in agreeableness were rated as such, regardless of their self-reported levels of self-monitoring. Taken as a whole, these results do not support Hypothesis 3.

Comparing the Interaction Terms

Finally, although at Time 1 the regression weight for the GMA by agreeableness interaction term ($\beta = .25$; 95% CI = .15, .35) was higher than that of the self-monitoring by agreeableness interaction term ($\beta = -.13$; 95% CI = -.23, -.03), in terms of the magnitude of the interactions (and not the direction) there was some overlap in the

confidence intervals for the two regression weights. At Time 2, the regression weights for the GMA by agreeableness interaction term ($\beta = .14$; 95% CI = .01, .27) was slightly lower than that of the self-monitoring by agreeableness interaction term ($\beta = .03$; 95% CI = -.09, .15), with considerable overlap between the confidence intervals for the two regression weights. Overall, these results do not support Hypothesis 6, which predicted that the regression weight for the interaction between GMA and agreeableness would better predict peer-ratings of agreeableness than would the interaction between self-monitoring and agreeableness.

Emotional Stability

Results for Time 1

Hypothesis 2 predicted an interaction between GMA and emotional stability such that those high in GMA and low in emotional stability would be perceived by their team members as being high in emotional stability and that individuals low in GMA and low in emotional stability would be perceived by their team members as being low in emotional stability. Hypothesis 3 predicted that any interaction between GMA and emotional stability would weaken as observer exposure to the target increased. Hypothesis 5 predicted that the interaction between self-monitoring and emotional stability would predict observer ratings of emotional stability, and Hypothesis 7 predicted that the interaction between GMA and emotional stability would be a better predictor than would the interaction between self-monitoring and emotional stability. Table 4 shows the results for each of these hypotheses. Just as with the results for agreeableness, Model 1 in Table 4 contains the control variables and each of the main study variables. In Model 2, the

interaction terms were added to the regression equation. Models 1 and 2 give the results for Time 1, while Models 1-T1 and 2-T2 show results for Time 2.

The results shown in Model 2 indicate that at Time 1 an interaction between GMA and emotional stability was not detected ($\beta = .04$; 95% CI = $-.05, .13$), which does not support Hypothesis 2. Figure 5 shows both the predicted and actual results of the regression analysis. The results indicate that those who had higher self-ratings of emotional stability also received higher observer-ratings of emotional stability. There was also a main effect for GMA ($\beta = .48$; 95% CI = $.37, .59$), with those higher in GMA receiving higher observer-ratings of emotional stability. Again, these results do not support Hypothesis 2.

On the other hand, Model 2 in Table 4 shows that while no main effects for self-monitoring were found ($\beta = -.09$; 95% CI = $-.18, .00$) an interaction between self-monitoring and emotional stability was detected at Time 1 ($\beta = -.33$; 95% CI = $-.43, -.23$), which initially seems to lend support to Hypothesis 5. However, inspection of the graph for this interaction reveals that it did not predict peer ratings of emotional stability as expected. This graph is shown in Figure 6, and shows that those low in self-monitoring received observer-ratings of emotional stability that corresponded to their levels of self-rated emotional stability. Those high in self-monitoring and low in emotional stability were perceived as having higher levels of emotional stability than were those low in self-monitoring and emotional stability, while those high in self-monitoring and emotional stability were perceived as having lower levels of emotional stability than were those low in self-monitoring but high in emotional stability. Again, an interaction between self-

monitoring and emotional stability was detected at Time 1, but the results were not as predicted. Thus, I concluded that Hypothesis 1 was not supported.

Results for Time 2

The results shown in Model 2-T2 in Table 4 indicate several things. First, though main effects were found for both emotional stability ($\beta = .30$; 95% CI = .16, .44) and GMA ($\beta = .26$; 95% CI = .11, .41), no main effect was found for self-monitoring ($\beta = -.02$; 95% CI = -.15, .11). Second, an interaction between GMA and emotional stability was not detected at Time 2 ($\beta = -.04$; 95% CI = -.17, .09). For informational purposes, these results are shown in Figure 7. Third, the results indicate that an interaction between self-monitoring and emotional stability was detected at Time 2 ($\beta = -.18$; 95% CI = -.32, -.04). As seen in Figure 8, the results are virtually identical to the results for Time 1. Specifically, those low in self-monitoring received observer-ratings of emotional stability that corresponded to their levels of self-rated emotional stability; those high in self-monitoring and low in emotional stability were perceived as having higher levels of emotional stability than were those low in self-monitoring and emotional stability; those high in self-monitoring and emotional stability were perceived as having lower levels of emotional stability than were those low in self-monitoring but high in emotional stability.

Comparing the Interaction Terms

Finally, at Time 1 the regression weights for the interaction between self-monitoring and emotional stability ($\beta = -.33$; 95% CI = -.43, -.23) was larger than the regression weight for the interaction between GMA and emotional stability ($\beta = .04$; 95% CI = -.05, .13). This was also the case at Time 2 ($\beta = -.18$; 95% CI = -.32, -.04 and $\beta = -.04$; 95% CI = -.17, .09 for self-monitoring and GMA, respectively). Notably, there is

some overlap between the confidence intervals surrounding the Time 2 estimates. These results do not lend consistent support to Hypothesis 6.

Post-hoc Analyses: Conscientiousness

Although the previous results failed to support the proposed hypotheses, I explored the possibility that conscientiousness might interact with GMA and self-monitoring. This exploration was informed by indirect evidence that project team members were motivated to display behaviors associated with conscientiousness (e.g., arriving to team meetings in a timely fashion, staying on task during meetings, meeting assignment deadlines) more than they were motivated to display behaviors associated with agreeableness or emotional stability. This evidence presented itself in the form of the peer evaluations that team members completed after each team project. The majority of the comments that were provided on the team members' performance evaluations indicated that behaviors associated with conscientiousness, specifically in the pursuit of high levels of task performance, were the behaviors that had the greatest impact on peer evaluations of performance. Therefore, I examined the interaction between GMA and conscientiousness and the interaction between self-monitoring and conscientiousness at both Time 1 and Time 2.

The results for conscientiousness are shown in Table 5. Models 1 and 2 show the results for Time 1, and Models 1-T2 and 2-T2 show the results for Time 2. There was a main effect for self-reported conscientiousness at Time 1 ($\beta = .16$; 95% CI = .05, .27) and at Time 2 ($\beta = .14$; 95% CI = .02, .26). There was also a main effect for GMA at Time 1 ($\beta = .54$; 95% CI = .42, .66) but not at Time 2 ($\beta = .10$; 95% CI = -.04, .24). In addition, the results indicate that GMA and conscientiousness interacted to predict peer ratings of

conscientiousness at Time 1 ($\beta = -.10$; 95% CI = $-.22, .02$), but not at Time 2 ($\beta = .02$; 95% CI = $-.12, .16$). As these results relate to Hypotheses 1 and 3, they lend support to the both hypotheses. These results are shown in Figures 9 and 10.

The results further indicate that a main effect for self-monitoring was found at Time 1 ($\beta = -.21$; 95% CI = $-.32, -.10$), but not at Time 2. Additionally, an interaction between self-monitoring and conscientiousness was detected at Time 1 ($\beta = .24$; 95% CI = $.14, .34$), and at Time 2 ($\beta = .17$; 95% CI = $.05, .29$). These results are shown in Figures 11 and 12. Although an interaction was detected at both Time 1 and Time 2, the results were not as expected. As shown below, at Time 1 and Time 2, individuals were rated by observers as having roughly the same level of conscientiousness regardless of their self-rated levels of conscientiousness or self-monitoring—with one exception. Unexpectedly, those high in self-monitoring and low in conscientiousness were rated as being lower in conscientiousness than any other participants. As these results relate to Hypotheses 4 and 5, they do not support those hypotheses.

Table 2. Means, Standard Deviations, and Correlations for Study Variables.

	M	SD	1	2
1 Approach motivation	5.51	1.18	.90	.10
2 Avoidance motivation	4.66	1.31	.09	.81
3 Agreeableness (AGR)	5.40	.97	.04	-.02
4 Emotional stability (ES)	4.12	.88	-.12	-.27
5 Conscientiousness (CON)	5.38	.92	.10	-.08
6 General mental ability (GMA)	25.24	4.71	.04	-.27
7 Self-monitoring (SM)	12.93	3.59	.11	.04
8 PAM T1	2.26	.42	.09	-.04
9 Peer-rated agreeableness T1	5.33	.69	-.04	.09
10 Peer-rated emotional stability T1	4.88	.67	-.03	.00
11 Peer-rated conscientiousness T1	5.29	.75	.00	-.05
12 PAM T2	2.41	.48	.08	.03
13 Peer-rated agreeableness T2	5.32	.76	.00	.08
14 Peer-rated emotional stability T2	4.90	.74	-.02	-.01
15 Peer-rated conscientiousness T2	5.15	.90	.10	-.05
16 GMAxAGR	.00	1.00	.08	.06
17 SMxAGR	.00	1.00	.00	-.06
18 GMAxES	.00	1.00	.06	-.05
19 SMxES	.00	1.00	.09	-.02
20 GMAxCON	.00	1.00	.10	-.01
21 SMxCON	.00	1.00	.08	-.03

Note: N = 109-310. PAM = Personal acquaintance measure. Observed correlations are shown in the lower diagonal. Correlations corrected for measurement error in the predictor and the criterion are shown in the upper diagonal. Reliabilities shown for peer ratings of personality are estimates of interrater reliability.

Table 2--continued

3	4	5	6	7	8	9	10
.05	-.16	.12	.06	.14	.10	-.09	-.04
-.03	-.37	-.10	-.41	.05	-.05	.22	.01
.73	.47	.46	-.20	-.05	.04	-.04	-.04
.33	.67	.36	-.07	-.20	.04	.12	.17
.33	.24	.69	-.06	-.16	.12	-.11	-.16
-.13	-.04	-.04	.84	.22	-.01	.28	.30
-.04	-.14	-.11	.14	.71	-.08	-.13	.06
.03	.03	.09	.00	-.06	.83	.20	-.39
-.02	.05	-.04	.09	-.05	.08	.21	.31
-.02	.09	-.09	.14	.03	-.23	.09	.40
-.09	.12	.06	.15	-.05	.11	.10	.07
.04	-.04	.05	-.04	-.05	.70	.07	-.17
.14	.04	-.02	-.02	-.01	.06	.13	.14
.06	.08	-.06	.07	.02	-.18	.10	.24
-.06	.02	.11	.02	-.08	.12	.09	.08
.14	-.02	-.07	-.14	.02	-.07	.01	.03
-.06	.03	-.09	.02	.00	-.03	-.03	-.05
-.02	.05	-.01	.07	-.08	-.03	.02	.03
.03	.13	-.03	-.08	-.10	-.14	-.09	-.08
-.08	-.01	.05	.06	-.01	-.02	.01	-.03
-.08	-.03	-.07	-.01	.04	.04	-.02	-.01

Table 2--continued

11	12	13	14	15	16	17	18
-.01	.09	.01	-.03	.14	.10	.00	.09
-.08	.04	.14	-.01	-.08	.08	-.09	-.07
-.16	.05	.26	.12	-.09	.21	-.10	-.03
.22	-.05	.08	.16	.03	-.03	.04	.08
.11	.06	-.04	-.13	.17	-.11	-.15	-.01
.32	-.05	-.05	.16	.04	-.25	.04	.12
-.10	-.06	-.02	.04	-.13	.03	.00	-.12
.20	.82	.11	-.34	.18	-.09	-.05	-.04
.35	.16	.47	.37	.27	.02	-.10	.06
.18	-.29	.37	.65	.17	.05	-.10	.06
.40	.13	.24	.35	.49	-.29	.04	.05
.08	.86	.36	-.07	.33	-.01	-.02	.00
.09	.20	.36	.51	.39	.16	.02	-.05
.13	-.04	.18	.34	.33	.15	-.05	-.04
.23	.22	.17	.14	.54	-.23	.03	.03
-.14	-.01	.07	.07	-.13	.62	.26	.53
.02	-.02	.01	-.02	.01	.15	.52	.22
.03	.00	-.02	-.02	.02	.31	.12	.56
-.07	-.06	-.08	-.07	-.10	.10	.26	.14
.00	.03	.01	-.03	.03	.38	.08	.19
.07	-.01	-.01	.02	.07	.05	.44	.02

Table 2--continued

19	20	21
.14	.14	.11
-.03	-.01	-.05
.04	-.13	-.14
.23	-.02	-.05
-.06	.07	-.12
-.16	.11	-.02
-.18	-.02	.06
-.22	-.03	.06
-.27	.03	-.08
-.19	-.06	-.02
-.17	-.01	.16
-.09	.04	-.02
-.20	.01	-.01
-.16	-.07	.05
-.21	.06	.14
.18	.64	.09
.04	.14	.86
.27	.33	.03
.48	.04	.48
.02	.58	.19
.23	.10	.50

Table 3: Self-rated agreeableness, GMA, and self-monitoring on observer ratings of agreeableness (AGR).

Variable	Time 1		Time 2	
	Model 1	Model 2	Model 1-T2	Model 2-T2
Personal acquaintance measure	0.23 (.13, .33)	0.25 (.17, .33)	0.35 (.23, .47)	0.36 (.24, .48)
Performance goal orientation	-0.17 (-.28, -.08)	-0.21 (-.31, -.11)	-0.06 (-.18, .06)	-0.08 (-.20, .04)
Avoidance goal orientation	0.57 (.47, .67)	0.56 (.46, .66)	0.19 (.05, .33)	0.21 (.08, .34)
Self-rated AGR	0.11 (.01, .21)	0.06 (-.04, .16)	0.28 (.16, .40)	0.26 (.14, .38)
General mental ability	0.68 (.57, .80)	0.75 (.63, .87)	0.12 (-.02, .26)	0.16 (.01, .31)
Self-monitoring	-0.28 (-.38, -.18)	-0.3 (-.40, -.20)	-0.02 (-.14, .10)	-0.03 (-.15, .09)
AGR X GMA		0.25 (.15, .35)		0.14 (.01, .27)
AGR X self-monitoring		-0.13 (-.23, -.03)		0.03 (-.09, .15)
Multiple R	0.66	0.69	0.47	0.49

Table 4: Self-rated emotional stability, GMA, and self-monitoring on observer ratings of emotional stability (ES).

Variable	Time 1		Time 2	
	Model 1	Model 2	Model 1-T2	Model 2-T2
Personal acquaintance measure	-0.38 (-.47, -.29)	-0.47 (-.56, -.38)	-0.05 (-.18, .08)	-0.07 (-.19, .05)
Performance goal orientation	-0.01 (-.10, .08)	0.06 (-.03, .15)	-0.03 (-.16, .10)	0.02 (-.11, .15)
Avoidance goal orientation	0.37 (.25, .49)	0.35 (.24, .46)	0.22 (.06, .37)	0.21 (.06, .36)
Self-rated ES	0.35 (.25, .45)	0.42 (.32, .52)	0.26 (.12, .40)	0.3 (.16, .44)
General mental ability	0.55 (.44, .66)	0.48 (.37, .59)	0.29 (.14, .44)	0.26 (.11, .41)
Self-monitoring	-0.06 (-.16, .04)	-0.09 (-.18, .00)	0.01 (-.12, .14)	-0.02 (-.15, .11)
ES X GMA		0.04 (-.05, .13)		-0.04 (-.17, .09)
ES X self-monitoring		-0.33 (-.43, -.23)		-0.18 (-.32, -.04)
Multiple R	0.63	0.69	0.3	0.35

Table 5: Self-rated conscientiousness, GMA, and self-monitoring on observer ratings of conscientiousness (CON).

Variable	Time 1		Time 2	
	Model 1	Model 2	Model 1-T2	Model 2-T2
Personal acquaintance measure	0.19 (0.09, 0.29)	0.17 (.07, .27)	0.31 (.19, .43)	0.32 (.20, .44)
Performance goal orientation	-0.07 (-0.18, 0.04)	-0.09 (-.20, .02)	0.12 (.00, .24)	0.09 (-.03, .21)
Avoidance goal orientation	0.18 (.06, 0.30)	0.22 (.10, .34)	-0.05 (-.19, .09)	-0.03 (-.17, .11)
Self-rated CON	0.12 (.01, .23)	0.16 (.05, .27)	0.12 (.00, .24)	0.14 (.02, .26)
General mental ability	0.49 (.37, .61)	0.54 (.42, .66)	0.08 (-.06, .22)	0.1 (-.04, .24)
Self-monitoring	-0.19 (-.30, -.08)	-0.21 (-.32, -.10)	-0.13 (-.26, .00)	-0.14 (-.27, .01)
CON X GMA		-0.12 (-.22, -.02)		-0.02 (-.14, .10)
CON X self-monitoring		0.24 (.14, .34)		0.17 (.05, .29)
Multiple R	0.48	0.53	0.40	0.43

Figure 1. Interaction between GMA and agreeableness at Time 1.

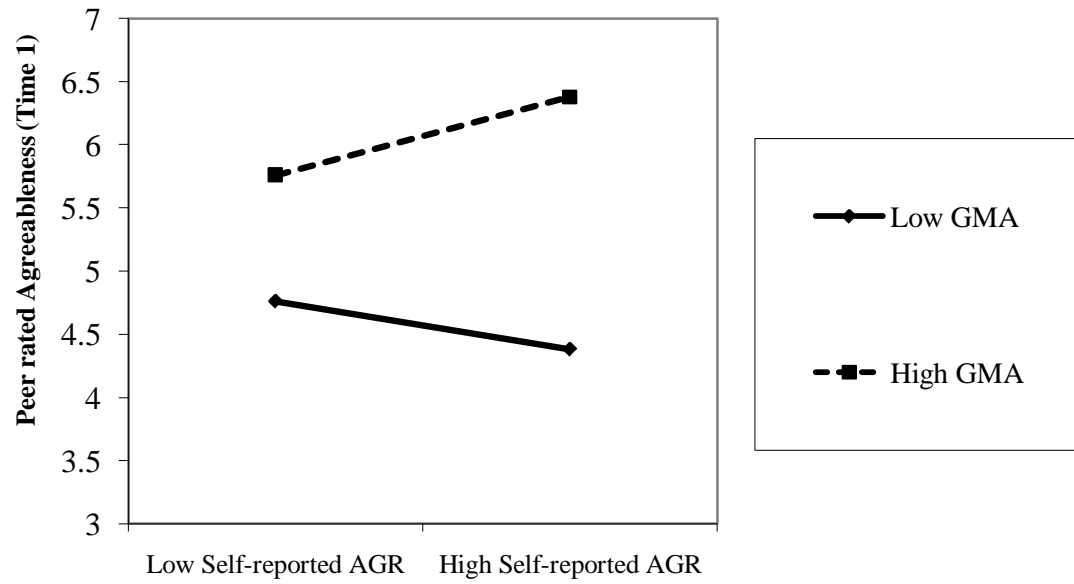


Figure 2. Interaction between self-monitoring and agreeableness at Time 1.

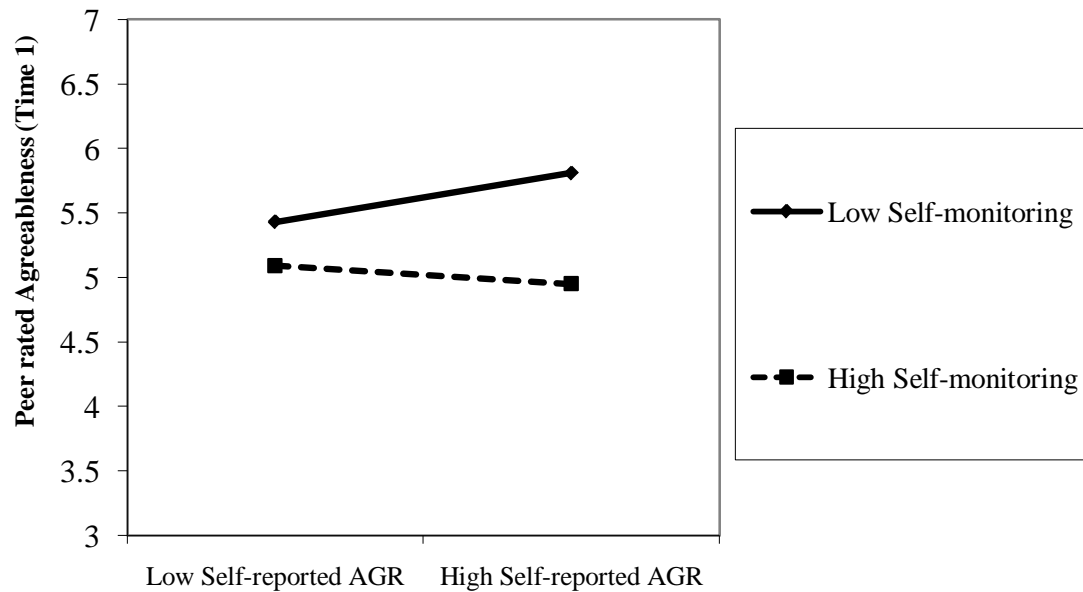


Figure 3. Interaction between GMA and agreeableness at Time 2.

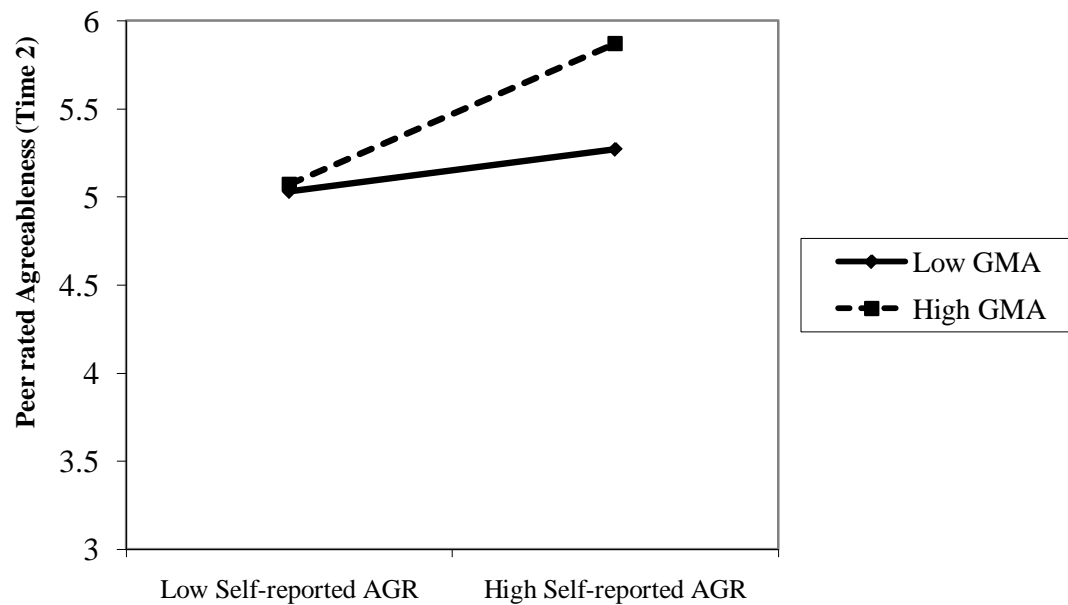


Figure 4. Interaction between self-monitoring and agreeableness at Time 2.

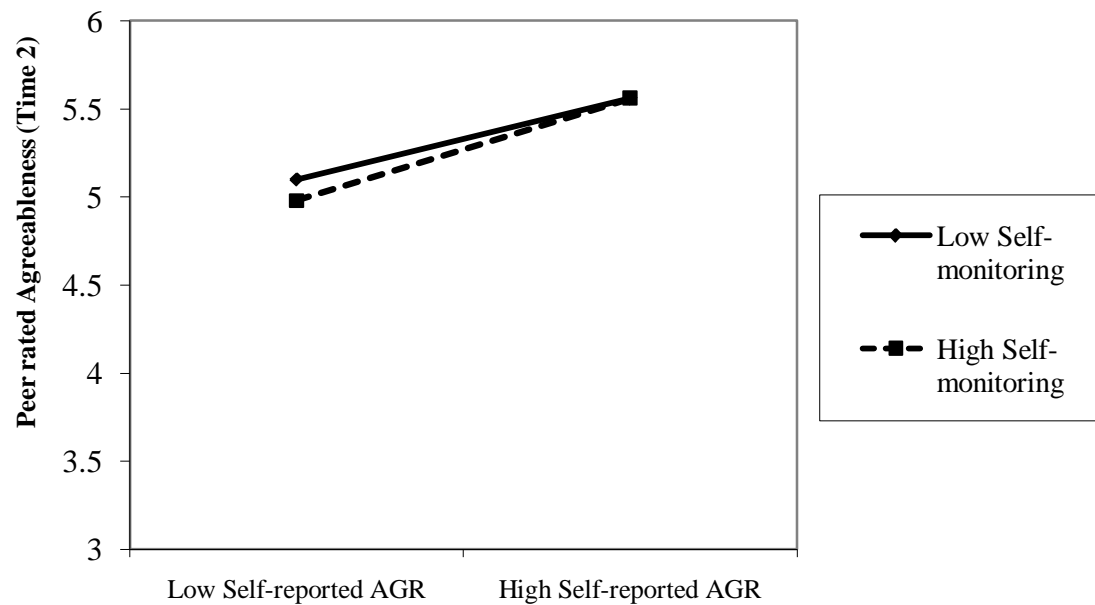


Figure 5. Interaction between GMA and emotional stability at Time 1.

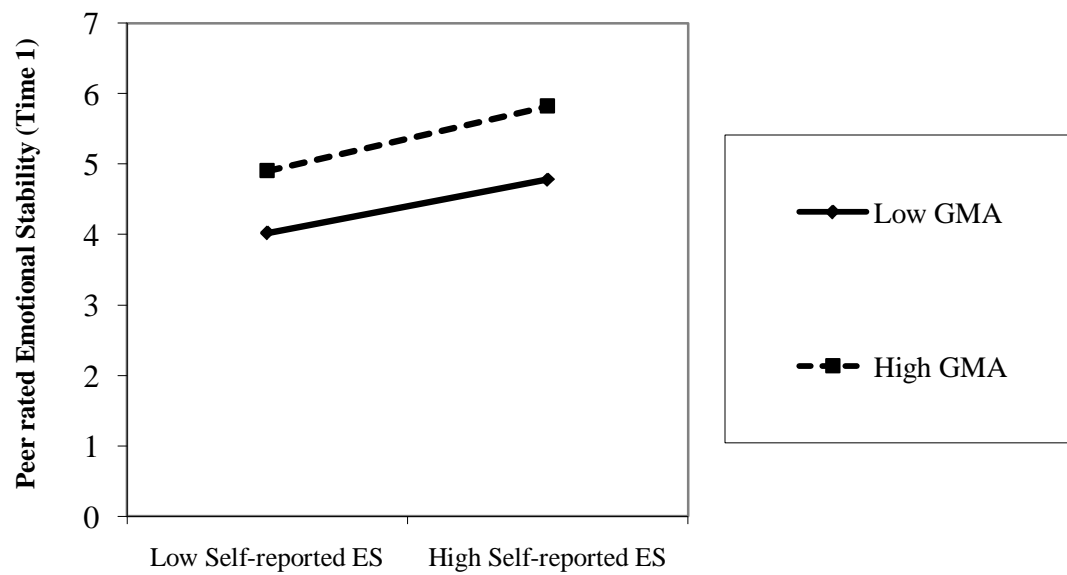


Figure 6. Interaction between self-monitoring and emotional stability at Time 1.

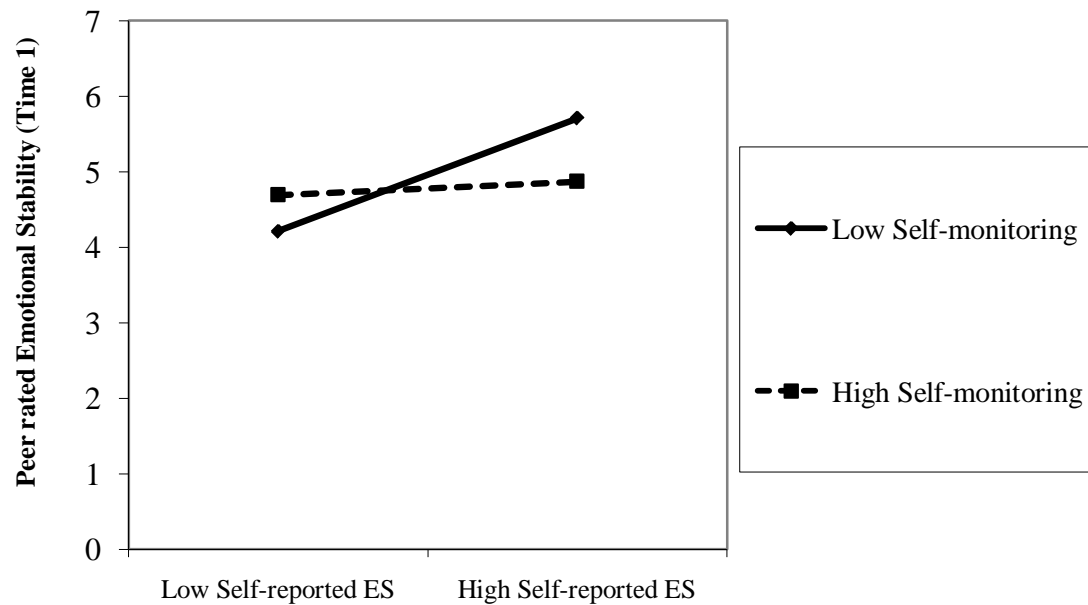


Figure 7. Interaction between GMA and emotional stability at Time 2.

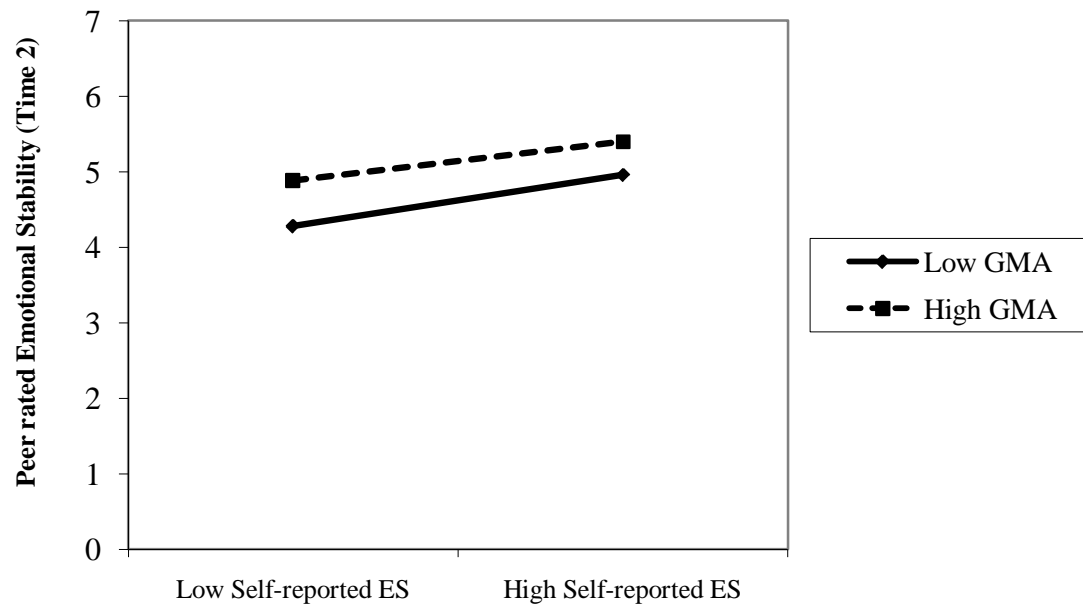


Figure 8. Interaction between self-monitoring and emotional stability at Time 2.

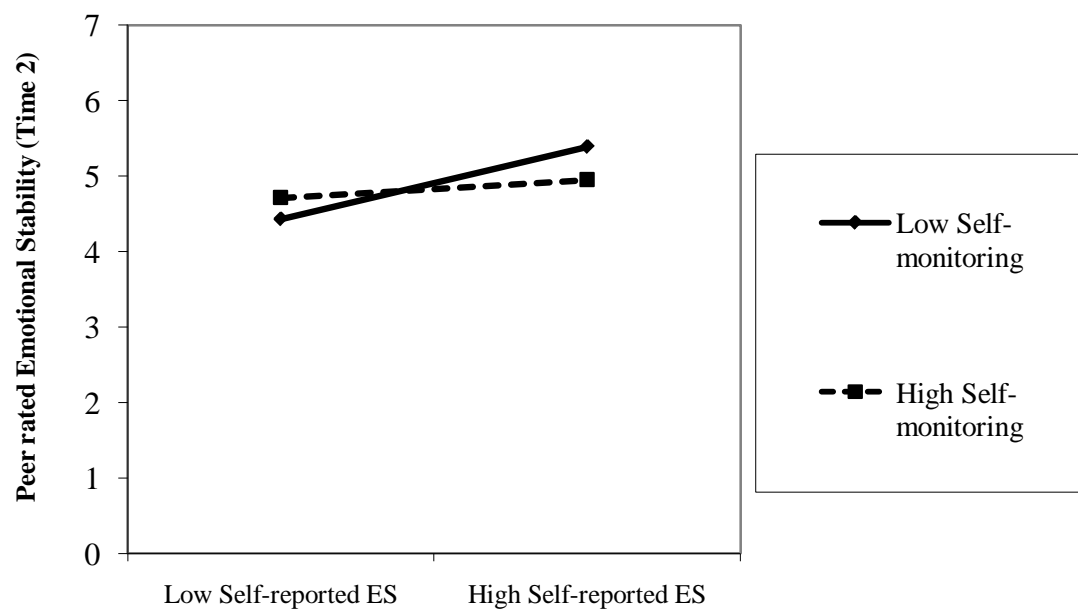


Figure 9. Interaction between GMA and conscientiousness at Time 1.

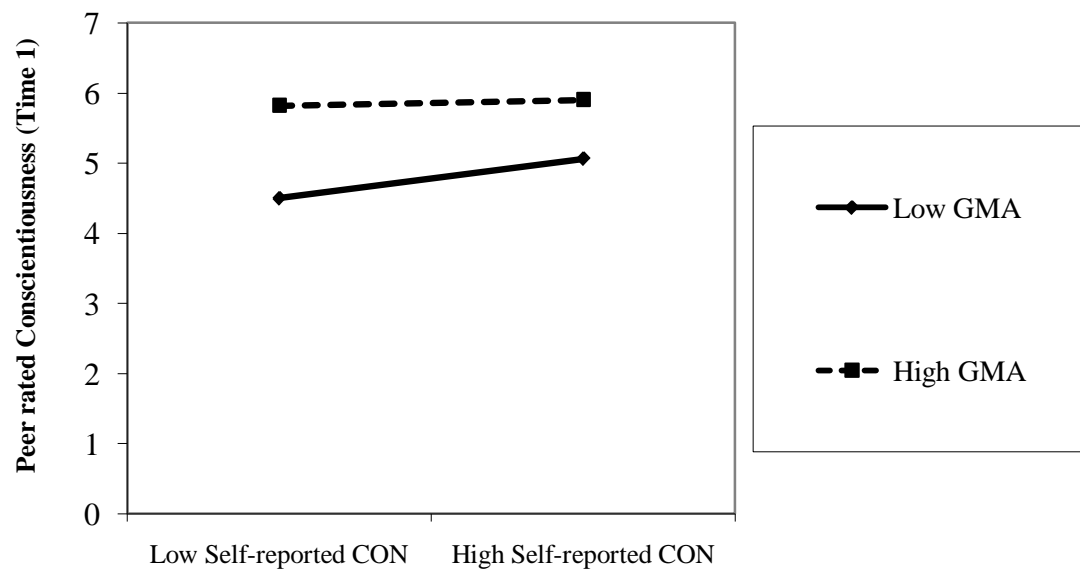


Figure 10. Interaction between GMA and conscientiousness at Time 2.

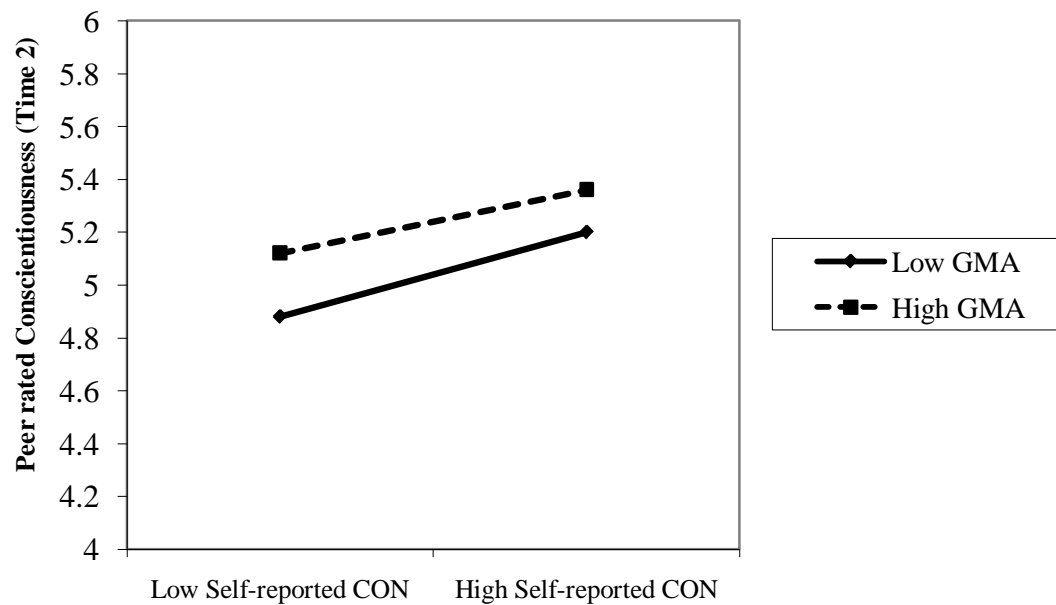


Figure 11. Interaction between self-monitoring and conscientiousness at Time 1.

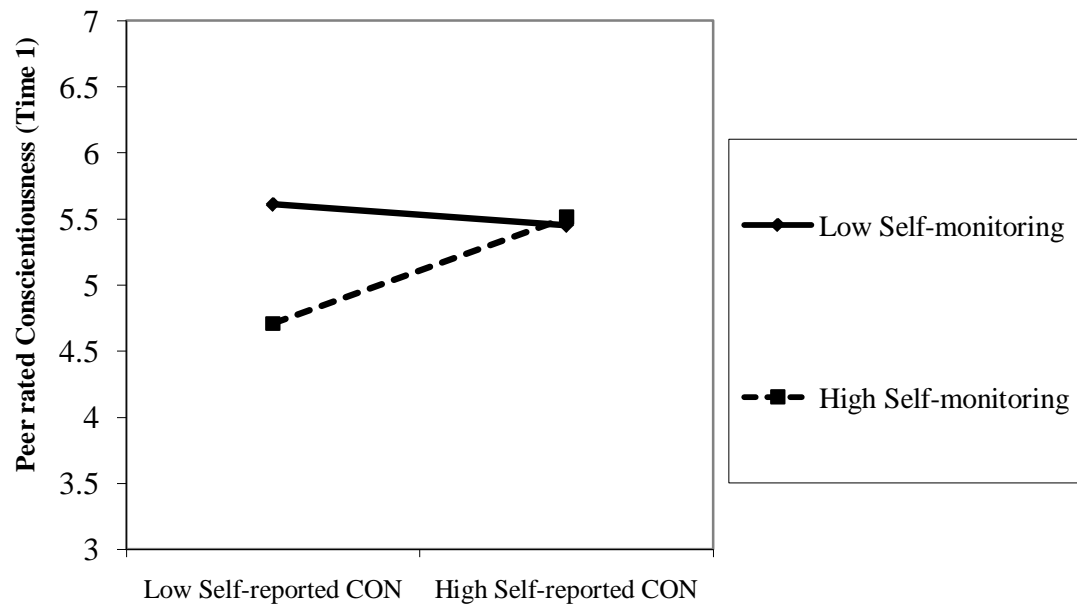
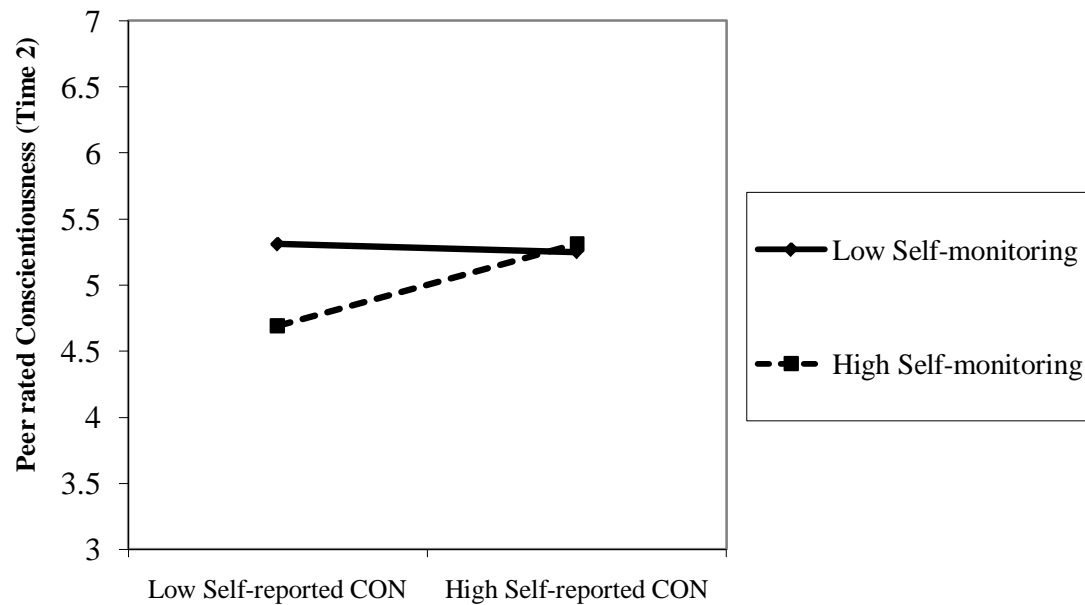


Figure 12. Interaction between self-monitoring and conscientiousness at Time 2.



CHAPTER 5: DISCUSSION

This study examined the possibility that GMA and personality would interact to predict peer ratings of personality such that individuals with higher levels of GMA and low levels of a given desirable personality trait would be rated by their peers as having high levels of that trait. Building on the theory of cognitive buffering, I proposed that this interaction would occur through a process of conscious inhibition of undesirable traits and selection of behaviors that are associated with a given desirable trait. Furthermore, I explored the possibility that this process may occur through the interaction of self-monitoring and personality. In this study I considered these interactions in conjunction with three specific Big Five traits: Agreeableness, emotional stability, and conscientiousness.

The major findings of this study can be summarized in three parts: 1) GMA and personality, 2) self-monitoring and personality, and 3) comparisons between GMA and self-monitoring. First, while the interaction between GMA and agreeableness did predict peer ratings of agreeableness at Time 1 and Time 2, the pattern of the interaction was not expected. Second, the interaction between GMA and emotional stability did not predict peer ratings of emotional stability. Finally, the interaction between GMA and conscientiousness predicted peer ratings of conscientiousness at Time 1, but not at Time 2. In addition, the pattern of the interaction was as I had initially expected. In general, these results seem to support the theory of cognitive buffering.

Second, interactions between self-monitoring and each of the three personality traits examined were detected at Time 1. However the pattern of results for these

interactions were not as expected. In addition, the interaction between self-monitoring and emotional stability seemed to persist over time.

Third, in the two instances in which an interaction between GMA and personality and between self-monitoring and personality was simultaneously detected, the interactions were of similar strength, but were in opposite directions. In three other instances interactions between self-monitoring and personality were detected while no interactions between GMA and personality were detected. Finally, in a single instance the interaction between GMA and personality was detected while no interaction between self-monitoring and personality was detected. These results might suggest that the interaction between self-monitoring and personality may be the better predictor of peer ratings of personality. However, given the inconsistency in these results and the fact that the main effects for GMA were consistently stronger than those of self-monitoring, no such conclusions can be made. Overall, these results may have implications for the self-monitoring construct. Of course, any interpretation of these results requires a more detailed examination of the self-monitoring construct. I address this issue later in the discussion section.

Contextual Issues

Levels of Acquaintanceship

There are several details regarding the context of this study that may help explain why the results did not consistently support the hypotheses. First, the level of communication that occurred between team members may have been different than I had originally expected. Specifically, my hypotheses were based on the notion that project team members would meet together for an amount of time that was adequate to develop a

reasonably accurate picture of the personality of their peers. In order to measure the extent to which participants felt that they knew one another, I administered the Personal Acquaintance Measure. However, without directly measuring the amount of time that team members spent with one another, I do not know how much face-to-face interaction they had. Data from the PAM allows for only rough inferences about the amount of time that team members spent with one another. Based on such data, I speculate that team members spent little time with each other. This is because there was no meaningful change in PAM scores between Time 1 ($M = 2.26$, $SD = .42$) and Time 2 ($M = 2.41$, $SD = .48$). Essentially, there was little, if any, change in the degree to which participants perceived that they had become acquainted with one another during the 12 weeks between the first and second administration of the PAM.

The fact that there was no real change in the PAM scores over the course of this study may help explain the results of this study, which showed minimal resemblance to the hypothesized patterns. Even when interactions were detected, the nature of those interactions seemed to have little in common with the hypotheses put forth in this study, and, in fact, little in common with each other. In order to judge accurately the agreeableness and emotional stability of an individual, it is necessary to spend time with that individual (Norman & Goldberg, 1966, Connolly, Kavanagh, & Viswesvaran, 2007). Based on the mean PAM scores obtained in this study, there is no indication that the participants became more acquainted with each other during the 12 week interval between PAM administrations. If participants did not spend much time with each other, it may have been difficult for them to make meaningful judgments of one another.

The Task-oriented Nature of the Project Teams

Second, the project teams were much more task-oriented than I had expected. It was possible for team members to complete all of their required projects while having only minimal face-to-face interactions with one another. This may have led to a situation in which many group members spent little time with one another, making agreeableness and emotional stability less salient to the performance evaluations that group members completed at the end of each project. Because the groups were so highly focused on task performance, conscientiousness may have been the personality trait that individuals deemed most important to their performance evaluations, making the expression of conscientiousness the process that they were most concerned with managing. The zero-order correlation between peer ratings of conscientiousness and peer ratings of performance indicate that peer ratings of conscientiousness were more strongly related to performance ratings than were any of the other Big Five traits (the corrected correlation was .54). The results suggest that in this particular context, conscientiousness should have been the main trait of interest in any test of the theory of cognitive buffering, and the results for conscientiousness did lend support to the theory of cognitive buffering as it relates to observer ratings of conscientiousness.

Overall, even assuming that social interaction took place between team members, the nature of those interactions is unclear. This does not mean that the study participants did not interact with one another during the course of this study. Interestingly, a prior study conducted in a similar introduction to management course did, in fact, collect extensive data on the nature of the interactions between team members—but did not collect peer-rated personality. I was aware of this before collecting the data for the

current study. However, it was decided that collecting both interaction data and repeated measures of team member personality would be too onerous for the participants.

Therefore, I was unable to collect both sets of measures.

If both sets of measures were available, it would be possible to test the cognitive buffering hypotheses differentially for each group depending on the nature of the interactions that they report. For example, if one set of groups worked mainly through face-to-face meetings in informal settings, it may be more relevant to examine the interaction between GMA and agreeableness for this set of groups. If another set of groups conducted their work mainly through electronic communication and focused on the division of tasks and setting deadlines, the interaction between GMA and conscientiousness may be more relevant. Future research on the theory of cognitive buffering should make greater efforts to theoretically link the study context with the personality traits under examination.

Theoretical Implications

Cognitive Buffering

There are several possible theoretical explanations as to why the hypotheses in this study were not consistently supported. First, and most unlikely, is that the theory of cognitive buffering is an invalid theory. Because the theory of cognitive buffering has been supported in numerous experimental and several field studies, this possibility seems improbable. Second, it may be the case that cognitive buffering does, in fact, affect the influence of personality traits on behavior, but not as strongly as this study predicted. For example, an intelligent introvert may be able to display extroverted behaviors when they are required, but may not be able to act so highly extraverted as to be considered an

extrovert by observers. That is, perhaps the change in behavior that is caused by the cognitive buffering process is not so drastic as to lead observers to completely misread the target of their observations. An intelligent introvert may be able to seem less introverted, but not highly extroverted.

A third possibility, and one that has important implications for further research on cognitive buffering, is that the theory of cognitive buffering by itself does not provide for accurate predictions about the expression of personality traits. Specifically, though the theory of cognitive buffering fails to account for extreme situational constraints on behavior. Personality research has long considered the view that the influence of personality traits on behaviors may be affected by situational demands (Barrick & Mount, 1993; Mischel, 1977). According to research on the person-situation interaction, situational demands can range from strong to weak. A strong situation is one that exerts considerable pressure on individuals to display behaviors that are judged to be congruent with behavioral expectations for that situation (Mischel, 1977). In contrast, weak situations are those in which individuals are not constrained to behave in prescribed ways and are instead free to exhibit whatever behaviors that they see fit.

Strong situations can reduce or even eliminate any relationship between the personality traits and the behaviors of a given individual because “in such situations the person is restricted in the range of behaviors that she or he may be both willing and able to exhibit” (Barrick & Mount, 1993:112). From a purely statistical standpoint two variables cannot be correlated if there is no variance in either one of the variables. Strong situations have the effect of reducing the correlation between personality traits and actual behaviors by decreasing the variance in behaviors that are displayed in that situation.

Thus, only when situation strength is adequately low it is possible for personality traits and behaviors to be correlated.

When both situational strength and cognitive buffering are considered simultaneously, predictions regarding the relationship between personality traits and behaviors should be somewhat different than those made in this study. Specifically, the juxtaposition of the two theories may imply a three-way interaction between personality, situation strength, and cognitive ability such that personality is most highly correlated with behavior only when both situation strength and cognitive ability are adequately low. One limitation of previous research in this area is that personality, situation, and cognitive ability have not been simultaneously examined (e.g., Hollenbeck, Brief, Whitener, & Pauli, 1988; Mount, Barrick, & Strauss, 1999; Sackett, Gruys, & Ellingson, 1998). Some support for a situation-personality interaction has been shown (Barrick & Mount, 1993), but such research has not included a measure of cognitive ability.

Shaffer and Postlethwaite (2010) recently conducted a meta-analysis that may serve as a preliminary test of a three-way interaction between GMA, personality, situation strength. In their paper, the authors hypothesized that personality and job performance would be most highly correlated with job performance in medium complexity jobs. Their argument was that low complexity jobs constrain employee behavior to the extent that personality and performance would not be correlated in such jobs, and that individuals in high complexity jobs possessed the levels of cognitive ability necessary to prevent undesirable personality traits from affecting their jobs performance. The results of the study confirmed their hypothesis, showing that for all but one of the Big Five personality traits, the correlation between personality and job performance was

near zero in low and high complexity jobs and was near .30 for medium complexity jobs. These results are not directly applicable to the current study, because the dependent variables of interest were not the same. However, the meta-analysis does suggest that the theory of cognitive buffering may be only one of multiple processes that affect the relationship between personality and expressed behaviors.

Finally, even if GMA does not interact with personality to predict observer ratings of personality, it is interesting to note that the results of this study showed that the main effects of GMA on observer-ratings of personality were very strong at Time 1 for all three of the personality traits examined (the results were moderately strong at Time 2 only for emotional stability), even though GMA was not correlated with self-rated personality and, in the case of agreeableness, was negatively correlated with self-rated personality. This suggests that while GMA and self-ratings of personality are not highly correlated, individuals that were high in GMA somehow were able to elicit positive evaluations from others. Again, though the hypotheses in this study were not supported, it seems that the GMA of the target of observer-rated personality assessments is an important factor in studies of self-observer agreement that has been overlooked in the past.

Self-monitoring

The results for self-monitoring were unexpected. The results for the interaction between self-monitoring and emotional stability, specifically, may suggest that individuals that are high self-monitors tend to be perceived as neither extremely high nor extremely low in emotional stability regardless of their self-reported levels. One possible explanation might be that it was difficult for observers to perceive accurately the

personality of high self-monitors, resulting in observer ratings that were neither high nor low in a given trait. That is, observers may have defaulted to rating high self-monitors as “average” in a trait that they could not judge with confidence. Though not directly related to personality assessment, a study by Friedman and Miller-Herringer (1991) found that in competitive social contexts, high self-monitors were less likely to display any kind of emotion in competitive situations, leading both trained coders and naïve raters to have difficulty judging their emotions. Snyder and Simpson (1984) reported that high self-monitors reported lower levels of intimacy with romantic partners than did low self-monitors. Thus, it may be the case that self-monitoring has similar effects on observer perceptions of personality. In the current study, perhaps high self-monitors provided few cues from which observers could judge their personality.

The explanation above is predicated on the notion that the personality of high self-monitors is simply difficult to judge—it says nothing about the actual motivation of high self-monitors. Another possible explanation for these results is that high self-monitors were purposefully attempting to present an image that was neither “too hot” nor “too cold”. If this were indeed the case, it would raise an interesting question concerning the behavioral motives of the participants.

As mentioned in the Chapter 1, Snyder (1974) proposed that high self-monitors may base their behaviors on a wide variety of motivations:

“... (a) to communicate accurately one’s true emotional state by means of an intensified expressive presentation; (b) to communicate accurately an arbitrary emotional state which need not be congruent with actual emotional experience; (c) to conceal adaptively an inappropriate emotional state and appear unresponsive and unexpressive; (d) to conceal adaptively an inappropriate emotional state and appear to be experiencing an appropriate one; (e) to appear to be experiencing some emotion when

one experiences nothing and nonresponse is inappropriate” (Synder, 1974; p. 527).

At present, it is generally accepted that the motivation of high self-monitors encompasses a broad desire to achieve social status (Gangestad & Snyder, 2000). In social groups, it has been shown that high self-monitors increase their status by becoming known as being friendly and helpful (Flynn et al., 2006). In the workplace, high self-monitoring may have an effect on ratings of job performance though influencing dimensions of job performance that are related to social interactions. Indeed, Warech and colleagues (1998) showed that in the workplace, self-monitoring was correlated with supervisor ratings of interpersonal effectiveness but was unrelated to ratings of business competence. Barrick, Parks, and Mount (2005) extended this study, showing that self-monitoring interacted with personality traits to predict job performance ratings such that personality was uncorrelated with interpersonal performance for those high in self-monitoring and was correlated with interpersonal performance for those low in self-monitoring. This interaction was not found in relation to task performance. Taken as a whole, these findings tend to support the notion that self-monitoring is generally associated with a motivation to gain social status, at least to a degree.

From the perspective of a researcher, I assumed that team members would be motivated to regulate the expression of the traits most relevant to individual performance and team performance (agreeableness and emotional stability). And, as mentioned previously, the focus of the project teams centered on task performance—as opposed to relationship building—which may have motivated participants to display behaviors that are typically associated with conscientiousness. However, from the perspective of an undergraduate student in an introduction to management course, other considerations—

completely unrelated to the actual projects on which they were working—may have influenced the degree to which participants engaged in controlling the expression of certain personality traits.

Given this study's sample, it is possible to imagine that a male team member, high on self-monitoring, was placed in a group with an attractive female. The male group member, if motivated to impress the female, may have presented himself as neither too low on emotional stability (to avoid being perceived as temperamental or nervous) or too high on emotional stability (to avoid being perceived as too unemotional or too calm). Such efforts might explain the pattern of results seen in this study. While this example might, on one hand, seem trite, it may not be outside the realm of normal behavior for undergraduate students.

The above scenario highlights the need to capture the specific behavioral intentions of high-self monitors in order to make accurate predictions regarding their behavior in social settings. There is little disagreement among researchers that high self-monitors “engage in more behaviors strategic to receiving positive outcomes for a particular social interaction” (Warech, Smither, Reilly, Millsap, & Reilly, 1998: 451); however, the definition of a positive outcome may vary between individuals, and self-monitoring theory is quite vague about this issue. Future research should consider specific motivations or behavioral intentions when building theoretical models of personality and cognitive buffering.

Limitations

This study possesses several key limitations. First and as previously mentioned, I administered no measures that captured the specific behavioral intentions of the

participants. I did measure the performance motivation of the participants for use as a control variable, but this measured only the motivation of the participants as directly related to their broad, overall performance in the class as opposed to their behavioral intentions in their project teams. Failing to capture specific behavioral intentions of the participants may have limited my ability to link the process of cognitive buffering to the most relevant personality traits given the setting of this study.

Second, the personality measure used in this study was chosen in part for its brevity. I chose this measure with the hope that its conciseness would encourage participants to complete all of the peer ratings of personality at each administration, thereby reducing the attrition rate of the study over time. Shorter personality measures are available, but their reliability is quite low. That said, the reliability of the measure used in this study was not excellent. To alleviate this concern as much as possible, I performed all of the regression analyses based on corrected correlations. Although correlations can be corrected for measurement error, measurement unreliability makes the detection of interaction terms very difficult. Future studies of personality and cognitive buffering should consider the use of longer, more reliable measures of personality. Related to this issue, the interrater reliability of peer ratings of personality was quite low, even after applying the Spearman-Brown formula. Again, though corrected correlations were used in all of the regression equations in my analyses, a more reliable measure of personality should be considered for future studies.

Third, participants were extremely lenient when rating the performance of their team members (on a 7-point scale, the Time 1 $M = 6.48$; $SD = .56$; Time 2 $M = 6.32$; $SD = .76$; Time 3 $M = 6.30$; $SD = .81$). Given that performance ratings were not a variable

that was analyzed in this study, this might seem like a tangential issue. However, the hypotheses that I put forth in this study were based on the assumption that students would be motivated to display behaviors related to the personality traits most strongly associated with team performance in order to elicit positive performance ratings from their team members. If team members were exceedingly lenient in their ratings of one another, there may have been little reason for the participants to manage their behaviors in such a way as to elicit positive evaluations from their peers. Thus, the process of cognitive buffering may not have been especially salient in this particular study context.

Related to the leniency in ratings, I found that the vast majority of students that received poor evaluations from their peers—including all of the students who were fired from their groups—did not participate in this study. I noticed this while entering data from the peer evaluations of performance, because entering this data required me to visually inspect all of the performance evaluations for the entire class in which this study was conducted. I only entered the data of those students who consented to participate in the study, but during this process I noticed that low performing students were less likely to have participated in this study. Again, because my hypotheses were based on the idea that students would be motivated to elicit positive evaluations from their peers, the exclusion of students that received poor evaluations from my final dataset could have affected the results of the study. Obviously, the exclusion of poor performers is a form of range restriction which could be corrected if the appropriate data were available. Unfortunately, I cannot estimate the extent to which this type of range restriction affected my results.

Finally, and related to range restriction, I am aware of no method that allows for correcting range restriction in interaction terms. Though I did correct for range restriction in GMA when possible, the data from which the interaction terms were computed were truncated due to range restriction in GMA, which was one of the components of the interaction terms. The correlations that I used in my analyses were derived from these truncated interaction scores, which, in turn, also were subject to a certain amount of range restriction for which I could not correct. If I had been able to fully correct for range restriction in the interaction terms, the corrected correlations between (a) the study variables and (b) the interactions between personality and GMA (columns 16, 18, and 20 of Table 2) would have been slightly larger and the regression weights of the personality by GMA interactions in Tables 3, 4, and 5 also would have been slightly larger.

Future Research

Future research could examine the theory of cognitive buffering in a variety of ways. First, although cognitive ability did not interact with Big Five personality traits as expected, it may be the case that the context of this study was not ideal for testing such an interaction. Future studies should consider testing the theory of cognitive buffering in a more controlled team setting. Experimental studies may be a good starting point for future studies. To date, experimental studies of cognitive buffering have been conducted on individuals in isolated contexts—not team contexts—and have not considered individual differences such as personality. At the same time, it is important to note that future experimental studies should occur in a far more applied context than have the experimental studies reviewed in Chapter 1.

Second, the two field studies that have tested the theory of cognitive buffering as it relates to job performance have reported an interaction between cognitive ability and emotional stability that predicts job performance (Perkins & Corr 2005;2006)

Additionally, meta-analytic results suggest that cognitive ability and personality interact to predict job performance (Shaffer & Postlethwaite, 2010). Thus, field work indicates that the process of cognitive buffering is relevant to workplace setting. However, the results of the current study suggest that the cognitive buffering process may not manifest itself consistently in peer ratings of personality, which leads to the possibility that personality is not the ideal individual difference variable to include in future studies of cognitive buffering. In their taxonomy of motivation sources Mitchell and Daniels (2001), refer to personality as a distal or “cold” motivator that is difficult to alter. The authors consider mood, emotions, and affect to be more proximal or “hot” motivators that are more easily controlled than is personality. Accordingly, it seems reasonable to hypothesize that personality traits may be less influenced by the process of cognitive buffering than are affective states. On a basic level, perhaps cognitive ability interacts with simple emotions in such a way that negative emotions do not affect the job performance of those high in cognitive ability. On a more complex level, cognitive ability may interact with more job relevant affective states. For example, meta-analyses have shown that organizational commitment has a positive correlation with job performance (Mathieu & Zajac, 1990; Riketta, 2005). Perhaps cognitive ability and organizational commitment interact such that for those high in cognitive ability the relationship between organizational commitment and job performance is attenuated. Similar hypotheses might

be made regarding the relationship between job satisfaction, leader-member exchange, or burnout and job performance.

Third, in the current study personality was measured broadly, but it does not have to be. Dudley, Orvis, Lebiecki, and Cortina (2006) expanded on the findings of Barrick and Mount (1991) by performing a meta-analysis of the validity of the facets of conscientiousness. Their analysis revealed that task performance, job dedication, interpersonal facilitation, and counterproductive work behaviors were best predicted by the dependability facet of conscientiousness. The same extension could be used in relation to personality and the theory of cognitive buffering. For example, perhaps individuals high in cognitive ability turn their efforts towards inhibiting specific facets of conscientiousness, such as cautiousness, that have little or no relation to job performance.

Summary

This study drew on the theory of cognitive buffering to examine interactions between general mental ability and Big Five personality traits. Using a large dataset collected from undergraduate students working in teams, regression analysis was used to determine whether GMA and personality interacted to predict peer ratings of personality. In addition, interactions between self-monitoring and personality were also explored.

The results indicated that GMA did not interact with agreeableness or emotional stability in such a way that would be predicted by the theory of cognitive buffering. An interaction between self-monitoring and emotional stability was found, but not in the expected direction. Post hoc analyses showed that GMA interacted with conscientiousness in such a way that would be predicted by the theory of cognitive

buffering. Self-monitoring and conscientiousness also interacted to predict peer ratings of conscientiousness, but not in the expected direction.

Several key contextual factors, including the nature of the project teams studied, the extent of the interactions between team members, and the motivations of the participants, may have affected the results of this study. Future research should examine cognitive buffering in a more controlled setting, or one in which the specific motivations of the subjects is measured. In addition, the theory of cognitive buffering should be examined as it relates to more temporal behavioral influences, such as mood and emotion.

APPENDIX: EMOTIONAL INTELLIGENCE

Another construct that has been proposed as an explanation of individual differences in the ability to control the outward expression of natural behavioral tendencies is emotional intelligence (EI). Salovey and Mayer (1990) first proposed the term EI, which traces its roots to Thorndike's (1920) theory of social intelligence and later to Gardner's (1983) theory of multiple intelligences. According to Mayer and Salovey (1997), EI refers to "the ability to perceive accurately, appraise and express emotion; the ability to access and/or generate feelings when they facilitate thought; the ability to understand emotion and emotional knowledge; and the ability to regulate emotions to promote emotional and intellectual growth" (p. 10). Jordan, Ashkanasy, and Hartel (2002) propose that in a given situation individuals high in EI are able to override the emotions that the situation may trigger and select behaviors that may not reflect their immediate emotional state but instead are appropriate for the situation at hand, while individuals low in EI are not. In the case of a highly emotionally intelligent employee who is experiencing a perception of job insecurity Jordan and colleagues suggest that this employee "may control and redirect their anxiety into productive behaviors that will help to make their job more secure" (p. 368) while a employee low in EI may choose less effective coping behaviors. It follows that individuals high in EI may possess a greater ability to control the expression of behaviors that are associated with the personality traits that have been linked to individual performance in groups than do those low in EI. However, while EI has been the subject of countless publications in the popular press (e.g. Goleman, 1995) it has received far less attention from academic researchers. As explained below, there is much debate surrounding the concept.

First, unresolved issues related to the exact construct validity and measurement of EI continue to plague this line of research. There are those who do not see the usefulness of the EI, holding that definitions of the construct are ambiguous and elusive (e.g., Becker, 2003). At the same time, even proponents of the inclusion of EI in models of organizational behavior cannot seem to agree on how EI should be conceptualized or measured. As a result two main models of EI have developed—the ability based model and the mixed model. The ability model views EI as a specific “type of intelligence.” EI measures based on the ability model ask respondents about their capacity to recognize emotions in themselves and in others or to solve problems that contain some sort of interpersonal or emotional dilemma. The mixed model sees EI as a combination of traits that are dispositional in nature and that have no latent connection to GMA (for a full review of this debate see Mayer, Salovey, and Caruso, 2008). At its core, the mixed model is similar, at least conceptually, to personality. Measures based on the mixed model may in fact be no more than a combination of items that are indistinct from items given in standard personality assessments.

As one might guess, the relationship of EI with GMA and personality is dependent on the type of approach that is used when measuring it. In a meta-analysis, Van Rooy, Viswesvaran, and Pluta (2005) found that the true score correlation between ability-based EI and GMA was .34, while the correlation between ability-based EI and Big Five personality traits ranged from only .06 to .18. In contrast, mixed model EI correlated .13 with GMA while its correlation with Big Five traits ranged from .27 to .40. As it relates to both ability-based models and mixed models, it is also of interest that Schulte, Ree, and Carretta (2004) found that when GMA (as measured by the Wonderlic

Personnel Test) and personality (as measured by the NEO personality inventory) were combined into a single model, the multiple observed correlation with EI as measured by the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) was .58 ($n = 102$). Had the authors corrected for measurement error and analyzed their regression model using true score correlations, the multiple correlation between GMA, personality, and EI would no doubt be even higher. Because EI seems to be confounded with personality and GMA, it has been suggested that social or emotional intelligence is not a separate type of intelligence but is, rather, a set of specific competencies or skills (Hough & Ones, 2001).

Second, although the previous two paragraphs offer some empirically derived information about EI, in general the empirical work on EI that is publicly available in academic journals remains sparse, and it has thus been difficult to reconcile the claims of commercial vendors of EI assessments with the findings of peer reviewed research when attempting to assess the usefulness of the construct of EI for organizational purposes. For example, Hays Associates, which has proprietary rights to the EI measure endorsed by Goleman, claims to have a vast database with a total sample size of over 60,000 that shows EI is far and away the best predictor of job performance (Landy, 2005). Yet, the most comprehensive, peer reviewed meta-analysis to date of EI and job performance includes only 19 studies and a total sample size of 2,652 (Van Rooy & Viswesvaran, 2004). The meta-analysis reported an operational validity for EI of only .24. This estimate approximates the validity of conscientiousness and comes nowhere close to the validity of GMA, which is widely accepted by academics as the single best predictor of job performance. Furthermore, the authors explain, "Although EI did demonstrate incremental validity over GMA, the increase is minimal at .02. However, when switched

around, the incremental validity of GMA over EI is substantial at .31” (p. 85). Thus, it seems that vendor claims as to the usefulness of EI seem to be exaggerated at best.

Third, most reviews of emotional intelligence seem to make a point of drawing attention to Thorndike’s 80 year old reflection on the possibility that there may such a thing as social intelligence. Proponents of emotional intelligence reference Thorndike as a way of showing support for the existence of EI as separate from GMA. While such references seems to lend credibility to the argument that EI may in fact be a separate type of intelligence, too often these same writers exclude references to subsequent factor analyses of various facets of intelligence that provide evidence for a general factor of intelligence (see Jordan et al., 2002; Law, Wong, & Song, 2004; Mayer et al., 2008 for examples). That is, it would seem that proponents of EI take a rather selective approach to reviewing the literature dealing with general intelligence. Thorndike (1936) himself, after hypothesizing that social intelligence existed as different from other types of intelligence, reported in a factor analysis that “abstract” and “social” intelligence both loaded onto a single factor. In this study, Thorndike included tests designed to measure observation of human behavior, recognition of mental state, and judgment in social situations—tests which seem similar in name to the definition of emotional intelligence—and stated that “...the first factor is overwhelmingly the most important. The first factor is weighted positively in every test, and corresponds roughly to what is general to all ten tests” (p. 232). Similar studies of a variety of supposedly unique facets of intelligence met with similar results: all suggested the existence of a single, general factor (e.g. Eysenck, 1939).

NOTES

1. The mixed results that have been reported in this area of research are due, in part, to two factors. First, as explained by Schmidt (1973) and Evans (1991), early work on the interactions between ability and motivation did not take into the effects that between-study differences in the way in which ability and motivation were measured could have on the final results of those studies. That is, though researchers were basically working with the same constructs, their measurement of those constructs varied greatly, which in turn led to large differences in the interaction terms used in their analyses. In terms of the implications of such differences for testing various multiplicative models of motivation, Evans notes: “As a result, a whole body of literature that relied on using such multiplicative composites to test expectancy theory was rendered suspect” (p. 7).

Second, later tests of the multiplicative model of performance did not consistently test the interaction of the same two constructs. Though all such studies included a more or less standardized measure of ability, the moderator variables included varied greatly. Self-esteem, locus of control (Hollenbeck et al., 1988), need for achievement (Wright et al, 1995), global measures of conscientiousness (Mount et al, 1999; O’Reilly & Chatman, 1994), and facet level measures of conscientiousness (Sackett et al., 1998) have all been examined as a moderator of the performance-ability relationship. Thus it is not surprising that results from these studies have been inconsistent.

2. It is important to note that Connonly, Kavanagh, and Viswesvaran (2007) did not explain how they placed studies into the stranger, peer, and close relative categories. In terms of coding studies contained ratings made by strangers, the authors note only that studies were coded as such when they “explicitly noted that the observers were strangers” (p. 113). A detailed examination of these studies shows that there seems to be general agreement on the basic definition of a stranger for the purposes of studies of this sort, but that the definition varies slightly between studies. Examples of the definition of “strangers” have been (a) subjects not previously acquainted with the target of their assessments who rated the target after watching a 60 second video in which the target described themselves (Borkenua & Liebler, 1992) or; (b) watched a 5 minute video of the target conversing with another individual (Funder & Colvin, 1988); (c) subjects who were not previously acquainted and who spend minimal time with each other during which they did not talk to each other and completed assessments of the other’s personality (Beer & Watson, 2008). Connelly & Ones (in press), in the most comprehensive treatment of this literature to date define a stranger as a rater who was “unacquainted with [the] target prior to [the] study beginning” (p. 83).

3. Some may consider concepts such as emotional intelligence (EI) or multiple intelligences to be relevant to this discussion. While such concepts are intuitively appealing and have taken root in popular literature, they have not held up under serious psychometric examination. As such, although I had initially intended to include EI in this study and develop hypotheses dealing with its potential role in the control of the expression of behaviors, after reviewing the EI literature, I found that it is difficult to identify (1) a widely accepted conceptualization of EI; (2) what the construct of emotional intelligence adds to our understanding of individual behavior beyond what we currently know about GMA and personality; (3) any theoretical justification for selecting

an appropriate measure of emotional intelligence for use in this study. Therefore, I have concluded that without any solid theoretical or empirical foundation from which to build hypotheses I cannot, at this time, include EI in this study. A short review of the EI literature is included in the Appendix.

4. Prior to analyzing any data, a number of outliers were deleted from the dataset. These outliers included several sets of personality ratings on which (a) all items were rated as “highly disagree” or (b) all items were rated as “highly agree”. I judged that the participants in these cases had not taken the study seriously and that their data was not usable. In addition, several participants achieved scores on the Wonderlic Personnel Test that were far below what might be considered normal (i.e., the scores potentially indicated the participants were illiterate). After cross-checking these scores against the names of these participants, I determined that most of the low scores were achieved by students whose first language was not English. These data points were also removed prior to data analysis.

REFERENCES

- Aiken, L.S., & West, S.G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Allik, J., & Realo, A. (1997). Intelligence, academic abilities, and personality. *Personality and Individual Differences*, 23, 809-814.
- Ashton, M.C., & Lee, K. (2005). Honesty-humility, the Big Five, and the Five-Factor Model. *Journal of Personality*, 73, 1321-1353.
- Barrick, M.R., & Mount, M.K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.
- Barrick, M.R., & Mount, M.K. (1993). Autonomy as a moderator of the relationship between the Big Five personality dimensions and job performance. *Journal of Applied Psychology*, 78, 111-118.
- Barrick, M.R., Mount, M.K., & Judge, T.A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *Personality and Performance*, 9, 9-29.
- Barrick, M.R., Parks, L., & Mount, M.K. (2005). Self-monitoring as a moderator of the relationship between personality traits and performance. *Personnel Psychology*, 58, 745-767.
- Becker, T. (2003). Is emotional intelligence a viable concept? *Academy of Management Review*, 28, 192-195.
- Beehr, A., & Watson, D. (2008). Personality judgment at zero acquaintance: Agreement, assumed similarity, and implicit similarity. *Journal of Personality Assessment*, 90, 250-260.
- Biesanz, J.C., West, S.G., & Millevoi, A. (2007). What do you learn about someone over time? The relationship between length of acquaintanceship and consensus and self-other agreement in judgments of personality. *Journal of Personality and Social Psychology*, 92, 119-135.
- Bjorklund, D.F., & Harnishfeger, K.K. (1995). The evolution of inhibition mechanisms and their role in human cognition and behavior. In F.N. Dempster & C.J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 141-173). San Diego: Academic Press.
- Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology*, 62, 645-657.
- Bradley, B. H., Postlethwaite, B. E., Burns, M. J., & Brown, K. G. (2008). Team processes and outcomes: A longitudinal examination. Unpublished manuscript.

- Breslau, N., Lucia, V.C., & Alvarado, G.F. (2006). Intelligence and other predisposing factors in exposure to trauma and posttraumatic stress disorder. *Archives of General Psychology*, 63, 1238-1245.
- Brewin, C.R., & Beaton, A. (2002). Thought suppression, intelligence, and working memory capacity. *Behavior Research and Therapy*, 40, 923-930.
- Brewin, C.R., & Smart, L. (2005). Working memory capacity and suppression of intrusive thoughts. *Journal of Behavior Therapy and Experimental Psychiatry*, 36, 61-68.
- Connelly, B.S., & Ones, D.S. (in press). An other perspective on personality: Meta-analytic integration of observer's accuracy and predictive validity. *Psychological Bulletin*.
- Connolly, J.J., Kavanagh, E.J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment*, 15, 110-117.
- Conway, A.R.A., Kane, M.J., & Engle, R.W. (2003). Working memory capacity and its relation to general intelligence. *TRENDS in Cognitive Sciences*, 7, 547-552.
- Costa, P.T., & McCrae, R.R. (1992). NEO-PI-R professional manual. Odessa, FL: Psychological Assessment Resources.
- Davies, M., Stankov, L., & Roberts, R.D. (1998). Emotional intelligence: In search of an elusive construct. *Journal of Personality and Social Psychology*, 75, 989-1015.
- Day, D.V., & Scheicher, D.J. (2006). Self-monitoring at work: A motive-based perspective. *Journal of Personality*, 74, 685-713.
- Day, D.V., Schleicher, D.J., Unckless, A.L., & Hiller, N.J. (2002). Self-monitoring at work: A meta-analytic investigation of construct validity. *Journal of Applied Psychology*, 87, 390-401.
- Diamond, S., & Balvin, R.S., & Diamond, F.R. (1963). *Inhibition and choice: A neurobehavioral approach to problems of plasticity in behavior*. New York: Harper & Roe.
- Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417-440.
- Dudley, N.M., Orvis, K.A., Lebiecki, J.E., & Cortina, J.M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits.
- Eysenck, H.J. (1939). Primary mental abilities. *British Journal of Educational Psychology*, 9, 270-275.

- Elliot, A.J., & Church, M.A. (1997). A hierarchical model of approach and avoidance achievement orientation. *Journal of Personality and Social Psychology*, 72, 218-232.
- Fleishman, E.A. (1958). A relationship between incentive motivation and ability level in psychomotor performance. *Journal of Experimental Psychology*, 56, 78-81.
- Flynn, F.J., Reagans, R.E., Amanatullah, E.T., & Ames, D.R. (2006). Helping one's way to the top: Self-monitors achieve status by helping others and knowing who helps whom. *Journal of Personality and Social Psychology*, 91, 1123-1137.
- French, E.G. (1958). The interaction of achievement motivation and ability in problem-solving success. *Journal of Abnormal of Social Psychology*, 57, 306-309.
- Funder, D.C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, 102, 652-670.
- Funder, D.C. (1999). *Personality judgment: A realistic approach to person perception*. Academic Press: San Diego, CA.
- Funder, D.C., & Colvin, C.R. (1998). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology*, 55, 149-158.
- Funder, D.C., & Dobroth, K.M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology*, 52, 409-418.
- Funder, D.C., Kolar, D.C., & Blackman, M.C. (1995). Agreement among judges of personality: Interpersonal relations, similarity, and acquaintanceship. *Journal of Personality and Social Psychology*, 69, 656-672.
- Gangestad, S.W., & Snyder, M. (2000). Self-monitoring: Appraisal and reappraisal. *Psychological Bulletin*, 126, 530-555.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. Basic Books, Inc: New York, NY.
- Gernsberger, M.A., & Faust, M.E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology*, 17, 245-262.
- Gill, C.M., & Hodgkinson, G.P. (2007). Development and validation of the Five-Factor model questionnaire (FFMQ): An adjectival-based personality inventory for use in occupational settings. *Personnel Psychology*, 60, 731-766.
- Goleman, D. (1995). *Emotional intelligence*. Bantam Books: New York, NY.

- Gottfredson, L.S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24, 79-132.
- Green, B.L., Grace, M.C., Lindy, J.D., Gleser, G.C., & Leonard, B.A. (1990). Risk factors for PTSD and other diagnoses in a general sample of combat veterans. *American Journal of Psychiatry*, 147, 729-733.
- Harnishfeger, K.K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F.N. Dempster & C.J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 141-173). San Diego: Academic Press.
- Hirschfeld, R.R., Lawson, L., & Mossholder, K.W. (2004). Moderators of the relationship between cognitive ability and performance: General versus context-specific achievement motivation. *Journal of Applied Social Psychology*, 34, 2389-2409.
- Herrnstein, R.J. (1971). I.Q. *The Atlantic Monthly*, pp. 43-64.
- Hobart, R. & Dunnette, M.D. (1967). Development of moderator variables to enhance the prediction of managerial effectiveness. *Journal of Applied Psychology*, 51, 50-64.
- Hogan, R. (2007). *Personality and the fate of organizations*. Lawrence Erlbaum Associates: New Jersey.
- Hough, L.M., & Ones, D.S. (2001). The structure, measurement, validity, and use of personality variables in industrial, work, and organizational psychology. In N. Anderson, D.S. Ones, H. Sinangil, and C. Viswesvaran (Eds.) *Handbook of Industrial, Work, and Organizational Psychology*, vol. 1. London: Sage.
- Hollenbeck, J.R., Brief, A.P., Whitener, E.M., & Pauli, K.E. (1988). An empirical note on the interaction of personality and aptitude in personnel selection. *Journal of Management*, 14, 441-451.
- Hunter, J.E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, 29, 340-362.
- Hunter, J.E. (1996). A multiple regression program in basic.
- Hunter, J.E., & Hunter, R.F. (1984). Validity and utility of alternate predictors of performance. *Psychological Bulletin*, 96 72-98.
- Hunter, J.E., & Schmidt, F.L., & Le, H. (2006). Implications of direct and indirect range restriction for meta-analysis methods and findings. *Journal of Applied Psychology*, 91, 594-612.

- James, L.R., Demaree, R.G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, 69, 85-98.
- Jensen, A.R. (1998). *The g factor: The science of mental ability*. London: Praeger.
- John, O.P., & Robins, R.W. (1992). Determinants of interjudge agreement on personality traits: The big five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, 61, 521-551.
- Jordan, P.J., Ashkanasy, N.M., & Hartel, C.E. (2002). Emotional intelligence as a moderator of emotional and behavioral reactions to job insecurity. *Academy of Management Review*, 27, 361-372.
- Judge, T.A., Ilies, R., Dimotakis, N. (2010). Are health and happiness the product of wisdom? The relationship of general mental ability to educational and occupational attainment, health, and well-being. *Journal of Applied Psychology*, 95, 454-468.
- Judge, T.A., Klinger, R.L., & Simon, L.S. (2010). Time is on my side: Time, general mental ability, human capital, and extrinsic career success. *Journal of Applied Psychology*, 95, 92-107.
- Judge, T.A., Colbert, A.E., & Ilies, R. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology*, 89, 542-552.
- Judge, T.A., Jackson, C.L., Shaw, J.C., Scott, B.A., & Rich, B.L. (2007). Self-efficacy and work-related performance: The integral role of individual differences. *Journal of Applied Psychology*, 92, 107-127.
- Keller, T. (1999). Images of the familiar: Individual differences and implicit leadership theories. *Leadership Quarterly*, 10, 589-607.
- Kipnis D. (1962). A noncognitive correlate of performance among lower aptitude men. *Journal of Applied Psychology*, 46, 76-80.
- Landy, F.J. (2005). Some historical and scientific issues related to research on emotional intelligence. *Journal of Organizational Behavior*, 26, 411-424.
- Law, K.S., Wong, C.S., & Song, L.J. (2004). The construct and criterion validity of emotional intelligence and its potential utility for management studies. *Journal of Applied Psychology*, 89, 483-496.
- Lawler, E.E. (1966). Ability as a moderator of the relationship between job attitudes and job performance. *Personnel Psychology*, 19, 153-164.

- Le, H., Schmidt, F.L., Putka, D.J. (2009). The multifaceted nature of measurement artifacts and its implications for estimating construct-level relationships. *Organizational Research Methods*, 12, 165-200.
- Locke, E.A. (1965). Interaction of ability and motivation in performance. *Perceptual and Motor Skills*, 21, 719-725.
- Luu, P., Tucker, D.M., & Derryberry, D. (1998). Anxiety and the motivational basis of working memory. *Cognitive Therapy and Research*, 22, 577-594.
- Maier, N.R.F. (1955). *Psychology in industry*. Boston: Houghton-Mifflin.
- Mabe III, P.A., & West, S.G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, 67, 280-296.
- MacLeod, C.M. (2007). The concept of inhibition in cognition. In D.S. Gorfein & C.M. MacLeod (Eds.), *Inhibition in cognition* (pp. 3-23). Washington DC, American Psychological Association.
- Macklin, M.L., Metzger, L.J., Litz, B.T, McNally, R.J., Lasko, N.B., Orr, S.P., & Pitman, R.K. (1998). Lower precombat intelligence is a risk factor for posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology*, 66, 323-326.
- Mathieu, J.E., & Zajac, D.M. (1990). A review and meta-analysis of the antecedents, correlates, and consequences of organizational commitment. *Psychological Bulletin*, 108, 171-194.
- Mayer, J.D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.) *Emotional development and emotional intelligence: Implications for educators*. Basic Books: New York, NY.
- Mayer, J.D., Salovey, P., & Caruso, D.R. (2008). Emotional intelligence—New ability or eclectic traits? *American Psychologist*, 63, 503-517.
- McCrae, R.R. (1993). Moderated analyses of longitudinal personality stability. *Journal of Personality and Social Psychology*, 65, 577-585.
- McCrae, R.R., & Costa, P.T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90.
- McNally, R.J., & Shin, L.M. (1995). Association of intelligence with severity of posttraumatic stress disorder symptoms in Vietnam combat veterans. *American Journal of Psychiatry*, 152, 936-938.
- Mitchell, T.R., & Daniels, D. (2001). Motivation. In W.C. Borman, D.R. Ilgen, & R.J. Klimoski (Eds.), *Comprehensive Handbook of Psychology, Vol. 12: Industrial and Organizational Psychology*. New York: John Wiley & Sons.

- Miller, J.S., & Cardy, R.L. (2000). Self-monitoring and performance appraisal: Rating outcomes in project teams. *Journal of Organizational Behavior*, 21, 609-626.
- Morgeson, F.P., Campion, M.A., Dipboye, R.L., Hollenbeck, J.R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, 60, 683-729.
- Morrison, K.A. (1997). Personality correlates of the Five-Factor model for a sample of business owners/managers: Associations with scores on self-monitoring, type A behavior, locus of control, and subjective well-being. *Psychological Reports*, 80, 255-272.
- Mount, M.K., Barrick, M.R., & Stewart, G.L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance*, 11, 145-165.
- Mount, M.K., Barrick, M.R., & Strauss, J.P. (1994). Validity of observer ratings of the Big Five personality factors. *Journal of Applied Psychology*, 79, 272-280.
- Mount, M.K., Barrick, M.R., & Strauss, J.P. (1999). The joint relationship of conscientiousness and ability with performance: Test of the interaction hypothesis. *Journal of Management*, 25, 707-721.
- Neuman, G.A., & Wright, J. (1999). Team effectiveness: Beyond skills and cognitive ability. *Journal of Applied Psychology*, 84, 376-389.
- Norman, W.T., & Goldberg, L.R. (1966). Raters, ratees, and randomness in personality structure. *Journal of Personality and Social Psychology*, 4, 681-691.
- Nunnally, J.C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Oh, I., Wang, G., & Mount, M.K. (2010). Validity of observer ratings of personality: A meta-analysis. Unpublished manuscript.
- Ones, D.S., & Viswesvaran, C. (1998). Faking on personality and integrity assessment for personnel selection. *Human Performance*, 11, 245-269.
- Ones, D. S., & Viswesvaran, C. (2003). Job-specific applicant pools and national norms for personality scales: Implications for range-restriction corrections in validation research. *Journal of Applied Psychology*, 88, 570-577.
- O'Reilly, C.A. & Chatman, J.A. (1994). Working smarter and harder: A longitudinal study of managerial success. *Administrative Science Quarterly*, 39, 603-627.
- Perkins, A.M., & Corr, P.J. (2005). Can worriers be winners? The association between worrying and job performance. *Personality and Individual Differences*, 38, 25-31.

- Perkins, A.M., & Corr, P.J. (2006). Cognitive ability as a buffer to neuroticism: Churchill's secret weapon? *Personality and Individual Differences*, 40, 39-51.
- Ployhart, R.E., Schneider, B., & Schmitt, N. (2006). *Staffing Organizations*. Lawrence Erlbaum Associates: New Jersey.
- Postlethwaite, B.E. (2009) Self-estimated and psychometric intelligence: A meta-analysis. Paper presented at the Academy of Management 2009 annual meeting, Chicago, IL, August 7-11.
- Ready, R.E., Clark, L.A., Watson, D., & Westerhouse, K. (2000). Self- and peer-reported personality: Agreement, trait ratability, and the "self-based heuristic". *Journal of Research in Personality*, 34, 208-224.
- Riketta, M. (2005). Organizational identification: A meta-analysis. *Journal of Vocational Behavior*, 66, 358-384.
- Rosen, V.M., & Engle, R.W. (1998). Working memory capacity and suppression. *Journal of Memory and Language*, 39, 418-436.
- Rotundo, M., & Sackett, P.R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy-capturing approach. *Journal of Applied Psychology*, 87, 66-80.
- Sackett, P.R., Gruys, M.L., & Ellingson, J.E. (1998). Ability-personality interactions when predicting job performance. *Journal of Applied Psychology*, 83, 545-556.
- Salovey, P., & Mayer, J.D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality*, 9, 185-211.
- Saucier, G. (1994). Mini-markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment*, 63, 506-516.
- Schmidt, F.L., Le. H., & Ilies, R. (2003). Beyond alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual-differences constructs. *Psychological Methods*, 8, 206-224.
- Schmidt, F.L., Shaffer, J.A., & Oh, I. (2008). Increased accuracy for range restriction corrections: Implications for the role of personality and general mental ability in job and training performance. *Personnel Psychology*, 61, 827-868.
- Schulte, M.J., Ree, M.J., & Carretta, T.R (2004). Emotional intelligence: not much more than g and personality. *Personality and Individual Differences*, 37, 1059-1068.
- Schutte, N.S., Malouff, J.M., Hall, L.E., Haggerty, D.J., Cooper, J.T., Golden, C.J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences*, 25, 167-177.

- Schmidt, F.L.(1973). Implications of a measurement problem for expectancy theory research. *Organizational Behavior and Human Performance*, 10, 243-251.
- Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, 30, 526-537.
- Shaffer, J.A., & Postlethwaite, B.E. (2010). Reevaluating the validity of personality: A meta-analysis of the moderating effects of job complexity (Unpublished manuscript).
- Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology*, 15, 201-292.
- Sternberg, R.J., & Wagner, R.K. (1986). *Practical Intelligence*. Cambridge University Press: New York, NY.
- Stewart, G.L. (2003). Toward and understanding of the multilevel role of personality in teams. In M.R. Barrick and A.M. Ryan (Eds.) *Personality and Work*. Wiley.
- Tett, R.P., Fox, K.E., & Wang, A. (2005). Development and validation of a self-report measure of emotional intelligence as a multidimensional trait domain. *Personality and Social Psychology Bulletin*, 31, 1-30.
- Thorndike, R.L. (1936). Factor analysis of social and abstract intelligence. *The Journal of Educational Psychology*
- Thurstone, L.L. (1931). Multiple factor analysis. *Psychological Review*, 38, 406-427.
- Van Rooy, D.L., & Viswesvaran, C. (2004). Emotional intelligence: A meta-analytic investigation of predictive validity and nomological net. *Journal of Vocational Behavior*, 65, 71-95.
- Van Rooy, D.L., Viswesvaran, C., & Pluta, P. (2005). An evaluation of construct validity: What is this thing called emotional intelligence? *Human Performance*, 18, 445-462.
- Vroom, V.H. (1960). *Some personality determinants of the effects of participation*. Englewood Cliffs, NJ: Prentice Hall.
- Wall Street Journal (Dec 13, 1994) (Eastern edition). New York, N.Y.: pg. A18.
- Warech, M.A., Smither, J.W., Reilly, R.R., Millsap, R.E., & Reilly, S.P. (1998). Self-monitoring and 360-degree ratings. *Leadership Quarterly*, 9, 449-473.
- Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology*, 57, 120-128.

- Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology*, 78, 546-558.
- Evans, M.G. (1991). The problem if analyzing multiplicative composites: Interactions revisited. *American Psychologist*, 46, 6-15.
- Wright, P.M., Kacmar, K.M., McMahan, G.C., & Deleeuw, K. (1995). $P = f(M \times A)$: Cognitive ability as a moderator of the relationship between personality and job performance. *Journal of Management*, 21, 1129-1139.