The Interaction of Neuroticism and Gender and Its Impact on Self-Efficacy and Performance

Neal Schmitt

Michigan State University,

Online Publication Date: 01 January 2008
Self-efficacy has been linked to performance in many areas of human endeavor. Examinations of personality correlates of self-efficacy suggest that emotional stability or neuroticism is consistently related to self-efficacy. Other findings suggest a gender difference in emotional stability. In this article, it is proposed that emotional stability and gender interact to affect self-efficacy and that efficacy, in turn, affects performance. The hypothesized interaction and mediation are confirmed using longitudinal data collected from college students. The model and results also provide evidence of one mechanism whereby personality affects performance even in the absence of direct personality–performance relationships.

Self-efficacy is a personal judgment of “how well one can execute courses of action required to deal with prospective situations” (Bandura, 1982, p. 122). According to Bandura, self-efficacy determines whether an individual’s coping behavior will be initiated, how much task-related effort will be expended, and how long the effort will be expended. The importance of these expectations is well documented in a wide variety of educational (e.g., Gore, 2006) and work contexts (e.g., Gist, Schwoerer, & Rosen, 1989). Stajkovic and Luthans (1998) identified 114 studies documenting the relationship between self-efficacy and performance in a wide variety of domains. They reported a meta-analytic correlation between self-efficacy and work-related performance of .38.

Moreover, self-efficacy is seen as more than merely an ability estimate. Bandura and Locke (2003) argued that it represents a motivation component as well or a willingness to expend effort consistent with one’s ability. Bandura (1997) also argued that traits, in particular personality measures, are not likely to account...
for much of the variance in human behavior across different tasks and circumstances. He cited successful manipulations of self-efficacy in a variety of experimental situations (Bandura & Locke, 2003). However, the importance of individual differences in perceived (as opposed to manipulated) self-efficacy suggests that personal characteristics of a more stable nature do influence self-efficacy.

Evidence regarding personality correlates of self-efficacy is provided in a meta-analysis performed by Judge and Ilies (2002). They reported that the average observed correlation between self-efficacy and neuroticism across three studies was –.29. The corrected correlation was –.35. Extraversion was similarly related to self-efficacy ($\rho = -0.33$) whereas conscientiousness and openness were less strongly related ($\rho = 0.22$ and 0.20, respectively) to self-efficacy. Judge and Ilies also reported that the correlation between neuroticism and self-efficacy was higher in academic (as opposed to work) settings, where most of the research was conducted and in concurrent studies, which was the research design most often employed to evaluate the neuroticism/self-efficacy relationship.

The objective of the research reported in this article is to explore further the nature of the relationship between self-efficacy and neuroticism and performance. A model in which the interaction of neuroticism and gender affects self-efficacy, which in turn impacts performance in an academic setting, is tested. In the remainder of this article, we use emotional stability to refer to neuroticism so that all hypothesized and empirical relationships are positive rather than negative.

There are empirical and logical arguments that support the model depicted in Figure 1. This model and the data reported here help provide an understanding of one of the reasons why personality is related to performance (Barrick & Mount, 1991). Barrick, Mitchell, and Stewart (2003) and Mount, Barrick, and Ryan (2003) have argued for the examination of similar models as a means of determining the mechanisms underlying personality–performance relationships. Barrick, Mount, and Strauss (1993) provided an empirical example of such a mediating model in which goal setting mediated the relationship between conscientiousness and sales performance.

There are several reasons to focus on gender when analyzing the relationship between emotionality stability and self-efficacy and their combined impact on performance. Gender differences in emotional stability have been reported on several

![FIGURE 1](image)

**FIGURE 1** Hypothesized mediation model relating gender, emotional stability, the interaction of gender and emotional stability, self-efficacy, and GPA.
measures of emotional stability. Hough (1998) reported a difference of .21 standard deviation units across various measures of the construct she labeled adjustment. In related research, Furnham and Thomas (2004) found that emotional stability significantly predicted self-estimates of intelligence, with more stable or less neurotic people providing higher estimates. Comijs, Deeg, Dik, Twisk, and Jonker (2002) reported that, in a large sample of older adults, emotional stability was negatively associated with memory problems even though these adults gave no objective evidence of cognitive decline. In a follow-up to their earlier research, Furnham and Buchanan (2005) showed that emotional stability was related to gender such that women reported being less emotionally stable than men \( (d = .26) \) and that women also reported lower levels of intelligence. This gender difference along with the finding that less emotionally stable persons provided lower estimates of intelligence suggests that these two variables in combination will disadvantage less emotionally stable women. Whether the combination is additive or interactive is an important question. It is also the case that the self-estimates of intelligence collected by Furnham and colleagues are quite general compared to the more domain-specific measures that Bandura (1991) and others have found related to performance in various domains, so these studies perhaps constitute only indirect evidence of a linkage between emotional stability and self-efficacy.

Stereotype threat research also provides evidence that women regard themselves as less able than men in areas related to mathematical/numerical skills (Spencer, Steele, & Quinn, 1999; Steele & Davies, 2003). This sense of low self-efficacy along with a feeling that mathematical skill is important to their self-identity is thought to comprise a self-handicap of sorts that diminishes women’s performance on at least measures of mathematical skill. Stereotype threat is thought to produce anxiety about measures critical to one’s self-identity that results in decrements of performance. This anxiety-provoking mechanism is central to stereotype threat as evidenced by the fact that it is sometimes used as a manipulation check in stereotype threat studies (e.g., Nguyen, O’Neal, & Ryan, 2003). It certainly seems plausible that less emotionally stable women would be more adversely affected by stereotype threat or related test anxiety than would men who are usually more emotionally stable. In the study presented here, the domain of interest is academic performance, which should be important to the self-identity of both male and female students. If the relatively lower emotional stability of female students makes them more susceptible to the anxiety produced by stereotype threat or test taking, it should have a greater impact on their self-efficacy than it would for male students. This expected interaction is depicted in Figure 1. It should be pointed out, though, that this study did not involve a manipulation or measure of stereotype threat, and one would not have been possible in the context of the college student performance that is the focus in this article.

A model similar to that depicted in Figure 1, but involving a different personality characteristic, was proposed by Zhao, Seibert, and Hills (2005). They proposed
that gender along with risk propensity, experience, and perceptions of learning acquired in an MBA course would lead to a sense of entrepreneurial self-efficacy and that self-efficacy would lead to intentions to engage in entrepreneurial activity. The relationship between gender and intent to engage in entrepreneurial activity was not mediated by self-efficacy but was directly related to entrepreneurial intentions. However, we posit that gender and personality (in our case, emotional stability) will interact in their impact on self-efficacy. That interaction was neither hypothesized nor tested in the Zhao et al. research.

The fact that women appear to have a lower sense of self-efficacy and report being less emotionally stable and perhaps more susceptible to the anxiety produced by stereotype threat and taking tests suggests an interaction between gender and emotional stability in its impact on self-efficacy; specifically, we propose the following hypothesis:

**H1:** Gender and neuroticism will interact in such a manner that the relationship between emotional stability and self-efficacy will be significantly greater for women than the same relationship for men.

The nature of this hypothesized interaction is depicted in Figure 2.

Based on the large body of literature on efficacy-performance relationships, we also predict that self-efficacy will be related to performance, that is, academic performance in this instance. Generalized self-esteem has long been seen as an important correlate of a wide variety of behavior across many work and nonwork domains (e.g., Brockner, 1983; Brockner & Elkind, 1985; Rosenberg, 1965), but

---

**FIGURE 2** Hypothesized interaction effect of gender and emotional stability on self-efficacy.

Downloaded By: [Michigan State University] At: 21:45 4 November 2008
more domain specific measures of self-efficacy have proven superior in most empirical research in academic and work contexts (Multon, Brown, & Lent, 1991; Stajkovic & Luthans, 1998) and are preferred by Bandura (1997; Bandura & Locke, 2003). Accordingly, our second hypothesis poses a relationship between academic self-efficacy and academic performance:

H2: Academic self-efficacy will be related positively to academic performance.

Our model posits that the relationship between the interaction of emotional stability and gender and academic performance will be mediated by self-efficacy. This is based on previous research (Furnham & Buchanan, 2005; Judge & Ilies, 2002) that documents that both gender and emotional stability correlate with academic self-efficacy. It is reasonable to posit that academic self-efficacy will be related to academic performance and that self-efficacy will mediate the relationship between gender, emotional stability, and their interaction and academic performance. To evaluate this notion, academic performance was regressed self-efficacy on Step 1 and gender, emotional stability, and their interaction on Step 2. The third hypothesis is as follows:

H3: In this regression analysis, self-efficacy will be significantly related to academic performance, but the variables entered at Step 2 will not contribute significantly to the prediction of academic performance.

METHOD

Sample

Participants were 891 incoming freshmen at 10 U.S. colleges and universities. These institutions were quite different and included large state schools, highly selective institutions, and historically Black colleges. These students were recruited to participate in a related data collection effort at the beginning of their college experience and provided their demographic information and permission to be recontacted at later dates. They also granted permission for us to obtain their grade point average (GPA) data from officials at their university. All students were paid $40 to complete the initial measures in 2 hr. The total initial sample size across all 10 schools was 2,771. The average age of our participants was just older than 18; over 97% of our sample was either 18 or 19 years old. Sixty-four percent of the sample was female, 96% were U.S. citizens, and 94% indicated that English was their native language. Ethnically, this sample was 55% White, 25% African American (because of intentional oversampling of this subgroup), 6% Hispanic, 7% Asian, and 7% other ethnicities. These students were contacted three additional times elec-
tronically and asked to complete a variety of different self-report outcomes and re-
actions to their college experience. Each time, they were offered and paid a $20 gift
certificate from Amazon.com. Academic self-efficacy measures described next
were collected at Time 4 (in November of their 2nd year in college). Cumulative
GPA measures were collected from university personnel at the end of the first se-
mester of their sophomore year (or at the end of four quarters). Complete data were
available from 891 of the original sample, of whom 65% were women and 35%
were men. Five percent of the respondents at Time 4 were Hispanic American,
12.8% were Asian, 8.1% were African American, and 65.8% were White. Eight
percent of the sample did not respond to the ethnic status question.

Measures

The Emotional Stability measure was the 10-item short form version of the Inter-
national Personality Item Pool (Goldberg, 1999). This measure is psycho-
metrically comparable to other commonly used measures of the Big Five con-
structs, such as the NEO-PI (Costa & McCrae, 1992). Goldberg reported the
coefficient alpha for the Emotional Stability scale to be .84. In this study, coeffi-
cient alpha was .87.

Academic self-efficacy was measured with a four-item scale created for this
study. Each item was answered on a 5-point Likert-type scale ranging from 1
(strongly disagree) to 5 (strongly agree). One example item is, “I am confident in
my ability to succeed as a college student.” Coefficient alpha for this measure in
this study was .76. Items were contextual modifications of similar items used to
measure self-efficacy in other studies. Consistent with Bandura’s (1991) conten-
tion that self-efficacy is domain-specific, our self-efficacy items all referred to the
students’ efficacy with respect to their academic performance in college. They are
also Likert-type items consistent with the research reported by Maurer and Pierce
(1998). These authors showed that this relatively simple approach to the measure-
ment of self-efficacy displayed similar measurement properties and validity as did
the more complicated measures assessing both magnitude and strength used by
other researchers (Lee & Bobko, 1994) and Bandura (1986).

As previously above, college GPA for each participant was collected from uni-
versity officials at the student’s institution. This was a cumulative index over the
first three semesters or first four quarters of the students’ college grades.

Data Analyses

To test H1, we regressed academic self-efficacy on the centered measure of emo-
tional stability, gender, and their product. The significance of the product term in
this regression constitutes a test of the hypothesized interaction. Because this hy-
pothesis assumes that the emotional stability measure addresses similar constructs
across gender groups, this regression analysis was preceded by tests that confirmed
the measurement equivalence of the emotional stability measure for men and
women using confirmatory factor analyses (Vandenberg & Lance, 2000).

Tests of H2 and H3 were conducted using the Baron and Kenny (1986) ap-
proach to tests of mediation. GPA was regressed on self-efficacy on Step 1 and
emotional stability, gender, and the interaction term on Step 2. We expected that
the variables entered at Step 2 would not be significant predictors in the full regres-
sion indicating mediation of the interaction-GPA relationship and self-efficacy.

RESULTS

Descriptive statistics for male and female participants in our sample as well as the
combined group are presented in Table 1. The table indicates that there are the ex-
pected significant correlations (and corresponding mean differences) between
gender and academic self-efficacy and self-efficacy and GPA. There are also larger
differences ($d = .57$) between gender groups on emotional stability.

Prior to regression analyses, we conducted tests of the equivalence of the mea-
sure of emotional stability across gender groups. A model that hypothesized that
all 10 items in this measure loaded on a single latent variable and that constrained
factor loadings and error variances associated with the item measures equal across
covariance matrices generated from male and female responses fit the data well,

<table>
<thead>
<tr>
<th>Variable</th>
<th>Female M</th>
<th>SD</th>
<th>Male M</th>
<th>SD</th>
<th>Self-Eff.a</th>
<th>ES</th>
<th>GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-eff.</td>
<td>3.93</td>
<td>.64</td>
<td>4.02</td>
<td>.59</td>
<td>1.00</td>
<td>.06</td>
<td>.35</td>
</tr>
<tr>
<td>ES</td>
<td>3.13</td>
<td>.74</td>
<td>3.53</td>
<td>.64</td>
<td>.23</td>
<td>1.00</td>
<td>-.06</td>
</tr>
<tr>
<td>GPA</td>
<td>3.22</td>
<td>.55</td>
<td>3.22</td>
<td>.59</td>
<td>.22</td>
<td>.03</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total sample</th>
<th>Genderb</th>
<th>Self-Eff.</th>
<th>ES</th>
<th>Interaction</th>
<th>GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1.65</td>
<td>.48</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-Eff.</td>
<td>3.96</td>
<td>.63</td>
<td>-.07</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ES</td>
<td>3.26</td>
<td>.73</td>
<td>-.26</td>
<td>.19</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>5.37</td>
<td>1.81</td>
<td>.72</td>
<td>.09</td>
<td>.46</td>
<td>1.00</td>
</tr>
<tr>
<td>GPA</td>
<td>3.22</td>
<td>.58</td>
<td>.00</td>
<td>.26</td>
<td>.00</td>
<td>.01</td>
</tr>
</tbody>
</table>

Note. Self-eff. = self-efficacy; ES = emotional stability; GPA = grade point average.

aThe correlations above the diagonal are those for the male group; those below the diagonal are for
the female group. Correlations above .11 are statistically significant, $p < .05$. bCorrelations above .06
are statistically, $p < .05$. 

TABLE 1
Descriptive Statistics for the Total Sample and for Men and Women
\( \chi^2(88) = 557.66, \, N = 891, \, p < .001 \) (root mean square error of approximation [RMSEA] = .07, NNFI = .97, CFI = .97). Equally important, a model that allowed for separate estimates of factor loading and error variances fit the data only slightly better than the constrained model. The \( \chi^2 \) difference was statistically significant, \( \chi^2(22) = 68.43, \, N = 891, \, p < .01 \), but the fit indices were nearly identical (RMSEA = .06, non-normed fit index [NNFI] = .97, comparative fit index [CFI] = .97). Examination of individual parameter estimates for the two groups revealed only small differences between subgroup standardized factor loadings and error variances (i.e., all were < .10). The conclusion was that equivalent constructs underlie the measurement of emotional stability in male and female subgroups.

A regression of self-efficacy on gender, emotional stability, and their product provided a test of H1. The results are displayed in Table 2. Gender is significantly related to self-efficacy, but most important for H2 is the significance of the Gender \( \times \) Emotional Stability interaction. This interaction is depicted in Figure 3, and as can be seen, the figure is consistent with the nature of the hypothesized interaction in Figure 2. There is a positive relationship between these two variables for female participants, but the relationship for male participants is nearly zero.

Given our relatively large sample size and the relatively small increment in \( R^2 \) due to the interaction, an important question relates to the practical significance of the interaction displayed in Figure 3. To address this question, we computed the standardized difference in self-efficacy between male and female students at 1 and 2 standard deviations above and below the mean on emotional stability. At 1 and 2 standard deviations below the emotional stability mean, the male–female difference in self-efficacy was .24 and .48 standard deviation units, respectively. Male–female differences at 1 and 2 standard deviations above the mean were .11 and .22, respectively, in this instance favoring female students rather than male. These differences are considered low to moderate as defined by Cohen (1977). These effect sizes reflect the impact of gender, emotional stability, and their interaction on self-efficacy, which in turn is hypothesized to affect actual student GPA.

### TABLE 2
Regression of Self-Efficacy on Gender, Emotional Stability, and Their Interaction

<table>
<thead>
<tr>
<th>Variable</th>
<th>Reg. Wt.</th>
<th>Std. Reg. Wt.</th>
<th>( R^2 ) Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>(-.51^*)</td>
<td>(-.39)</td>
<td></td>
</tr>
<tr>
<td>Emotional stability</td>
<td>(-.08)</td>
<td>(-.10)</td>
<td>(.038^*)</td>
</tr>
<tr>
<td>Interaction</td>
<td>(.14^*)</td>
<td>(.41)</td>
<td>(.005^*)</td>
</tr>
<tr>
<td>Constant</td>
<td>(3.48^*)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Reg. wt. = regression weight; std. = standardized.

\(^*p < .05\).
The regression weight for self-efficacy displayed in Table 3 indicates that about one fourth of this effect is translated into individual differences in GPA.

Results of a second regression testing H2 and H3 are displayed in Table 3. As expected, self-efficacy is significantly related to GPA. Consistent with H3, the effect of gender, emotional stability, and the interaction of these two variables is nonsignificant and near zero, indicating complete mediation of any effect of gender, emotional stability, or their interaction on GPA by self-efficacy.

![Interaction effect of gender and emotional stability on self-efficacy.](image)

**FIGURE 3** Interaction effect of gender and emotional stability on self-efficacy.

The regression weight for self-efficacy displayed in Table 3 indicates that about one fourth of this effect is translated into individual differences in GPA.

Results of a second regression testing H2 and H3 are displayed in Table 3. As expected, self-efficacy is significantly related to GPA. Consistent with H3, the effect of gender, emotional stability, and the interaction of these two variables is nonsignificant and near zero, indicating complete mediation of any effect of gender, emotional stability, or their interaction on GPA by self-efficacy.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-efficacy</td>
<td>.25*</td>
<td>.27*</td>
<td>.068*</td>
</tr>
<tr>
<td>Gender</td>
<td>-.12</td>
<td>-.13</td>
<td>-.10</td>
</tr>
<tr>
<td>Emotional stability</td>
<td>-.10</td>
<td>-.13</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>.04</td>
<td>.12</td>
<td>.002</td>
</tr>
<tr>
<td>Constant</td>
<td>2.58*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Reg. wt. = regression weight; std. = standardized.
$p < .05.$
DISCUSSION

The results reported in this article have several important theoretical and practical implications. As reported in some earlier studies, women report significantly lower levels of self-efficacy when in fact men and women have identical average grades. Women also report lower levels of emotional stability which is related to self-efficacy for women, but not for men. As is repeatedly reported, self-efficacy is related to performance. The combination of these factors accounts for the results of our regression analyses and suggests that emotional stability contributes to higher levels of self-efficacy among women but not men. Because self-efficacy is a relatively important determinant of performance, this combination of effects has a potential negative impact on women relative to men.

Theoretically, these results suggest that personality may have an indirect impact on performance even when the zero order correlations are near zero. In this case, emotional stability is a precursor of self-efficacy, which in turn appears to be a proximal determinant of academic performance. What makes this effect more interesting is the fact that this causal sequence appears to be more important for women than men. Recent research on personality–performance relationships (e.g., Barrick & Mount, 1991) has rekindled applied interest in the use of personality measures in selection contexts. At the same time, Mount et al. (2003) called for closer examination of the manner in which personality might impact performance. It is argued that these “process” investigations might provide more meaningful explanations of relationships between abstract measures of individual differences and human performance. This position is further supported by the Judge and Ilies (2002) meta-analysis, which documents the impact of various personality constructs on the central concepts in three motivational theories. Such investigations may also provide more evidence for the importance of personality constructs in the understanding and prediction of human performance than the simple zero order validity coefficients that are usually reported in personnel selection research.

The results relative to gender differences may also be relevant to stereotype threat research in that we find gender differences in the extent to which emotional stability is related to self-efficacy. Research relating measures of these constructs to the magnitude of a stereotype threat effect may help us understand the nature of this effect. Of course, extension of similar research to minority–majority subgroup differences may also provide interesting insights into stereotype threat phenomena. It should be pointed out, as mentioned in the introduction, that we did not manipulate or measure stereotype threat. However, stereotype threat is thought to be anxiety provoking (Nguyen et al., 2003). The results presented in this article are consistent with the notion that stereotype threat would produce more anxiety for those low in emotional stability (i.e., women) and that the relationship between emotional stability and self-efficacy would be greater for women than men. Although this represents a plausible explanation of the data described in this article,
future research that includes measures related to this hypothesized explanation is warranted.

An obvious question arising from the research reported in this article is what can be done to enhance the emotional stability and self-efficacy of female students. This is a more difficult question most likely rooted in the differential socialization of men and women. If it is a socialization issue, then cross-cultural research replicating the findings reported in this article might be interesting. One obvious implication is that interventions to increase the realistic self-efficacy of students (possibly members of both genders) may be useful. Although the positive correlation between emotional stability and self-efficacy for women suggests that these differences are trait based and may be resistant to change, Bandura (1997) and colleagues have reported numerous successful interventions designed to enhance self-efficacy.

The research reported in this article has at least two obvious limitations. First, we have only correlational data and the models we evaluated suggest causal mechanisms that operate over time. Our data are longitudinal in the sense that emotional stability measures were collected prior to college attendance, but academic self-efficacy and grades were collected later and nearly simultaneously. Because students in some of these courses were likely receiving feedback on their performance during the course, the self-efficacy measure may have been influenced by this feedback. The knowledge of their relative success or failure in the course may have influenced their responses to the self-efficacy items. Thus, it is obvious that causal statements need to be made with caution.

Second, the generalizability of our results might be questioned. Our sample is relatively large and represented students from a variety of institutions (i.e., large state universities, private institutions, and historically Black colleges) of various backgrounds. However, there was also no attempt to randomly sample from any population of American students. It was also the case that a smaller proportion of African American students responded to our instruments at the last data collection time than was true of the initial data collection. Race was not correlated with the study variables with the exception of GPA. On this variable, all racial subgroups that responded to the last survey were superior in GPA to those in the original sample. Another generalization question would be the extent to which these results apply to other performance domains, in particular, work contexts. Replication of this study in work or other contexts would have interesting theoretical and practical implications.

In summary, the relationship between emotional stability and self-efficacy appears to be stronger for women than men, and this interaction does have a relatively small but statistically significant indirect effect on academic performance. This study then contributes to the growing body of research (Barrick et al., 2003; Barrick et al., 1993; Judge & Ilies, 2002) on the processes whereby personality impacts performance.
ACKNOWLEDGMENTS

The support of the College Board in the conduct of the research described in this article is acknowledged and appreciated. The data collection and management support of my colleagues Fred Oswald, Brian Kim, Alyssa Friede, Stephanie Merritt, and Anna Imus was also an important contribution to the work described in this article.

REFERENCES

Goldberg, L. R. (1999). A broad-bandwidth public-domain personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf


