Gender and Affect: Linguistic Predictors of Successful Academic Performance among Economically Disadvantaged First Year College Students

Richard Lewine, Ashlee Warnecke, Darlene Davis, Alison Sommers, Kayla Manley, and Ben Calebs

University of Louisville

Received 12 March 2018; Accepted 22 October 2018

Individuals entering college from disadvantaged economic backgrounds often face multiple obstacles to successful academic performance. Nonetheless, many such students are successful. In this study, we explore the personal characteristics of students from poverty who do well academically in comparison to their economic peers who were less successful academically. Pre-admission, written applications were analyzed using the computerized linguistic analysis tool, LIWC, to predict first semester GPA in a group of 48 students, all of whom came from economic backgrounds that were 150% or more below Federal guidelines. Significant poverty level and sex differences were found. Men's GPA was highly correlated with Total Word Count, while women's GPA was significantly correlated with Reward and Tentativeness. Most striking was the strong positive correlation between GPA and Positive Affect among women from the lowest economic group. The findings suggest further research to clarify and confirm the role of cognitive styles and affect in academic performance as moderated by both sex and degree of poverty, even among those traditionally viewed as belonging to a homogenous economic group.

INTRODUCTION

Contemporary college students face a range of challenges to academic success, including inadequate preparation, anxiety and depression, social heterogeneity, and an uncertain future (Harackiewicz et al., 2002; Hunt & Eisenberg, 2010; Krieg, 2013; Richard-son et al., 2012; Simon, 2017). As economic disparity grows, many students start college at a disadvantage. Education researchers have reported, for example, that students from economically impoverished backgrounds come to campus less prepared academically, with less economic and social capital, and experiencing greater stress than their more well-off peers (Beegle, 2000; Borrego, 2008; Brock, 2010; Cabrera & La Nasa, 2001; Duffy, 2007; Goldrick-Rab, 2006; Green, 2006; Howard, 2001; Kahnlen-berg, 2004; Lareau, 1987; Stinebrickner & Stinebrickner, 2002; Ting, 1998; Walpole, 2003; Yoder, 2005). As completion of college has been shown to provide social and economic advantages for financially disadvantaged students (Hout, 2012), it is important that we understand what contributes to success in this group of students. We should note that “impoverished” background is not limited to money; obviously, there are major disadvantages associated with living in violent neighborhoods, having a seriously dysfunctional family, exposure to substance abuse, etc. The tricky part for research is untangling causality as there is a high empirical correlation among these variables. In terms of college and in this article, however, economic poverty is the primary focus.

Not surprisingly in the face of these obstacles, data suggest that students from poor backgrounds take longer to graduate, have lower grades, and drop out more often than others. It is important to note, however, that this is only part of the story, based on a nomothetic approach. That is, the studies cited above all compare poor students as a group to non-poor students as a group. When studies have examined how and why economically disadvantaged students do well (the implicit assumption being that this is a surprise), it is generally attributed to “resilience,” an often ambiguous term reflecting the fact that students have done well who were not expected to do so (Bonanno & Mancini, 2008; Fergus & Zimmerman, 2005; Infurna & Gress, 2010; Luecken & Gress, 2010; Luther et al., 2000; Waxman et al., 2003).

This report is from an on-going, longitudinal study of economically disadvantaged students pursuing academic degrees at a large, public, four-year research university in the mid-South. We are examining a broad range of potential predictors of college performance, including high school GPA and standardized test scores, previous life experiences, students’ perceived academic success at admission, history of stress prior to college admission, and affective and cognitive styles. The goal is to examine variables that distinguish economically disadvantaged students who do well from those who struggle. In contrast to programs addressing the needs of economically disadvantaged students that approach the problem as a “one size fits all” institutional support requisite (Tinto, 2012), our broader goal is to ask what strengths allow otherwise disadvantaged students to perform well and how those strengths might be supported and encouraged.

In this analysis, we focus on the predictive utility of students’ linguistic characteristics reflected in college application essays using the Linguistic Inquiry and Word Count (LIWC). The LIWC is a computerized, formal analysis of word usage based on a fixed dictionary of words that have been assigned to categories covering cognition, affect, motivation, and social characteristics (Pennebaker & King, 1999; Pennebaker & Graybeal, 2001; Pennebaker et al., 2003). With the exception of Total Word Count (TWC), each characteristic is expressed as a percent of total words. Various linguistically expressed characteristics have been correlated with a wide range of health, personality, and functional behaviors (see Pennebaker et al. 1999; 2001; 2003; 2014; 2015).

Of particular relevance to our study, LIWC analysis has been used with college writing samples to predict academic performance (grade point average, GPA) with some success. Robinson et al. (2013) examined the utility of predicting term GPA using the LIWC to analyze written samples from students taking the same course. Students in a personality course were asked to write a brief “self-introduction” with the resulting LIWC scores correlated with final course GPA. Four variables correlated significantly with course GPA: Total Word Count (.11), Certainty (.11), Negative Emotion (.11), and Present Tense (-.11). TWC has been interpreted as reflecting “... intellectual engagement with the world...” (Robinson et al, 2013, p. 8) and as reflecting complexity of thinking (Graybeal et al, 2002; Vercellonne-Smith et al,
university program targeting students who meet Federal guidelines for poverty were all admitted to the college support program (CCP), a local IRB approved the study. Participants.

further study. To our knowledge there are no LIWC analyses predicting academic performance that have examined sociodemographic factors other than sex/gender as possible moderator variables. The data regarding sex/gender differences are somewhat inconsistent but tend to show male students exhibiting more certainty and assertiveness and female students exhibiting more social orientation and uncertainty in their writing (Pennebaker al al, 2003). There do not appear, however, to be any studies that examine how such linguistic characteristics may differentially predict academic performance by sex. We therefore extend the research by using formally analyzed admission essays to predict college GPA separately for female and male students and by focusing on a group of economically disadvantaged students. We ask what characterizes the essays of those students who might be expected to struggle in college, but instead do well.

**METHOD**

**Participants.** The local IRB approved the study. Students in the study were all admitted to the college support program (CCP), a university program targeting students who meet Federal guidelines for poverty (Table 1), qualify for standard admission to the university, and file for external scholarships. In fact, all the students fell below a family income of $60,000 per year (See sample description below). CCP provides the full cost of tuition, room and board, books, and a work-study job for four years. A total of 58 students were admitted in the Fall 2016 CCP class; one student was under 18 years of age and therefore not eligible per protocol. Of the 57 remaining students, 54 (33 women, 19 men, 2 sex not listed) completed the study measures and were given a letter describing the study, procedures for withdrawing, and contact information for addressing concerns and questions. Pre-admission application essays were available for 48 students. Occasional data are missing due to students choosing not to answer an item.

**Procedures.** Prior to admission, students submitted an application essay that addressed two questions: (1) How will receiving the CCP impact your ability to go to college? (2) How will receiving a college degree impact your future? Questionnaire (sociodemographic information, life events, affect, cognitive style) packets were distributed and completed during orientation week for those students admitted to the university and entering the CCP (see Davis, 2017; Warnecke, 2017 for full details of the study). Term GPA was obtained from student academic transcripts at the end of the first semester. All data were de-identified for analysis in accordance with IRB requirements.

**Measures.** Essays were scanned, submitted to Optical Character Recognition, and converted to Word documents for computer analysis (LIWC). Analyzable essays were available for 48 (28 females; 18 males; 2 sex not listed) of the 54 students who had complete data sets. Given the current sample size and the large number of LIWC variables (approximately 370; see <https://utexas.box.com/s/9ncte8lmq5sxemw3q1x> for manual and complete list of variables), we selected 17 variables that were of immediate interest and/or had been reported by others to be correlated with academic performance and which reflected cognitive style, affective style, and motivation (see Table 2 for list and word examples). Our selection of variables is conceptually consistent with Richardson et al’s (2012) emphasis on non-intellectual factors that predict academic performance. They argue that admission to college reduces the intellectual differences among students making personality, motivation, and self-regulation more predictive. Intellectual functioning is represented in total word count (WC) which has been associated with academic performance as reviewed above.

As family incomes for the CCP students fell well below Federally determined poverty levels, they were coded into four categories: less than $9,999; $10,000-$19,999; $20,000-$39,999; and $40,000-$59,999. In light of the substantial difference in poverty levels among the families of students, an exploratory analysis examining GPA and LIWC patterns was conducted for extreme poverty, EP (< $9,999) and moderate poverty, MP (/> $10,000) students, separately by sex. Due to the small number of extremely poor men (n = 2), three groups were examined: extreme poverty women, EPW (n = 13), moderate poverty women, MPW (n = 16), and moderate poverty men, MPM (n = 14).

Five of the 32 students responding to the “highest parental education” item reported less than a high school education. Of the remaining students, college was attended by at least one of their parents. Of these, three received an associate degree, two a bachelor degree, and one a professional degree. Level of poverty was not associated with parental education.

We obtained term GPA from students’ academic transcripts at the end of the Fall 2016 semester.

**RESULTS**

**Sample description.** Female students were almost identical in age to male students, 18.0 (0) and 18.2 (.5) for women and men, respectively. Approximately half of both sexes were White/European (54.5% of women and 57.9% of men). The remainder of students were distributed among African-American/Black (13.2%), Hispanic/Latino (5.7%), Asian/Pacific (9.4%), and Mixed (13.2%). With respect to family income, however, significantly more female students (50%) came from families at the lowest end of the income scale (< $9,999) than male students (17.6%; p = .027).

All students had strong high school GPAs, with men having a mean of 3.6 (.43) and women a mean of 3.7 (.37; n.s.). Women had a higher first term college GPA (3.0; .76) than men (2.5; 1.26; p = .082). As reflected by the GPA standard deviations, within group variability of academic performance increased from high school to college for both the male (F<sub>17,13</sub> = 2.930, p < .025) and

<table>
<thead>
<tr>
<th>Table 1. Financial Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number in Family</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
</tbody>
</table>

*For each additional person add, $6,240*
female students ($F_{31,1} = 2.054, p < .05$), with somewhat greater variability among the men [$F_{31,17} = 1.658, p = .10$]. Familial poverty was not associated with GPA, 2.8 (.79) and 2.9 (1.09) for extreme and moderate poverty, respectively. We examined the distribution of GPAs and found no significant sex differences. GPAs ranged from 0.0-3.98 for men and 1.34-4.00 for women (Pearson chi-square, df = 92, $p = .194$). No students stood out as outliers who might have contaminated the correlational analyses.

**Linguistic characteristics.** Table 2 summarizes the means (standard deviations) of the LIWC measures separately for each sex. Three (fewer than one expected by chance) sex differences reached a conventional level of statistical significance ($p < .05$): men’s essays exhibited more positive emotion and risk taking compared to women’s. 

### Table 2. Means (standard deviations) of LIWC variables by sex*

<table>
<thead>
<tr>
<th>LIWC</th>
<th>Women (n=28)</th>
<th>Men (n=18)</th>
<th>t</th>
<th>df</th>
<th>sig</th>
<th>Word Examples**</th>
</tr>
</thead>
<tbody>
<tr>
<td>WC</td>
<td>513.4 (150.3)</td>
<td>516.3 (153.8)</td>
<td></td>
<td></td>
<td></td>
<td>happy, cried</td>
</tr>
<tr>
<td>Affect</td>
<td>4.3 (1.2)</td>
<td>4.8 (1.2)</td>
<td></td>
<td></td>
<td></td>
<td>love, nice, sweet</td>
</tr>
<tr>
<td>PosEmotion</td>
<td>3.2 (1.1)</td>
<td>3.9 (1.3)</td>
<td>2.25</td>
<td>44</td>
<td>.029</td>
<td>hurt, worried, hate, crying</td>
</tr>
<tr>
<td>NegEmotion</td>
<td>1.0 (0.6)</td>
<td>8.6 (1.6)</td>
<td></td>
<td></td>
<td></td>
<td>mate, talk, they</td>
</tr>
<tr>
<td>Social</td>
<td>7.6 (2.4)</td>
<td>6.3 (1.8)</td>
<td>-2.03</td>
<td>44</td>
<td>.048</td>
<td>cause, know, ought</td>
</tr>
<tr>
<td>CogProc</td>
<td>11.8 (2.3)</td>
<td>12.2 (3.6)</td>
<td></td>
<td></td>
<td></td>
<td>think, know</td>
</tr>
<tr>
<td>Insight</td>
<td>2.4 (7.1)</td>
<td>2.1 (8.1)</td>
<td></td>
<td></td>
<td></td>
<td>maybe, perhaps</td>
</tr>
<tr>
<td>Tentative</td>
<td>1.7 (8.1)</td>
<td>2.4 (1.4)</td>
<td></td>
<td></td>
<td></td>
<td>always, never</td>
</tr>
<tr>
<td>Certainty</td>
<td>1.6 (8.1)</td>
<td>1.4 (7.1)</td>
<td></td>
<td></td>
<td></td>
<td>ally, friend, social</td>
</tr>
<tr>
<td>Affiliation</td>
<td>101.1 (2.1)</td>
<td>100.0 (2.8)</td>
<td></td>
<td></td>
<td></td>
<td>win, success, better</td>
</tr>
<tr>
<td>Achievement</td>
<td>2.4 (1.5)</td>
<td>2.1 (1.0)</td>
<td></td>
<td></td>
<td></td>
<td>superior, bully</td>
</tr>
<tr>
<td>Power</td>
<td>3.6 (1.2)</td>
<td>3.9 (1.5)</td>
<td></td>
<td></td>
<td></td>
<td>take, prize, benefit</td>
</tr>
<tr>
<td>Reward</td>
<td>3.6 (9.1)</td>
<td>3.4 (1.1)</td>
<td></td>
<td></td>
<td></td>
<td>danger, doubt</td>
</tr>
<tr>
<td>Risk</td>
<td>2.2 (8.1)</td>
<td>2.8 (8.1)</td>
<td>2.31</td>
<td>44</td>
<td>.025</td>
<td>today, is now</td>
</tr>
<tr>
<td>PastFocus</td>
<td>.3 (2.3)</td>
<td>.3 (3.3)</td>
<td></td>
<td></td>
<td></td>
<td>may, will, soon</td>
</tr>
<tr>
<td>PresentFocus</td>
<td>3.5 (2.1)</td>
<td>2.8 (1.5)</td>
<td></td>
<td></td>
<td></td>
<td>may, will, soon</td>
</tr>
<tr>
<td>FutureFocus</td>
<td>10.4 (2.2)</td>
<td>11.0 (2.2)</td>
<td></td>
<td></td>
<td></td>
<td>may, will, soon</td>
</tr>
</tbody>
</table>

*Two students chose not to indicate sex **Pennebaker et al (2015)

was positively correlated with GPA ($r = .54, p = .046$) among MP women.

Men in the MP group were quite similar to MP women in the correlation between Affect and GPA, $r = -.56, p = .039$. Neither Reward ($r = -.25$) nor Certainty (.07) were statistically significant, although WC and GPA were highly correlated among MP men ($r = .69, p = .007$).

### DISCUSSION

Not surprisingly, there is an increase in inter-student variability in academic performance from high school to first term college, variability being more salient among male than female students. As it has been suggested that selection criteria reduce intellectual characteristics among students admitted to college (Richardson et al., 2012), that the variance is more marked for men than women, raises the possibility that selection criteria are not as effective for men as for women. We may further speculate that men are less-well prepared in non-intellectual areas and/or that female students may be likely to partially mask their intelligence due to social pressure (as we discuss below).

**Intellectual characteristics.** There was a strikingly positive relationship between total word count and term GPA among male students in our study, whereas word count and GPA were independent among female students. This finding partially fits the proposed model suggested by Pennebaker and colleagues that contemporary higher education rewards categorical thinking (hypothesized to be reflected in total word count), although in our sample only for men. In contrast, female students tended to do better when more tentative and less externally rewarded. Why should this be? Are there gender differences in learning style and/or gender-influenced differences in the way instructors interact with students that explain this gender difference? Tentativeness has typically been interpreted to reflect insufficient or inadequate cognitive complexity (Joksimovic et al., 2015; Taussczik & Pennebaker, 2010). This position would not predict our findings; indeed, the findings of this study counter the traditional interpretation. We offer a competing interpretation based on gender styles. On the basis of our own advising and teaching experiences, we hypothesize that cultural progress aside, women more often than men feel pressure to couch their intellectual abilities in socially acceptable postures. Like the “Resting Bitch Face,” unmasked intellectual expression may be viewed as hostile and intrusive when engaged in by women, but assertive and helpful when engaged in by men. These sorts of interactions should be addressed in future studies.

**Affect and academic performance**

**Overall GPA and affect.** There is a substantial body of evidence supporting the role that affect plays in cognition generally and academic performance more specifically. Interaction between affect and cognition has been reported from emotional intelligence (Elder, 1996) to differential cognitive strategies directed by posi-
positive and negative affect (Fredrickson & Joiner, 2002; Fredrickson & Losada, 2005) to the association of depression and anxiety and college grades (Lewine et al., 2011; Lewine et al., 2015; Lewine & Sommers, 2016). Of particular note from our earlier studies, the relationship between affect and academic performance may be moderated by sex, and as found in the current study, further moderated by level of poverty.

Specifically, the relationship between Affect and GPA is strongly moderated by level of poverty. Among the extremely poor, the higher the Affect, the higher the GPA; among the moderately poor, the higher the Affect, the lower the GPA (among both men and women). As the LIWC measure Affect includes both positive and negative affect, we examined the relative contributions of each to GPA. Positive affect clearly emerged as the more important contributor (r = .50) than negative affect (r = .18) among the extreme poverty group. In contrast, positive and negative affect contributed equally to the Affect-GPA relationship in the MP women (r = -.27 and -.27 for positive and negative emotion, respectively) and MP men (r = -.35 and -.22 for positive and negative emotion, respectively). We note that affect in this analysis reflects an individual’s emotional tone at the time of the essay before admission to and attendance at college, thereby providing genuinely predictive information.

We must, of course, acknowledge the limitations of this study and initial report. The analyses are largely exploratory in nature and large in number, which in combination with a modest sample size biases the finding of significant correlations. By design, the sample is restricted to those students who come from economic disadvantage thereby restricting the generalizability of the findings.

Moving forward, it will be important to include students from a much broader range of economic backgrounds while remembering not to categorically lump students into “poor” and “not poor” groups. In addition, the potential moderating role of urban and rural living environments on economic standing needs to be addressed, an issue we are examining as we follow the cohort reported in this study through its academic career. And ideally, others will begin to examine similar issues among their students. In defense of the findings, we point to the unique nature of the sample, the use of computerized analyses of pre-admission writing, and the longitudinal design that will yield within subject, repeated measures as the study progresses.

SUMMARY AND IMPLICATIONS

Men and women appear to exhibit different characteristics important to academic success. Even more striking is the role of positive affect in predicting academic success in female students from particularly harsh economic backgrounds. This suggests that positive affect may be a significant component of resilience in this group of students and serves as an important reminder that poverty as conceptualized by current federal standards does not yield a homogeneous group with similar characteristics important for academic success. Further, as many have suggested (e.g. Brookfield, 1987; 2002; Elder, 1996; Halonen, 1995; Linnenbrink, 2007), learning is not just about cognitive strategies or learning styles, but just as importantly about the role of affect. Among students who can be assumed to be motivated to engage in college (as true of the students in our sample), students’ emotional tone prior to and about their admission may well inform future success. Advising and instructing should be informed by the emotional and social context of the individual learner, something more easily acknowledged in principle than in practice. Equally important is the value of recognizing the heterogeneity of poverty, with poverty level, sex, and affect deserving careful study in the future.

ACKNOWLEDGEMENTS

We would like to thank those individuals who have supported and actively participating in the Cardinal Covenant program and quite literally without whose help and support this program of research would never have happened: Dale Billingsley, PhD, Chelsea McKendree, MA, Tony Robinson, MSW, and Bernadette Walter, PhD.

REFERENCES


