What can we learn about college retention from student writing?

Daniel McCaffrey, Steven Holtzman, Jill Burstein & Beata Beigman Klebanov
Educational Testing Service
{dmccaffrey, sholtzman, jburstein, bbeigmanklebanov}@ets.org

ABSTRACT: Low retention rates in college is a policy concern for US postsecondary institutions, and writing is a critical competency for college (Graham, 2019). This paper describes an exploratory writing analytics study at six 4-year universities aimed at gaining insights about the relationship between college retention and writing. Findings suggest that AWE is useful for exploring the relationship between college retention and writing, and have implications for gathering diagnostic retention analytics from student writing.

Keywords: writing analytics, automated writing evaluation, higher education, retention

1 INTRODUCTION

College retention is an issue of national concern. The U.S. Department of Education, National Center for Educational Statistics (2020) reports that among first-time, full-time undergraduate students who started a 4-year Bachelor's degree in Fall 2012, only 62% completed the degree within six years – i.e., by 2018. Previous research has shown relationships between coursework writing and academic success factors. Harackiewicz et al. (2016) showed that higher utility value scores -- i.e., scores based on how a writer expresses personal relevance about technical material in a STEM writing assignment -- was correlated with STEM course retention. Using writing data from Harackiewicz et al. (2016), Beigman Klebanov et al (2017) showed that *utility value* words (e.g., our, family) were indicative of writing responses with higher human rater *utility value* scores. Allen, Dascalu, McNamara et al (2016) showed how linguistic properties in college students' writing can be used to model individual differences in students' vocabulary knowledge and reading comprehension skills. Burstein, McCaffrey, Elliot et al., (2020) used AWE to examine relationships between writing and broader academic skills and success factors (e.g., college GPA).

The study examines the question: *What relationships exist between college retention and writing*?

2 METHODS

2.1 Participants

Six four-year public universities participated in the study. One site was a Historically Black College, and a second site was a Hispanic-Serving Institution. Data from 418 students enrolled in one of the six sites were included in this study.

2.2 Data

All 418 students submitted one or more *coursework* writing assignments (n=997). Assignments were from one of these courses: first-semester English composition, Business, History, and STEM, and from

argumentative, informative, or reflective genres (Burstein et al, 2019). Median coursework assignment word count was 753. A subset of 366 students completed a *timed, argumentative standardized writing assessment*; median word count was 220.

2.3 Automated writing evaluation (AWE) features

Automated writing evaluation (AWE) tools generate linguistic features from text (see Burstein et al 2017). In this study, AWE tools were used to generate 36 writing features representing six writing subconstructs: *Vocabulary* (e.g., word complexity), *English Conventions* (e.g., grammar errors), *Organization and Development* (e.g., text coherence), *Argumentation* (e.g., claim terms), *Sentence Structure* (e.g., use of clauses), and *Utility-Value language* (i.e., personal relevance terms, such as, "me", "friends"; See Beigman Klebanov, et al 2017). AWE features represent linguistic characteristics in the writing samples. To create a univariate measure for each subconstruct, the feature scores were combined into a weighted composite score. Weights equaled the loadings of the first principal component from a Principal Components Analysis fit separately for each subconstruct. Individual features were centered by genre to have mean zero. The final composite scores were standardized to a mean of zero and a variance of one and averaged across writing assignments to yield one score per composite per student. Analyses were run at the student level, and separately for the assessment and course writing data.

3 PREDICTING DROPOUT

Participating students' enrollment was tracked from 3 to 5 semesters after their participation in the study using administrative data provided by the participating universities. Random effects Cox proportional hazards regression was used to model dropout as a function of the AWE subconstruct composite score, controlling for the students' SAT/ACT score, high school GPA (HSGPA), university, and writing sample length. The models also include random effects for the course-section in which students were enrolled when participating. This accounted for possible unmodelled dropout risk factors associated with different section assignments. Separate models were fit for each feature composite score for coursework assignments, and for standardized writing assessments.

4 RESULTS

Two of the six composite features were predictive of dropout in the regression models; others were not. A standard deviation increase in the *Utility-Value language (UVL)* composite feature predicted a 26% increase in dropout hazard (i.e., dropout probability based on students continued enrollment or graduation) for both coursework ($p < 0.05$) and standardized assessment ($p < 0.10$). In addition, a standard deviation unit increase in the *Vocabulary* (VCB) composite feature in the standardized assessment predicts a 15% decrease in the hazard of dropout ($p < 0.10$). Analyses using individual component features (in the composites) showed dropout risk related positively (more risk) to pronoun use, and negatively (less risk) to use of longer words.

5 DISCUSSION

The relationship between *UVL* and college writing has not been widely studied. Beigman Klebanov et al (2017) found student success positively associated with UVL when writing assignments explicitly elicited utility value. In this study, results suggest that *UVL* could be a valuable predictor for dropout. Reviews of some student writing samples from study participants found *UVL use* reflected difficulty effectively integrating personal elements into academic writing. As discussed earlier, vocabulary has been found to be associated with various measures of academic skills and success. The results from this study extend those findings. The results suggest that exploring *vocabulary usage* with AWE might be used to identify students at risk of dropping out. More research will be required to draw clearer inferences about relationships between use of *UVL* and *VCB,* and college retention. Overall, study findings suggest relationships between AWE feature measures and retention. This insight has implications for AWE as a potential means to gather diagnostic retention analytics for stakeholders who monitor students' progress. For example, we could envision AWE integration into a learning management system in order to provide *not only* personalized learning for writing, but retention analytics for students, educators and other stakeholders to signal success and potential obstacles.

ACKNOWLEDGMENTS

Research presented in this paper was funded by the Institute of Education Science, United States Department of Education, Award Number R305A160115 any opinions, findings, conclusions, or recommendations are those of the authors and do not necessarily reflect the views of the IES.

REFERENCES

Allen, L., Dascalu, M., McNamara, D. S., Crossley, S., & Trausan-Matu, S. (2016). Modeling individual differences among writers using ReaderBench. In EDULearn16: 8th International Conference on Education and New Learning Technologies (pp. 5269-5279). IATED Academy.

Beigman Klebanov. B., Burstein, J., Harackiewicz, J. M., Priniski, S. J., & Mulholland, M. (2017). Reflective Writing About the Utility Value of Science as a Tool for Increasing STEM Motivation and Retention–Can AI Help Scale Up? International Journal of Artificial Intelligence in Education, 27(4), 791-818. https://doi.org/10.1007/s40593-017-0141-4

Burstein, J., McCaffrey, D., Beigman Klebanov, B., & Ling, G. (2017, September). Exploring Relationships between Writing & Broader Outcomes with Automated Writing Evaluation. In Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications (pp. 101-108). Retrieved from https://www.aclweb.org/anthology/W17-5011

Burstein, J., McCaffrey, D., Elliot, N., Beigman Klebanov, B., Molloy, H., Houghton, P. & Mladineo, Z. (2020). Exploring Writing Achievement and Genre in Postsecondary Writing . In Companion Proceedings in the 10th International Conference on Learning Analytics & Knowledge (LAK20), 53-55.

Graham, S. (2019). Changing how writing is taught. Review of Research in Education, 43(1), 277-303. https://doi.org/10.3102%2F0091732X18821125

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2016). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. Journal of personality and social psychology, 111(5), 745.

U.S. Department of Education, National Center for Education Statistics. (2020). *The Condition of Education 2020* (NCES 2020-144), Undergraduate Retention and Graduation Rates.