SAT Does Not Spell Success: How Non-Cognitive Factors Can Explain Variance in the GPA of Undergraduate Engineering and Computer Science Students

Matthew Scheidt
School of Engineering Education
Purdue University
Lafayette, IN
mscheidt@purdue.edu

Ryan Senkpeil
School of Engineering Education
Purdue University
Lafayette, IN
rsenkpei@purdue.edu

John Chen
Dept. of Mechanical Engineering
Cal Poly
San Luis Obispo, CA
jchen24@calpoly.edu

Allison Godwin
School of Engineering Education
Purdue University
Lafayette, IN
godwina@purdue.edu

Edward Berger
School of Engineering Education
Purdue University
Lafayette, IN
bergere@purdue.edu

***Abstract*—This work-in-progress research paper uses multiple regression of both cognitive and non-cognitive factors to model current GPA of engineering and computer science students. High school GPA and ACT/SAT scores are among the most common scores used as admission criteria, which result in a relatively homogeneous engineering population. Prior research, however, shows that these scores do not account for the variance in GPA once these students start undergraduate studies. In this work, we explore how students' cognitive (e.g., study skills, test performance, regulatory behaviors, etc.) and non-cognitive factors (e.g., identity, motivation, personality, etc.) predict student success in their engineering pathways. The data for this initial study comes from a pilot survey, deployed in the summer of 2017, of 490 engineering and computing students from two large, public institutions, one on the West Coast, the other in the Midwest. We used multiple linear regression to control for demographic variables while examining the predictive value of particular cognitive and non-cognitive factors for student academic achievement (i.e., GPA) in university. Our analysis shows, not surprisingly, that standardized test scores only explain a small portion of the variance of undergraduate GPA. Including non-cognitive and affective factors into a regression model produced a marked increase in the explained variance. Our work is novel in examining a constellation of possible factors that predict undergraduate student GPA, by combining both cognitive and non-cognitive factors as predictors. This analysis begins to unpack particular factors that have potential to predict GPA of engineering and computer science students.**

Keywords—student success, multiple regression, non-cognitive

I. Introduction

This work in progress is a continuation of other work exploring non-cognitive and affective factors as predictors of first year engineering student GPA [1]. In that work, we used multiple regression to explore the relationship of personality (from a variation of the Big Five Survey), grit (determined from the Short Grit survey), and test anxiety and time and study environment (determined from portions of the Motivated Strategies for Learning Questionnaire), with first year college GPA. In this paper, we expand on this work by regressing college GPA on a plethora of non-cognitive factors from the SUCCESS survey. A very brief explanation of the items in the SUCCESS survey is included below. A more detailed explanation of the items, including all of the questions used for each construct and why these constructs were chosen over others, along with the exploratory factor analysis conducted on all of the constructs used in this analysis is included in another work [2]. The goal of the SUCCESS project is to map the wide array of latent diversity among engineering and computing students and to identify areas where interventions can be made to help students succeed. This analysis uses pilot data collected as part of this project for survey development to begin to explore the many latent non-cognitive and affective factors that can contribute to student success beyond that of SAT scores.

In this paper, we address the question: can non-cognitive and affective factors better predict the GPA of engineering and computing students beyond that of some of the cognitive measures that are currently used as admission criteria? This project begins to answer the call from the National Academy of Sciences to "address gaps...to address how intra- and interpersonal competencies may be related to students' success in 2- and 4-year science, technology, engineering, and mathematics programs and majors" [3]. Inter- and intrapersonal competencies that they identified include conscientiousness, belonging, mindset, motivation, and identity. We examine these inter- and intrapersonal competencies as well as other factors such as test anxiety and time management in order to begin to explore more robust predictors of student success.

We hypothesize that the additional non-cognitive measures will account for significant additional variance in college GPA above and beyond that of cognitive and previously tested non-cognitive measures, controlling for institution, sex, race, and year of study.

This material is based upon work supported by the National Science Foundation under Grant Nos. DUE-1626287 (Purdue), DUE-1626185 (Cal Poly), and DUE-1626148 (UTEP). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

978-1-5386-1174-6/18/$31.00 ©2018 IEEE

II. METHODS

We collected pilot data electronically via Qualtrics® in the summer of 2017 from students at one West Coast and one Midwest university (*n* = 490). We used this data for exploratory factor analysis of the many different constructs in the survey and found acceptable validity evidence for the items used [2]. We measured: Big 5 personality [4]–[6], Community [7], Grit [8], [9], Thriving [10], Identity [11], Mindset [12], Motivation [13], [14], Time and Study Environment [15], Test Anxiety [15], Perception of Faculty Caring [16], Self-Control [17], Stress [18], Gratitude [19], Belongingness [20], and Mindfulness [21], [22]. All the items that make up each factor were asked on an anchored numeric scale between one and seven, with one generally representing anchors of "strongly disagree" or "least like me," and seven representing anchors of "strongly agree" or "most like me." We combined items that loaded into factors through arithmetic means which became the factors used within this analysis. In this survey we also collected demographic information (race, gender, academic-level) as well as self-reported GPA.

We also collected registrar information. From the registrar, we obtained admissions information such as SAT/ACT scores, demographic information, and transcript data including all of the classes taken by the participant and the grades that they received for those classes. We then combined the results from the survey with the registrar data for analysis. After listwise deletion, (removing all cases that were not complete for all data points), we had a total of 403 participants.

We coded the demographic variables in a way that is useful for this preliminary analysis. Although our survey includes a broad spectrum of gender identities, we use the binary sex identity available from the registrar data with 0 representing male (*n* = 252) and 1 representing female (*n* = 151). Our survey allowed for students to identify as multiple races/ethnicities ("American Indian or Alaska Native" [*n* = 6], "Asian" [*n* = 112], "Black or African American" [*n* = 2], "Hispanic, Latino, or Spanish origin", [*n* = 28], "Middle Eastern or North African" [*n* = 6], "Native Hawaiian or Other Pacific Islander" [*n* = 6], "White" [*n* = 302]), resulting in counts greater than our overall sample size. For students who did not provide race/ethnicity information on the survey, we used the registrar data equivalent in its place. Overall, we were able to obtain race/ethnicity for all but three students. Because of the small number of participants in some of the race/ethnicity categories, we combined those traditionally underrepresented in engineering together in one group, maintaining the Asian and White groups separately. To account for students' multiple potential answers, each race/ethnicity was coded dichotomously with 1 representing identifying as that race and 0 representing not identifying of that race. We also asked our participants what year (academic level) they were in school, with this variable being considered categorical with 1 (*n* = 99), 2 (*n* = 152), 3 (*n* = 58), and 4+ (*n* = 94) representing different years in college.

We checked to ensure that our predictor variables were suitable for multiple regression. With the exception of the demographic variables, all of the predictor variables were relatively normal (|skewness| < 2 and kurtosis < 7). Multicollinearity was also checked using variance inflation factor (VIF) calculations, with all VIF scores for each predictor variable less than 3.6. As a result, no variables were removed due to multicollinearity. VIF of 10 or more indicates serious multicollinearity [23].

We conducted several multiple regression analyses, regressing self-reported GPA onto composite SAT/ACT and non-cognitive and affective factors, while controlling for demographics. To understand the unique variance accounted for by demographics, we regressed GPA onto the institution, year, sex, and race variables. We then hierarchically added composite SAT/ACT score, to determine the unique variance accounted for by this cognitive measure. This model with demographics and SAT/ACT scores became the base model for to assess the additional variance accounted for by non-cognitive and affective factors. R statistical software Version 3.4.4 was used for this analysis [24].

III. RESULTS AND DISCUSSION

A. Demographics as Predictor Variables

In this analysis, we first looked at the unique variance explained by the demographic variables. Initially, we regressed self-reported GPA from the survey onto sex, race, institution, and year in school individually, and then as a set. We found that when GPA was regressed onto sex alone, there was no significant difference by sex. We also found institution to be a non-significant predictor of GPA. Using students in the first year as the reference group, years 2, 3, 4+ had statistically significantly lower GPAs than year 1, (β = -0.14, -0.19, and -0.14 respectively, $p < 0.05$). We found no statistical differences between years 2, 3, and 4+. Students who identified as "White" had statistically significant ($\beta = 0.16$, $p < 0.05$) higher GPAs than their peers. In the overall model regressing GPA onto sex, year, and race, only year in school was a significant predictor.

From the above analysis, as students in our sample advanced throughout college, their GPAs were lower, with students in their third year having the lowest average GPA. The literature about STEM GPA over time is inconsistent, with GPA going up [25] and down [26], after freshman year. Overall, the demographic variables were only able to account for 5.05% ($p < 0.01$) of the overall variance in GPA.

B. Composite SAT/ACT as a Predictor Variable

Regressing self-reported GPA onto the composite SAT/ACT score alone generated a significant ($\beta = 0.25$, $p < 0.0001$) model and predicted 6.44% of the variance in GPA. Adding the composite SAT/ACT score to the model with demographics serving as predictors of GPA is also significant ($p < .001$) explaining 10.26% of the variance. The details of this model are shown in Table I. This model which includes sex, race, and year in school along with SAT/ACT composite score as predictors of self-reported GPA will serve as the baseline by which to compare non-cognitive factors as predictors of college GPA. The only significant predictors were composite SAT/ACT ($\beta = 0.24$, $p < .0001$) and year 3 (β = -0.18, $p < 0.01$). Composite SAT/ACT score alone explained 6.44% of the variance in GPA, and when added into the model

TABLE I. Summary for hierarchical regression analysis for variables predicting self-reported GPA (n= 401), with regression coefficients (B), standardized coefficients (β), and standard errors of the regression coefficients ($SE\ B$).

	Baseline Model			**Baseline with Non-Cognitive Factors**		
Variable	***B***	***β*** [a]	***SE B***	***B***	***β*** [a]	***SE B***
Intercept	3.32	0****	0.111	3.35	0****	0.010
Sex	0.05	0.06	0.042	0.09	0.11*	0.039
Year 2	-0.10	-0.12	0.054	-0.05	-0.05	0.046
Year 3	-0.21	-0.18**	0.069	-0.08	-0.07	0.060
Year 4	-0.10	-0.10	0.062	-0.06	-0.06	0.053
Institution	0.00	0.00	0.053	-0.01	-0.01	0.047
Under-Represented Minorities	-0.02	-0.02	0.070	-0.01	-0.01	0.060
Asian	0.03	0.03	0.067	0.01	0.01	0.058
White	0.12	0.12	0.070	0.02	0.02	0.063
Composite SAT/ACT	0.04	0.24***	0.008	0.03	0.23****	0.007
Stress (Frustrations)				-0.01	-0.03	0.015
Self-Control (Acting)				0.00	0.01	0.017
Self-Control (Restraint)				0.02	0.05	0.025
Perception of Faculty Support				0.03	0.08	0.015
Test Anxiety				-0.04	-0.17***	0.013
Time and Study Environment				0.06	0.14**	0.022
Motivation (Expectancy)				0.12	0.35****	0.022
Big 5 (Conscientiousness)				0.02	0.06	0.014
Grit Perseverance of Effort				0.04	0.09	0.020
Eng. Identity (Perf. Competence)				-0.06	-0.15*	0.027

a. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001

with demographics, explained 5.21% additional variance above that of demographics alone. This suggests that there is some shared variance across composite SAT/ACT score and demographics.

C. *Non-Cognitive Factors as Predictors*

For this portion of the analysis, each construct was initially added individually to the base model, consisting of the demographics and the composite SAT/ACT score, to identify which non-cognitive factors were individually significant predictors, beyond the baseline model. Each of the significant predictors was evaluated to determine the unique variance that was explained by that predictor above the SAT/ACT composite score after controlling for demographics. Proceeding in this manner, a case of suppression (where one variable hides the relationship with another, possibly eliminating irrelevant variance [23]) was observed among latent variables. With mindset, when both increment (growth) and entity (fixed) mindset were in the model, entity mindset significantly contributed negatively to GPA. When increment mindset was removed from the model because it did not contribute significantly, entity mindset became non-significant. For consistency throughout this analysis, because they did not hold significance independently, they were both excluded from the modeling shown in Tables I and II.

All of the factors shown in Table II contributed from 3.05% to 17.82% of the variance in self-reported GPA beyond that of the demographics and the composite SAT/ACT score. We found similar results to previous studies that have used some of these factors to predict GPA. Grit (persistence of effort) was positively related to GPA ($r = 0.27$, $p < 0.001$) with our sample compared with Duckworth, Peterson, Matthews, and Kelly's overall grittiness score correlation [8]. Additionally, Big 5 (conscientiousness) has a positive relationship with GPA ($\beta = 0.17$, $p < 0.001$) which is similar to our previous work [1]. Test anxiety was negatively associated with GPA ($\beta = -0.30$, $p < 0.0001$) and time and study environment was positively associated with GPA ($\beta = 0.34$, $p < 0.0001$), which is also consistent with our previous work [1]. What was unexpected, was finding that when test anxiety was added to the base model, gender became statistically significant ($\beta = 0.13$, $p < 0.01$), also indicating suppression.

The resulting semi-partial change in variance (unique total variance explained by adding a predictor or set of predictors to a model [23]) with the addition of each individual factor is shown in Table II. While the base model is able to only account for 10.26% of the variance, the non-cognitive and affective factors account for considerably more variance above that of the baseline model. Combining all of the non-cognitive and affective factors shown in Table II into one linear regression model accounts for 36.63% of the total variance in GPA. This model is shown in Table I alongside the baseline model.

D. *Non-Cognitive Factors in Explaining Student Success*

This study is unique because we have access to different non-cognitive factors that have not been combined for this type of analysis before. We describe the effects for each of these constructs here. Although a growth mindset has been shown to predict student success through GPA [27], we did not observe a

TABLE II. Unique variance accounted for by base model and each factor added to base model individually.

Variables	R^2 ***[%]***	sr^2 ***[%]***[b]
Base Model	10.26	
Big 5 (Conscientiousness)		3.05
Grit Perseverance of Effort		7.24
Eng. Identity (Perf. Competence)		7.14
Motivation (Expectancy)		17.82
Test Anxiety		7.54
Time and Study Environment		11.09
Perception of Faculty Support		7.38
Self-Control (Acting)		4.13
Self-Control (Restraint)		5.98
Stress (Frustrations)		4.46
All of above in model		26.37

b. All p-values for changes in model <0.001 as calculated through ANOVA

statistically significant relationship. Engineering identity (performance competence) was positively related to GPA (β = 0.28, p < 0.0001). Motivation (expectancy) and motivation (instrumentality) both had significant relationships with GPA (β = 0.51, p < 0.0001 and β = -0.16, p < 0.01, respectively) when in the model together. However, when GPA is regressed on these two motivation factors separately, controlling for demographics and composite SAT/ACT, these relationships change, (β = 0.44, p < 0.0001 and β = 0.05, non-significant, for expectancy and instrumentality, respectively), also suggesting suppression of the effect of instrumentality. We also found that when motivation (expectancy) was added to the base model, gender became statistically significant (β = 0.12, p < 0.01), indicating suppression of the effect of gender as well. Perception of faculty support was positively related to GPA (β = 0.28, p < 0.0001). Self-control (acting) and self-control (restraint) were also related to GPA (β = -0.15, p < 0.05 and β = 0.20, p < 0.0001, respectively) when added to the baseline model together and also individually (β = -0.21, p < 0.0001 and β = 0.25, p < 0.0001, respectively). The increase in effect size and statistical significance indicates shared variance across the self-control factors. Finally, stress (frustration) was negatively associated with GPA (β = -0.22, p < 0.0001).

When we regressed self-reported GPA on all of the factors shown in Table II, we were able to explain an additional 26.37% (p < 0.0001) of the variance beyond that of the baseline model. This result, shown alongside the baseline model in Table I, confirms our hypothesis that non-cognitive measures account for significant additional variance in college GPA above and beyond that of cognitive and previously tested non-cognitive measures, controlling for institution, sex, race, and year of study. These relationships illustrate the complex relationships between different non-cognitive and affective factors as predictors of student success.

IV. Limitations

There are several limitations within this work in progress, some of which will be addressed as part of our future study into predicting student success. The first is that in this paper we use GPA as a measure of success. While this is not unfounded from prior literature, there are many other conceptualizations of student success that we did not consider. Some of those include persistence to graduation or post-college achievement. Another limitation of this work is that we do not include high school GPA as an independent variable. Other work looking at non-cognitive factors as predictors of GPA showed that high school GPA was a significant predictor of first year engineering college GPA [1]; however, we could not access this variable across all of the institutions in the dataset. Additionally, more than ACT/SAT scores and high school GPA are used for entrance into engineering programs. For example, one study named "subject matter expectations…high school class rank…strength of HS curriculum…[and] personal background and experience" [7, p. 279] among others as factors that are considered when admitting students. We also recognize that our categorization of students' gender (binary) and race/ethnicity (one group for traditionally underrepresented students) is not as reflective of the overall population included in the sample and does not currently reflect the full richness of gender and racial identification within our dataset. A larger sample in the future may be able to address this issue. In spite of these limitations, this analysis shows potential for using non-cognitive and affective factors as predictors for student success beyond that of SAT/ACT.

V. Conclusion

This regression analysis reveals many new insights into potential predictors of student success, above and beyond that of composite SAT/ACT, when controlling for institution, year in school, sex, and race. Motivation (expectancy) and the time and study environment, taken individually, explained over two times variance above our baseline model, suggesting that these two non-cognitive measures serve as better predictors of GPA than SAT/ACT scores. Additionally, when combined with other non-cognitive factors, the model predicted over three times the variance in GPA than SAT/ACT scores and controls for demographics alone. While this study is only preliminary, it begins to present a different and improved way of determining student success. Finally, these results may also have implications for policy such as admittance decisions or allocation of university resources.

Our future work includes also using high school GPA as a predictor of college GPA. Also, as we continue this research, and look across more institutions, we will further probe the finding of decreasing GPA from the first year to see if it is engineering specific, institution specific, or just within this sample of students. We plan to use these findings to explore more relations, such as interaction effects, as well as more accurately represent our participants' identities within the data.

Acknowledgment

We would like to thank our participants for taking this survey, without which this work would not be possible.

REFERENCES

[1] R. R. Senkpeil and E. J. Berger, “Impact of non-cognitive factors on first-year performance,” in *ASEE Annual Conference and Exposition*, 2016.

[2] M. Scheidt *et al.*, “Validity evidence for the SUCCESS survey: Measuring non-cognitive and affective traits of engineering and computing students.,” in *Proceedings of the American Society of Engineering Education Annual Conference & Exposition.*, 2018.

[3] National Academy of Sciences, “Supporting students’ college success: The role of assessment of intrapersonal and interpersonal competencies,” Washington, D.C., 2017.

[4] G. Dam and S. Kaufmann, “Computer assessment of interview data using latent semantic analysis.,” *Behav. Res. Methods*, vol. 40, no. 1, pp. 8–20, Feb. 2008.

[5] T. A. Judge, C. A. Higgins, C. J. Thoresen, and M. R. Barrick, “The big five personality traits , general mental ability , and career success across the life span,” *Pesonnel Psychol.*, vol. 52, pp. 621–652, 1999.

[6] B. J. Williams, “Investigation of broad and narrow personality traits in relation to major satisfaction for students in engineering, education and psychology majors,” . University of Tennessee Honors Thesis Projects, 2009.

[7] E. Litzler, C. C. Samuelson, and J. A. Lorah, “Breaking it down: Engineering student STEM confidence at the intersection of race/ethnicity and gender,” *Res. High. Educ.*, vol. 55, no. 8, pp. 810–832, 2014.

[8] A. L. Duckworth, C. Peterson, M. D. Matthews, and D. R. Kelly, “Grit: Perseverance and passion for long-term goals,” *J. Pers. Soc. Psychol.*, vol. 92, no. 6, pp. 1087–1101, 2007.

[9] A. L. Duckworth and P. D. Quinn, “Development and validation of the short grit scale (Grit-S),” *J. Personal. Assessment2*, vol. 91, no. 2, pp. 166–174, 2009.

[10] R. Su, L. Tay, and E. Diener, “The development and validation of the comprehensive inventory of thriving (CIT) and the brief inventory of thriving (BIT),” *Appl. Psychol. Heal. well-being*, vol. 6, no. 3, pp. 251–279, 2014.

[11] A. Godwin, “The development of a measure of engineering identity,” in *123rd American Society for Engineering Education Annual Conference & Exposition*, 2016, pp. 1–15.

[12] D. S. Yeager and C. S. Dweck, “Mindsets that promote resilience: When students believe that personal characteristics can be developed,” *Educ. Psychol.*, vol. 47, no. 4, pp. 302–314, Oct. 2012.

[13] A. Kirn and L. Benson, “Engineering students’ perceptions of the future: Exploratory instrument development,” in *Proceedings of 2015 American Society for Engineering Education Conference.*, 2015.

[14] D. F. Shell and J. Husman, “Control, motivation, affect, and strategic self-regulation in the college classroom: A multidimensional phenomenon.,” *J. Educ. Psychol.*, vol. 100, no. 2, pp. 443–459, 2008.

[15] P. Pintrich, D. Smith, T. Garcia, and W. McKeachie, “A manual for the use of the motivated strategies for learning questionnaire (MSLQ),” Ann Arbor, MI, 1991.

[16] M. Hoffman, J. Richmond, J. Morrow, and K. Salomone, “Investigating ‘sense of belonging’ in first-year college students,” *J. Coll. Student Retent.*, vol. 4, no. 3, pp. 227–256, 2003.

[17] P. W. Maloney, M. J. Grawitch, and L. K. Barber, “The multi-factor structure of the Brief Self-Control Scale: Discriminant validity of restraint and impulsivity,” *J. Res. Pers.*, vol. 46, no. 1, pp. 111–115, 2012.

[18] B. M. Gadzella, M. Baloglu, W. G. Masten, and Q. Wang, “Evaluation of the student life stress inventory-revised,” *J. Instr. Psychol.*, vol. 39, no. 2, pp. 82–91, 2012.

[19] T. L. Renshaw and S. J. Bolognino, “The college student subjective wellbeing questionnaire: A brief, multidimensional measure of undergraduate’ s covitality,” *J. Happiness Stud.*, vol. 17, pp. 463–484, 2016.

[20] A. Kirn, A. Godwin, L. Benson, H. Boone, and D. Verdin, “Intersectionality of non-normative identities in the cultures of engineering,” in *Proceedings of 2016 American Society for Engineering Education Conference.*, 2016.

[21] K. W. Brown and R. M. Ryan, “The benefits of being present: Mindfulness and its role in psychological well-being.,” *J. Pers. Soc. Psychol.*, vol. 84, no. 4, pp. 822–848, 2003.

[22] J. MacKillop and E. J. Anderson, “Further psychometric validation of the Mindful Attention Awareness Scale (MAAS),” *J. Psychopathol. Behav. Assess.*, vol. 29, no. 4, pp. 289–293, 2007.

[23] J. Cohen, P. Cohen, S. G. West, and L. S. Aiken, *Applied multiple regression/correlation analysis for the behavioral sciences*, Third. New York, NY: Routledge, 2015.

[24] R Core Team, “R: A language and environment for statistical computing.,” *R Foundation for Statistical Computing, Vienna, Austria.*, 2018. [Online]. Available: https://www.r-project.org/.

[25] W.- Cheng and J. Mau, “Characteristics of US students that pursued a STEM major and factors that predicted their persistence in degree completion,” *Univers. J. Educ. Res.*, vol. 4, no. 6, pp. 1495–1500, 2016.

[26] D. Lafreniere, C. Plouff, W. Mokhtar, and P. Plotkowski, “Extended abstract -Organizational socialization through a three-stage co-op model yields slump- to-jump transition during Engineering mid-years,” in *Mid Years Engineering Experience (MYEE) Conference*, 2016, p. W1A–1–W1A–4.

[27] J. Aronson, C. B. Fried, and C. Good, “Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence,” *J. Exp. Psychol.*, vol. 38, pp. 113–125, 2002.

[28] B. M. Holloway, T. Reed, P. K. Imbrie, and K. Reid, “Research□informed policy change: A retrospective on engineering admissions,” *J. Eng. Educ.*, vol. 103, no. 2, pp. 274–301, 2014.