Predicting STEM and Non-STEM College Major Enrollment from Middle School Interaction with Mathematics Educational Software

Maria Ofelia Z. San Pedro[1], Jaclyn L. Ocumpaugh[1], Ryan S. Baker[1], Neil T. Heffernan[2]
[1]Teachers College Columbia University, 525 W 120th St. New York, NY 10027
[2]Worcester Polytechnic Institute, 100 Institute Rd. Worcester, MA 01609

mzs2106@tc.columbia.edu, ocumpaugh@tc.columbia.edu, baker2@exchange.tc.columbia.edu, nth@wpi.edu

ABSTRACT
The worldwide increase in demand for qualified workers in science, technology, engineering, and mathematics (STEM) fields has resulted in a greater focus on preparing students to enroll in postsecondary STEM programs. The processes that lead students to become interested in and equip them for STEM careers begin years earlier. Previous research indicates that family background, financial resources, and prior family academic achievement can be used to predict whether a student will enroll in a STEM major. In this paper, we consider another class of factors that may be predictive while being more actionable. In this paper, we use predictive analytics, based on previously-validated automated detectors of student learning and engagement, to predict which students will choose a STEM major. With data from 363 college students who used ASSISTments during their regular middle school math classes, we develop a model that can successfully distinguish 66% of the time if a student will choose a STEM major or a non-STEM major when they enter college. In doing so, we offer steps towards providing educators with more actionable information on the STEM trajectories of individual students.

Keywords
STEM, Affect Detection, Knowledge Modeling, Educational Data Mining, Predictive Analytics, Gaming the System

1. INTRODUCTION
Science, technology, engineering, and mathematics (STEM) jobs have played a significant role in driving the modern economy, with growth as high as three times faster than that of non-STEM jobs in the United States over the last decade [13]. Many STEM jobs require a postsecondary degree or other advanced technical training. However, research shows a gap between the number of students who express interest in STEM degree programs and the number who actually enter them, which is driven by inadequate preparation for higher level STEM skills and other aspects of college readiness [21]. This lack of preparation often begins as early as middle school. For instance, the National Mathematics Advisory Panel argues that difficulties with concepts like fractions hinder students from further achievement in mathematics, including algebra [15].

Since the motivation and interest that guides students to enter STEM careers can often be traced to middle school [12], it may be valuable to work on creating better understanding of the factors and processes in middle school students' learning and engagement that connect to eventual decisions to pursue STEM degrees and careers. Studies show that family background, financial resources, and prior family academic achievement have a significant impact to a student's interest and intention to major in STEM [21]. However, current predictive models are generally insufficient to help classroom teachers identify which students are on track, which need further support, and what types of interventions are likely to have the greatest impact [14]. Part of the challenge is one of getting the right data – predictive models have typically relied on course-level data like grades [4] or high-level indicators of general student interest in STEM careers [12, 21], making it difficult to make predictions actionable before a student is already significantly off-track.

Recent work, however, has taken advantage of the increasing deployment of educational software that logs student behavior (often in fine-grained detail), developing automated detectors that assess student learning and engagement [3, 8, 10, 18]. This development creates new opportunities to assess students on a broader range of constructs than previously possible and to predict long-term student outcomes, such as college enrollment several years after using a learning system [20]. This study builds on these models, finding that enrollment in STEM degree programs (among those in college) can be inferred from learning and engagement during middle school mathematics learning. Using previously developed automated detectors of knowledge, affect, and disengaged behavior, we develop a prediction model to distinguish whether or not students who attend college will enroll in a STEM major. By identifying these constructs, we argue, we can better identify which students are most in need of interventions, helping educators to better serve their students.

2. METHODOLOGY
2.1 The ASSISTments System
This study predicts student outcomes from their interactions with the ASSISTments system [19], a free web-based mathematics tutoring system for middle-school mathematics, provided by Worcester Polytechnic Institute (WPI). ASSISTments *assesses* a student's knowledge while *assisting* them in learning, providing teachers with formative assessment of students as they acquire specific knowledge components. Within the system, each mathematics problem maps to one or more cognitive skills. When students answer correctly, they proceed to the next problem. When they answer incorrectly, the system scaffolds instruction by dividing the problem into component parts, stepping students through each before returning them to the original problem (see Figure 1). Once the original problem is correctly answered, the student advances to the next.

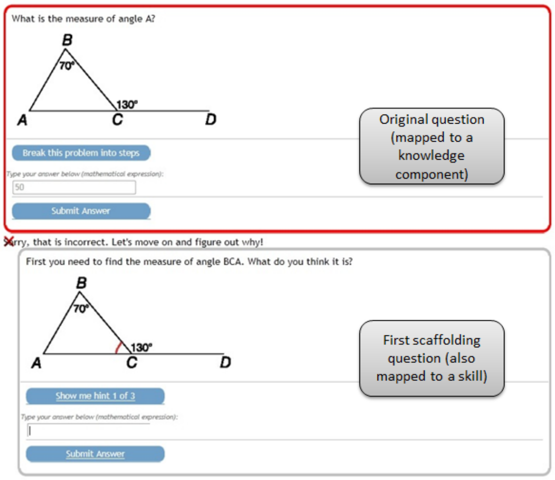

Figure 1. Example of an ASSISTments problem.

2.2 Data

2.2.1 Post-High School Survey Data

For this study, over 2,500 students who had used ASSISTments during their middle school mathematics classes were invited to participate in a survey about their post-high school academic and career achievements. As of this writing, a total of 425 students have responded, for a retention rate of about 20%. Students were drawn from school districts in the Northeastern US who used ASSISTments during the 2004-2005 to 2006-2007 school years (with a few continuing usage for more than one year). Of the 3 districts targeted for this research, 1 was a low-performing district in an urban area, primarily serving English language learners eligible for free or reduced-price lunches; the other 2 districts drew from suburban middle-class populations.

The survey asked students to specify what degree program(s) they were enrolled in, whether they were engaged in full or part-time employment, and what their current employment was. Out of the 425 respondents, 363 were enrolled in college and will be analyzed within this paper. Student answers were coded to reflect their enrollment or participation in a STEM major as defined by the National Science Foundation [16].

2.2.2 ASSISTments Data

Action log files from ASSISTments were obtained for each of the 363 respondents, generating a total of 329,565 actions within the system, across a total of 166,188 original and scaffolding problems. (Actions include answering questions or requesting help.) On average, this resulted in 457 problems per student. Knowledge, affect, and behavior models were applied to this dataset, generating features used for our predictive model of STEM major enrollment.

2.3 Feature Distillation

The features used to predict college major classifications (STEM vs. non-STEM) were generated using a discovery with models approach, leveraging automated detectors of student engagement and learning that were previously developed and validated for ASSISTments. These included existing models of student knowledge, disengaged behaviors (carelessness, gaming the system, and off-task behavior), educationally-relevant affective states (boredom, engaged concentration, confusion, frustration), and other information about student usage (the proportion of correct actions and the total number of actions made by the student, a proxy for overall usage).

Corbett and Anderson's [9] Bayesian Knowledge Tracing (BKT) model, a proven knowledge-estimation model used in a number of ITS systems, was applied to the data for this study by employing a brute-force grid search. BKT infers students' latent knowledge from their performance on previous problems involving the same set of skills. Each time a student attempts a problem or problem step for the first time, BKT recalculates the estimates of that student's knowledge for the skill (or knowledge component) involved in that problem. BKT estimates were calculated at the student's first response to each problem and were applied to each of the student's subsequent attempts on that problem.

To obtain assessments of affect and disengaged behaviors, we leverage existing detectors of student affect and behavior within the ASSISTments system [17, 18]. These included boredom, engaged concentration, confusion, frustration off-task behavior, gaming the system, and carelessness. Data from students who attended urban schools were labeled using affect models optimized for students in urban schools [17, 18], and data from students who attended suburban schools were labeled using affect models optimized for students in suburban schools [17].

Except for carelessness (explained below), the affect and behavior detectors were developed in a two-stage process. First, student affect labels were acquired from BROMP field observations, which records them using HART, an Android app (reported in [18]). Then those labels were synchronized with the log files generated by ASSISTments. This process resulted in automated detectors that can be applied to log files at scale, specifically the data set used in this project (action log files for the 363 students). The detectors were constructed using only log data from student actions within the software occurring concurrently or prior to each BROMP observation, achieving state-of-the-art model goodness [17, 18], and were applied to the data set used in this paper to produce confidence values for each construct for each student actions. Detector confidences were rescaled in order to correct for bias caused by resampling during training [18, 20].

Carelessness is operationalized using contextual slip estimates—the probability that despite knowing the skill to answer an item, a specific incorrect action made by the student for that item is the result of slip or carelessness (see [2]). The probability of carelessness/slip is assessed contextually and is different depending on the context of the student error. As such, the estimate of probability of carelessness/slip is different for each student action. This study uses carelessness models that were previously constructed for ASSISTments [18].

2.4 Modeling STEM Major Enrollment

Within this paper, we develop a logistic regression model predicting STEM major enrollment from combinations of features. Using logistic regression allows for relatively good interpretability of the resultant model, while matching the statistical approach used in much of the work predicting long-term transitions from K-12 education to college [3, 6, 11, 20].

For each of the assessments (learning, affect, and disengaged behaviors), aggregate student-level predictor variables were created by taking the average of the predictor feature values for each student. (In other words, taking the average boredom per student, average confusion per student, etc.) A simple backward elimination feature selection, based on each parameter's statistical significance was used. All predictor variables were standardized using z-scores to increase interpretability of the resulting odds ratios. (Note that this does not impact model goodness or predictive power in any fashion.)

3. RESULTS

First, we looked at our original, non-standardized features and how their values compare between those who were pursuing a STEM major in college and those who were not (Table 1).

Table 1. Feature comparison for STEM Major students (1, n=194) and Non-STEM Major (0, n=169).

	STEM Major	Mean	Std. Dev.	t-value	Cohen's d
Carelessness	0	0.204	0.118	-4.437	0.460
	1	**0.267**	**0.154**	(p < 0.01)	
Student Knowledge	0	0.340	0.196	-4.853	0.508
	1	**0.447**	**0.223**	(p < 0.01)	
Correctness	0	0.418	0.171	-5.184	0.547
	1	**0.508**	**0.161**	(p < 0.01)	
Boredom	0	0.222	0.072	0.286	0.030
	1	0.219	0.078	(p = 0.78)	
Engaged Concentration	0	0.660	0.064	1.500	0.162
	1	0.652	0.044	(p = 0.14)	
Confusion	0	0.085	0.058	0.636	0.067
	1	0.081	0.062	(p = 0.53)	
Frustration	0	0.171	0.078	1.602	0.166
	1	0.155	0.101	(p = 0.11)	
Off-Task	0	0.206	0.086	-0.709	0.076
	1	0.212	0.062	(p = 0.48)	
Gaming	**0**	**0.181**	**0.174**	5.269	0.573
	1	0.100	0.108	(p < 0.01)	
Number of Actions	**0**	**1049.5**	**1569.2**	1.984	0.218
	1	784.53	794.65	(p < 0.05)	

An independent samples t-test (Table1) shows that students in STEM majors had higher mean values for average student knowledge, average carelessness, and average correctness, while students in non-STEM majors had higher mean values for average gaming and average number of actions. Effect sizes for these features were computed using Cohen's d which measures the standardized mean difference of the features between two groups – in this paper, the students pursuing a STEM major and those taking a non-STEM major. As shown in Table 1, gaming the system has the largest effect size (d=0.573), indicating that students who took a non-STEM major had a mean gaming percentage 0.573 standard deviations higher during middle school than students who took a STEM major. It is worth noting that the effect size of gaming for predicting STEM major is substantially larger than the effect size of gaming for predicting whether students attended college or not, where d was 0.293 [43].

These observations align with the individual effects of each feature on the prediction of STEM major enrollment. For example, there is a strong positive relationship between enrolling in a STEM major and average correct answers, indicating that success in mathematics using ASSISTments is associated to higher probability of pursuing a STEM major. The same strong positive relationship is seen between STEM major enrollment and student knowledge estimate as the student learns with ASSISTments. Two non-intuitive results are found in these data. The first concerns the relationship between carelessness and STEM major enrollment. Taken by itself, the more a student becomes careless, the more likely the student is to choose a STEM major, evidence in keeping with past results that careless errors are characteristic of more successful students [7]. The second non-intuitive result concerns the amount of interaction the student has had with the system. Our results show that the number of actions per student is negatively related to majoring in a STEM program, perhaps indicative of struggling students whose actions consist mostly of help requests and scaffolded attempts (which indicate that the student got many problems wrong on the first try). Additionally, the more a student games the system, the less likely that student is to enroll in a STEM major – a result compatible with past evidence that gaming is associated with poorer learning in mathematics [8].

A model for STEM Major enrollment including a combination of data features was developed using Logistic Regression and cross-validated at the student level (6-fold). Our final model (Table 2) achieves a cross-validated A' of 0.663 and a cross-validated Kappa of 0.257. This model is statistically significantly better than the null model, χ^2 (df = 2, N = 363) = 38.010, p < 0.001 and achieved a fit of R^2 (Cox and Snell) = 0.099, R^2 (Nagelkerke) = 0.133, indicating that its predictors explain 9.9-13.3% of the variance of those who attended college. As seen in Table 2, the predictors (student knowledge and gaming) maintained the same directionality as they demonstrated individually (Table 1).

Table 2. Model of STEM major enrollment

Features	**Coefficient**	**Chi-Square**	**p-value**	**Odds Ratio**
Student Knowledge	0.357	8.859	0.003	1.429
Gaming	-0.492	13.792	< 0.001	0.611
Intercept	0.133	1.418	0.234	1.142

4. DISCUSSION AND CONCLUSION

This paper presents a logistic regression model which indicates that a combination of features of student engagement and student success in ASSISTments can distinguish a student who will take a STEM major 66.3% of the time. Success within middle school mathematics (indicated by correct answers and high probability of knowledge in ASSISTments) is positively associated with STEM major enrollment, a finding aligned with studies that conceptualize high performance and developing aptitude during early schooling as a sign of STEM major readiness and predictor of later enrollment in STEM programs [21]. The disengaged behavior of gaming the system during middle school mathematics is found to be negatively associated with pursuing a STEM degree. Previous research has shown that that gaming negatively impacts learning [8], but it is also a particularly strong indicator of disengagement with mathematics, suggesting that students' lack of interest in STEM careers may manifest early. It has been shown that gaming behaviors can be successfully remediated either through alternate opportunities to learn the material that students bypassed or through metacognitive interventions which explain why gaming is ineffective for learning [1, 8]. The relationship between gaming and the choice of college major is relatively large, larger than its relationship to whether a student attends college [20], suggesting that gaming remediation could be an important component of efforts to encourage more students towards STEM degree programs.

Our model also finds that affective states are not particularly strong predictors of whether a student will pursue a STEM major, in contrast to work which found that affective states were predictive of college attendance [20]. A possible explanation is that student affect may be less relevant for college major choice than how students respond to that affect (e.g. a student who just becomes careless when he or she gets bored might be more likely to maintain the STEM track than a student who games the system in response to his or her boredom). It also may be that affect during schooling largely plays a role in determining whether students choose higher education at all; once we analyze only the students who choose higher education (e.g. the current sample), affect plays a much smaller role than domain-specific learning or choices. Negative affective states should still be attended to, as they impact both learning outcomes and college attendance [18, 20]. It may be a valuable area of future work to explore whether the interactions of affective states with other factors can influence these predictions. For example, gaming the system and carelessness may be mediating some of the relationships between affect and college major selection.

One possible use of these findings is to give educators and career counselors a new lens on early indicators of disinterest or disengagement from STEM content and instruction, allowing them to develop counseling strategies that will sustain student interest in pursuing STEM degrees and careers. In doing so, it is important to note that despite considerable current societal emphasis on encouraging students to pursue STEM majors, some students will have other interests and goals. At the same time, the demand for STEM professionals considerably outstrips supply and there is value in a citizenry that has basic STEM literacy [5], regardless of their career choices. If these predictions, based on interactions within a mathematics tutor, can be used to provide targeted help to students that builds on their strengths and strengthens their weaknesses, we stand the chance of both identifying students who are particularly gifted in mathematics and creating greater options for students who struggle for one reason or another. As online learning spreads to other domains of K-12 education, we will be able to provide similar support within other subject domains, supporting all students in reaching their maximum potential.

5. ACKNOWLEDGMENTS

This research was supported by the NSF (#DRL-1031398, #SBE-0836012) and the Bill & Melinda Gates Foundation (#OPP1048577). We also thank Kevin Ramirez for his support in data collection and Irina Shklovski for her advice.

6. REFERENCES

[1] Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Mehranian, H., Fisher, D., Barto, A., Mahadevan, S. and Woolf, B. 2007. Repairing disengagement with non-invasive interventions. In *Proc. AIED*, 195–202.

[2] Baker, R.S.J.d., Corbett, A.T., and Aleven, V. 2008. More Accurate Student Modeling through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing. In *Proceedings of the 9th ITS Conference*, 406-415.

[3] Baker, R.S., Corbett, A.T., and Koedinger, K.R. 2004. Detecting Student Misuse of Intelligent Tutoring Systems. In *Proceedings of the 7th ITS Conference*, 531-540.

[4] Bowers, A.J., Sprott, R., and Taff, S.A. 2013. Do we Know Who Will Drop Out? A Review of the Predictors of Dropping out of High School: Precision, Sensitivity and Specificity. *The High School Journal*. 96(2), 77-100.

[5] Bybee, R. W. (1997). *Achieving scientific literacy: From purposes to practices*(Vol. 88). Portsmouth, NH: Heinemann

[6] Cabrera, A. F. 1994. Logistic regression analysis in higher education: An applied perspective. *Higher Education: Handbook of Theory and Research*, 10, 225-256.

[7] Clements, K. 1982. Careless errors made by sixth-grade children on written mathematical tasks. *Journal for Research in Mathematics Education*, 13, 136–144.

[8] Cocea, M., Hershkovitz, A., and Baker, R.S.J.d. 2009. The Impact of Off-task and Gaming Behaviors on Learning: Immediate or Aggregate? In *Proc. AIED*, 507-514

[9] Corbett, A.T., and Anderson, J.R. 1995. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction* 4, 4, 253-278.

[10] D'Mello, S.K., Craig, S.D., Witherspoon, A. W., McDaniel, B. T., and Graesser, A. C. 2008. Automatic Detection of Learner's Affect from Conversational Cues. *User Modeling and User-Adapted Interaction*, 18 (1-2), 45-80.

[11] Eccles, J. S., Vida, M. N., and Barber, B. 2004. The relation of early adolescents' college plans and both academic ability and task-value beliefs to subsequent college enrollment. *Journal of Early Adolescence*, 24, 63-77.

[12] Hayden, K., Ouyang, Y, Scinski, L., Olszewski, B., and Bielefeldt, T. 2011. Increasing student interest and attitudes in STEM: Professional development and activities to engage and inspire learners. *Contemporary Issues in Technology and Teacher Education*, 11(1), 47-69.

[13] Langdon, D., McKittrick, G., Beede, D., Khan, B., and Doms, M. 2011. STEM: Good jobs now and for the future. *ESA Issue Brief# 03-11*. US Department of Commerce.

[14] Lent, R. W., Lopez Jr, A. M., Lopez, F. G., and Sheu, H. B. 2008. Social cognitive career theory and the prediction of interests and choice goals in the computing disciplines. *Journal of Vocational Behavior*, 73(1), 52-62.

[15] National Mathematics Advisory Panel. 2008. *Foundations for success: The final report of the National Mathematics Advisory Panel*. US Department of Education.

[16] National Science Foundation, National Center for Science and Engineering Statistics. 2013. Science and Engineering Degrees: 1966–2010. *Detailed Statistical Tables NSF* 13-327. Arlington, VA.

[17] Ocumpaugh, J., Baker, R., Gowda, S., Heffernan, N., Heffernan, C. (in press) Population validity for Educational Data Mining models: A case study in affect detection. To appear in the *British Journal of Educational Technology*.

[18] Pardos, Z.A., Baker, R.S.J.d., San Pedro, M.O.C.Z., Gowda, S.M., Gowda, S.M. 2013. Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. In *Proceedings of the 3rd LAK Conference*, 117-124

[19] Razzaq, L. M., Feng, M., Nuzzo-Jones, G., Heffernan, N. T., Koedinger, K. R., Junker, B., ... & Rasmussen, K. P. (2005, May). Blending Assessment and Instructional Assisting. In *AIED* (pp. 555-562).

[20] San Pedro, M.O.Z., Baker, R.S.J.d., Bowers, A.J., Heffernan, N.T. 2013. Predicting College Enrollment from Student Interaction with an Intelligent Tutoring System in Middle School. In *Proceedings of the 6th EDM Conference*, 177-184.

[21] Wang, X. 2013. Why Students Choose STEM Majors Motivation, High School Learning, and Postsecondary Context of Support. *American Educational Research Journal*, 50(5), 1081-1121.