Assessing the Roles of Student Engagement and Academic Emotions within Middle School Computer-Based Learning in College-Going Pathways

Maria Ofelia Z. San Pedro
Teachers College Columbia University
525 W 120th St.
New York, NY 10027
mzs2106@tc.columbia.edu

ABSTRACT

This dissertation research focuses on assessing student behavior, academic emotions, and knowledge from a middle school online learning environment, and analyzing their potential effects on decisions about going to college. Using students' longitudinal data ranging from their middle school, to high school, to postsecondary years, I leverage quantitative methodologies to investigate antecedents to college-going outcomes that can occur as early as middle school. The research first looks at whether assessments of learning, emotions and engagement from middle school computer-based curriculum are predictive at all of college-going outcomes years later. I then investigate how these middle school factors can be associated with college-going interests formed in high school, using the same assessments during middle school, together with self-report measures of interests in college when they were in high school. My dissertation then culminates in developing an overall model that examines how student interests in high school can possibly mediate between the educational experiences students have during middle school technology-enhanced learning and their eventual college-going choices. This gives a richer picture of the cognitive and motivational mechanisms that students experience throughout varied phases in their years in school.

Keywords

College Choices, Academic Emotions, Behavior, Knowledge, Social Cognitive Career Theory, Interests

1. Introduction

College enrollment and completion are key steps towards career success for many learners. However, well before this point, many students effectively drop out of the pipeline towards college quite early. According to Social Cognitive Career Theory (SCCT) [10], academic and career choices are shaped throughout middle school and high school by environment supports and barriers, where higher levels of interest emerge within contexts in which the individual has higher self-efficacy and outcome expectations, and these interests lead to the development of intentions or goals for further exposure and engagement with the activity [10]. Traditional studies also show that family background, financial resources, and prior family academic achievement have significant impacts on where students find themselves after high school. All of these factors, however, are fairly strong displays of disengagement. By the time these indicators are commonplace, students may be in such a precarious situation that many interventions may fail. In general, current models about successful access to postsecondary education may be insufficient to help educators identify which students are on track and which need further support [11]. Fine-grained assessments of student behaviors and academic emotions (emotions that students experience during learning and classroom instruction) have been found to influence learning outcomes [12, 13]. Hence, there is an argument to be made that engagement and academic emotions in middle school play an essential early role in the processes described in SCCT. In SCCT, students' initial vocational interests are modified by their self-efficacy, attitudes, and goals towards career development (i.e. college enrollment, career interest), which are themselves influenced by the student's learning and engagement when encountering the increasingly challenging content in middle school [1, 12] – as poor learning reduces self-efficacy whereas successful learning increases self-efficacy [cf. 2]. As such, student academic emotions, learning, and engagement during middle school may be indicative of their developing interests in career domains which may in turn influence their choice to attend college [6, 9].

For the reasons aforementioned, my research attempts to answer Bowers' [5] call to identify much early, less acute signals of disengagement, the sort that occur when students' engagement is still malleable enough for interventions to succeed. Specifically, I investigate antecedents to college attendance that occur during middle school, using assessments of engagement and disengagement to better understand how these factors interact so that I can develop possible paths to re-engagement before students develop more serious academic problems. The models I create and the analyses I conduct involve the context of an online learning environment, and hence, this work provides both a new perspective on the efficacy of the system and an opportunity to explore how the system and its data can be used to predict long-term educational outcomes – in the case of my dissertation research, intervention and support in keeping students on track towards the pathway to college.

2. Data and Related Methodologies

My dissertation leverages data acquired from both traditional research methods as well as methodologies from machine learning and student modeling in assessing the constructs I analyze in my data, which I then use in developing the outcome models I propose. For middle school measures, I use the ASSISTment system (ASSISTments) as my source for middle school interaction data, and assessed measures of student knowledge, academic emotions, and behavior by using individual models developed to infer them. ASSISTments is a free web-based tutoring system for middle school mathematics that assesses a student's knowledge while assisting them in learning, providing teachers with detailed reports on the skills each student knows [14]. Interaction data from the ASSISTment system were obtained for a population of middle school students who used the system at various school years, from 2004-2005 to 2008-2009. These students are drawn from urban and suburban districts who used the ASSISTment system systematically during the year. I assessed

a range of constructs from interaction data in ASSISTments, which include student knowledge estimates, student academic emotions (boredom, engaged concentration, confusion), student disengaged behaviors (off-task, gaming the system, carelessness), and other information of student usage. These form the features in our final model of college-going outcomes. Aside from educational software data, I also use survey data from the same students who used the system in middle school, consisting of information about the their attitude about the subject (mathematics) and about the system itself. These survey data were acquired around the same time they used the software in middle school.

For my high school measures of interest, students who used the system during their middle school years and who are now in high school, were administered with two surveys: the first is a short questionnaire that asked the highest level of math and science courses that the student completed in high school and asks the student what his/her educational and career plans are upon graduation. The second survey is the an CAPA survey, designed by Fred Borgen and Nancy Betz [4]. It is an online survey with Likert scale inputs from students that gauges their interest and confidence on certain domains and skills, and then assesses their overall self-efficacy and vocational interests using existing instruments.

A subset of our student sample who were expected to be in postsecondary stage of education by the time of data collection were identified for their postsecondary education status. For their college enrollment information, records were requested from the National Student Clearinghouse, with information such as whether they were enrolled in a college or not, the name of the university, date of enrollment, and college major enlisted if available. We supplemented this data with college selectivity classification of the said postsecondary institutions, taken from the Barron's College Selectivity Rating which classifies colleges into ten categories [7, 16], from most selective or 'Most Competitive' to 'Special' which consist of specialty institutions such as schools of music, culinary schools, art schools, etc. Another source of data includes survey data about post-high school academic and career achievements that was administered to this subset of students.

3. Preliminary Work

In developing an overall integrated model, I initially tested the predictive power of the middle school factors on separate postsecondary outcomes. First, I applied fine-grained models of student knowledge, student academic emotions (boredom, engaged concentration, confusion, frustration) and behavior on middle school interaction data to understand how student learning and engagement during this phase of learning can predict college enrollment. A logistic regression model was developed and can distinguish a student who will enroll in college (68.6% of the time, an above average performance for models created from "discovery with models"). In particular, boredom, confusion, and slip/carelessness are significant predictors of college enrollment both by themselves and contribute to the overall model of college enrollment. The relationships seen between boredom and college enrollment, and gaming the system and college enrollment indicate that relatively weak indicators of disengagement are associated with lower probability of college enrollment. Success within middle school mathematics is positively associated with college enrollment, a finding that aligns with studies that conceptualize high performance as a sign of college readiness [15] and models that suggest that developing aptitude predicts college attendance [8].

Next, I also modeled whether students will attend a selective college, combining data from students who used the ASSISTment system with data on college enrollment, and ratings from Barron's on college selectivity. These were used to model another logistic regression model that could distinguish between a student who will attend a selective college and a student who will not attend a selective college 76% of the time when applied to data from new students. This model indicated that the following factors are associated with lower chance of attending a selective college: gaming the system, boredom, confusion, frustration, less engaged concentration, lower knowledge, and carelessness.

I finally looked at college major classification based on middle school student learning and engagement, specifically whether the major belonged to a STEM (Science, Technology, Engineering, Mathematics) or Non-STEM category. The logistic regression model developed could distinguish between a student who took a STEM college major and a student who took an non-STEM college major 66% of the time when applied to data from new students This model indicated that the following factors are associated with lower chance of enrolling in a STEM college major: gaming the system, lower knowledge, and carelessness.

4. Proposed Work

The initial individual models above support existing theories about indicators of successful entry to postsecondary education (academic achievement, grades). It sheds light on behavioral factors a student may experience in classrooms – which are more frequently and in many ways more actionable than the behaviors which result in disciplinary referrals – and how they can be predictive and be associated with long-term student outcomes.

With middle school assessments, I investigate at how student learning, academic emotions, and behavior as early as middle school may contribute as causal factors to a particular postsecondary decision (a in Figure 1 below) – an individual choice that is composed of answering the following questions: 1) Does the student decide to attend college?; 2) Does the student attend a selective college?; 3) What type of major does the student enroll in? I employ multivariate analysis on this part of my research work, for a richer and more realistic view of our postsecondary outcome, which is more than just one dependent variable of interest. Also this type of analysis allows us for causality to be deduced, as well as the inherent or underlying structure that can describe the data in a simpler fashion – in terms of latent variables. I also investigate interaction of features and how it affects our multivariate model via logistic regression, factor analysis and other appropriate statistical and machine learning algorithms that can be employed in our data to further understand the research problem.

In this phase of my dissertation research, I am starting to test the hypothesis of the possible existence of a mediating or indirect effect of high school college (and career) interests in predicting the multivariate postsecondary outcome based on middle school factors. I will establish this by looking at the causal influence of middle school factors to high school data (b in Figure 1 below). By integrating student data of their previous middle school interaction data, interests during their high school years, up to their postsecondary information, I will look at the possible causality of middle school factors to high school factors, as well as causality of high school factors to their postsecondary information. Like in previous analysis, I employ appropriate statistical and machine learning algorithms in trying to establish the indirect effect of high school factors (for our overall mediated

model later on). First, I look at how the middle school measures of student learning, engagement and academic emotions are predictive of the high school questionnaire responses, through multinomial logistic or decision tree algorithms. Then, I explore the association between the high school questionnaire responses with the multivariate postsecondary outcomes using structural equation modeling (factor analysis, regression, or path analysis).

Finally, by integrating emergent relationships and causal effects of middle school and high school factors on postsecondary outcomes conducted in the previous analyses, I will develop a multivariate predictive mediated model (c in Figure 1 below). Using student data that have complete information from middle school, to high school, to postsecondary years, I conduct causal modeling by fitting a mediational pathway model and evaluate how each of the variables influence one another over time [3]. In particular, using structural equation modeling (SEM), I develop a pathway starting from the middle school factors to the postsecondary outcomes, with high school factors as intervening or mediating factors. With significant zero-order correlations between the constructs (middle school factors, high school factors, postsecondary outcomes) established from the previous analyses, I employ a multiple regression analysis predicting postsecondary outcomes from both middle school and high school factors. It is expected that any partial effect (indirect effect) of high school factors (controlling for middle school factors) to be significant, decreasing the direct effect of middle school factors on postsecondary outcomes. Other SEM variants, such as factor analysis and path analysis are expected to be used as well for this analysis phase, to test the mediation model. This causal modeling has been used in educational research modeling motivational phenomena over time [3].

(c)

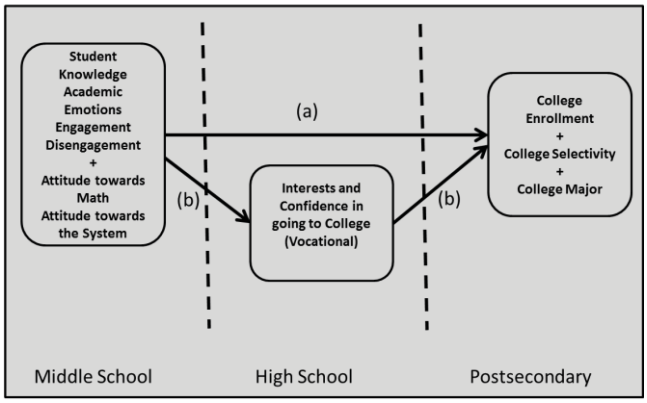

Figure 1. Modeling Postsecondary Outcomes from Middle School and High School factors: (a) Middle school factors predicting postsecondary outcomes; (b) Middle school factors predicting high school factors, High school factors predicting postsecondary outcomes; (c) Overall mediation model.

5. References

[1] Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., and Koedinger, K. 2008. Why students engage in "gaming the system" behavior in interactive learning environments. *Journal of Interactive Learning Research*, 19, 2, 185-224.

[2] Bandura, A. 1997. *Self efficacy: The exercise of control*. New York,. NY: W. H. Freeman & Company.

[3] Blackwell, L., Trzesniewski, K., and Dweck, C. 2007. Implicit Theories of Intelligence Predict Achievement Across an Adolescent Transition: A Longitudinal Study and an Intervention. *Child Development*, 78, 1, 246-263.

[4] Borgen, F. and Betz, N. 2008. Career self-efficacy and personality: Linking the Career Confidence Inventory and the Healthy Personality Inventory. *Journal of Career Assessment*, 16, 22-43.

[5] Bowers, A. J. 2010. Grades and graduation: A longitudinal risk perspective to identify student dropouts. *The Journal of Educational Research*, 103, 3, 191-207.

[6] Chen, X. 2009. Students Who Study Science, Technology, Engineering, and Mathematics (STEM) in Postsecondary Education. Stats in Brief. *NCES 2009-161*. National Center for Education Statistics.

[7] College Division of Barron's Education Series (Ed.). 2012. *Barron's profiles of American colleges (30th ed.)*. Hauppauge, NY: Barron's Educational Series, Inc.

[8] 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.

[9] Griffith, A. L. 2010. Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review*, 29(6), 911-922.

[10] Lent, R. W., Brown, S. D., and Hackett, G. 1994. Toward a unifying social cognitive theory of career and academic interest, choice and performance. *Journal of Vocational Behavior*, 45, 1, 79–122.

[11] 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.

[12] Mcquiggan, S. W., Mott, B. W., and Lester, J. C. 2008. Modeling self-efficacy in intelligent tutoring systems: An inductive approach. *User modeling and user-adapted interaction*, 18, 81-123.

[13] Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., and Perry, R. P. 2010. Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, 102, 3, 531.

[14] Razzaq, L., Feng, M., Nuzzo-Jones, G., Heffernan, N. T., Koedinger, K. R., Junker, B., and Rasmussen, K. P. 2005. The Assistment project: Blending assessment and assisting. In *Proceedings of the 12th Annual Conference on Artificial Intelligence in Education*, 555-562.

[15] Roderick, M., Nagaoka, J., and Coca, V. 2009. College readiness for all: The challenge for urban high schools. *The Future of Children*, 19, 1, 185-210.

[16] Schmidt, W., Burroughs, N., Cogan, L., and Houang, R. 2011. Are College Rankings an Indicator of Quality Education? In *Forum on Public Policy Online*, 2011, 3. Oxford Round Table.