**Student characteristics and online retention:**
**Preliminary investigation of factors relevant to mathematics course outcomes**

Claire Wladis
Borough of Manhattan Community College and the Graduate Center at the
City University of New York

Alyse C. Hachey
Borough of Manhattan Community College at the City University of New York

Katherine Conway
Borough of Manhattan Community College at the City University of New York

*There is evidence that students drop out at higher rates from online than face-to-face courses, yet it is not well understood which students are particularly at risk online. In particular, online mathematics (and other STEM) courses have not been well-studied in the context of larger-scale analyses of online dropout. This study surveyed online and face-to-face students from a large U.S. university system. Results suggest that for online courses generally, student parents and native-born may be particularly vulnerable to poor online-versus-face-to-face course outcomes. The next stage of this research will be to analyze the factors that are relevant to online versus face-to-face retention in mathematics (and other STEM) courses specifically.*

*Key words:* online learning; retention; student characteristics

As more and more courses move to online formats, higher education is undergoing a virtual transformation. By 2013, over 40 million college students took online classes worldwide; by 2017, that number is expected to reach over 120 million post-secondary students globally (Atkins, 2013). On the positive side, online courses may provide increased access to college, removing impediments to college progression by providing the flexibility that "non-traditional" students need. However, because they often have higher attrition (the reasons for which have yet to be determined), online courses may also be detrimental to degree completion (Jaggars, 2011).

Many questions remain about factors that impact course outcomes in online versus face-to-face courses. Further, the factors that impact course outcomes in the online versus face-to-face medium may be different for mathematics courses than for courses in other subjects, yet almost no larger-scale studies have focused on online mathematics courses specifically. In order for policies and advisement to be grounded in research evidence, mathematics education research must address the rapid growth in online learning and the need to focused research on factors impacting outcomes.

## Research questions

This initial exploratory study seeks to determine the relationship between student characteristics and online course-taking in order to inform later research as to which factors may impact mathematics course retention and grades specifically:

1. Which student characteristics make a student more likely to enroll in online than face-to-face courses?

2. Which student characteristics exacerbate or mitigate differences in rates of online versus face-to-face course retention and successful course completion?
3. Are there specific groups (e.g. women, racial/ethnic minorities, "non-traditional students") that are particularly successful or particularly vulnerable when taking courses online?
4. How do these patterns differ when comparing mathematics courses to other STEM or non-STEM courses?

## Theoretical framework and prior research

Doubling from just under a decade ago, thirty-two percent of U.S. college students enrolled in online courses in 2011-2012 (U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics (NCES), 2008). Further, since 2010, online enrollment has increased 29% (Allen and Seaman 2010; 2013; CCRC 2013). Numerous studies, including a 200 study meta-analysis, found no significant difference in learning outcomes in online versus face-to-face courses (e.g.(Bernard et al., 2004; Bowen, Chingos, Lack, & Nygren, 2012). Despite these findings, online course dropout in the U.S. ranges from 20-40%, and online attrition rates are reported as 7-20 percentage points higher than those for face-to-face courses (e.g. Hachey, Wladis & Conway, 2013; (Nora & Snyder, 2009; Patterson & McFadden, 2009).

Further, recent research suggests that the gap in attrition between the same courses offered online versus face-to-face can be larger for STEM than for non-STEM courses (Wladis, Hachey, & Conway, 2012). There is also some research that suggests that this gap may be larger for mathematics than for English gatekeeper courses, although differences in the gap between subjects was not tested for statistical significance (Xu & Jaggars, 2011). This may mean that there are factors in the online environment which impact mathematics and other STEM courses differently or more strongly than courses in other subjects. However, previous findings on student characteristics cannot necessarily be generalized to mathematics and other STEM courses specifically. In addition, there is currently little rigorous research on factors affecting retention in online STEM courses specifically. Given the rapid growth in online courses and the already high rates of dropout in many mathematics and STEM courses, it is essential to identify which students are at higher risk in online mathematics (and other STEM) courses, in order to target appropriate support services.

Previous research has found that online learners are more likely to be female, older, married, active military or to have other responsibilities (such as full-time work and/or children) (Shea & Bidjerano, 2014; C. Wladis, Hachey, & Conway, 2015). Additional studies have also found that online students tend to have higher G.P.A.'s, to be white, native English speakers, and to have applied for or received financial aid (Conway, Wladis, & Hachey, n.d.; Jaggers & Xu, 2010; Xu & Jaggars, 2011). Further, online student are more likely to have other "non-traditional" characteristics (e.g. delayed college enrollment; no high school diploma; part-time enrollment; financially independent) (e.g. Shea & Bidjerano, 2014; C. Wladis et al., 2015), and to be first-generation college students (Athabasca University, 2006). And non-traditional characteristics have been shown to be more significant as predictors of online enrollment for STEM than for non-STEM students (Wladis, Hachey, & Conway, 2015).

However, research on demographic variables is conflicting (Jones, 2010) and it remains unclear how different characteristics interact with each other to affect retention in online courses. For instance, Bernard, Brauer, Abrami and Surkes (2004) found that self-direction and beliefs were significant positive predictors of online course grade, however, the evidence showed that G.P.A. was a stronger predictor of online course outcomes. Waschull (2005) reports that self-discipline/motivation was significantly correlated with online course grades, but these same factors may predict success in both online and face-to-face classes. Aragon and Johnson (2008) found that online completers were more likely to be female, enrolled in more classes, and had a higher G.P.A., but unlike Abrami and Surkes (2004), they found no significant difference in academic readiness or self-directed learning.

In a similar vein, other investigations of student characteristics have also been inconclusive. Several studies investigating gender found no differences, whereas others report that females outperform males in terms of outcomes (for a review, see (Xu & Jaggars, 2013)). Angiello (2002) and Xu and Jaggars (2013) report differences in outcomes based on ethnicity while Welsh (2007), Aragon and Johnson (2008) and Wladis, Conway and Hachey (2015) found that ethnicity did not have an impact on online course outcomes. G.P.A is cited as a significant factor impacting online course outcomes in some studies, (e.g. Xu & Jaggars, 2013), but was found to be non-significant in others (e.g. Hachey, Wladis, & Conway, 2012). In one study on STEM courses specifically, older students did significantly better in online STEM courses, and women did significantly worse (although still no worse than men) online, than would be expected based on their outcomes in comparable face-to-face STEM courses, but there was no significant interaction between the online medium and ethnicity (C. Wladis et al., 2015).

Studies focused on online mathematics courses specifically have tended to compare student outcomes across mediums, without attempting to assess which factors predict or contribute to those outcomes (see e.g. Ashby, Sadera, & McNary, 2011; Bowen & Lack, 2012), or they have tended to explore factors that predict successful outcomes in online mathematics classes, without comparing them to face-to-face courses, so that it is impossible to determine whether the factors studied are relevant to learning mathematics online or just to learning mathematics more generally (see e.g. Kim, Park, & Cozart, 2014).

To accurately assess whether a factor puts a student at greater risk in the online environment, it is critical to analyze the *interaction* between that factor and course medium, while simultaneously controlling for self-selection into online courses. This is the only way in which it is possible to determine the extent to which particular factors are important in the online medium specifically, and not simply predictors of academic outcomes more generally. This study addresses an important gap in the literature by doing just that. It is an initial step in determining which factors may need to be explored as impacting outcomes in online mathematics courses specifically.

## Methodology

### Data source and sample

This study uses a sample of 9,663 students with 37,442 course records from the 18 two- and four-year colleges in the City University of New York (CUNY) system in the U.S. Students were selected if they were enrolled in a course in the sample frame, which consisted of all online and comparable face-to-face courses offered during the 2014 fall semester at one of the CUNY colleges. Of the courses that determined the sample frame, roughly 25% were STEM courses and roughly 10% were mathematics courses. At the end of the semester, students in the sample were invited to participate in an online survey.

### Measures

Two measures of student outcomes were utilized: *course retention*, defined as whether a student dropped a course (officially or unofficially); and *successful course completion*, defined as whether the student successfully completed a course with a C- or higher (chosen because it is the typical standard to receive major or transfer credit).

The main independent variable (IV), course medium, was dichotomized to face-to-face or fully online, based on Sloan Consortium definitions (Allen & Seaman, 2010); fully online courses have 80% or more of the course content online, and face-to-face courses have 33% or less of the content online. Previous studies contend that students who take hybrid courses (33-80% online content) are similar to students who take face-to-face courses and further, that their outcomes are similar (Xu & Jaggars, 2011).

The other IVs investigated were chosen because there is evidence that they may: 1) predict online course enrollment; 2) be related to course outcomes more generally 3) be related to face-to-face mathematics and STEM course outcomes specifically; or 4) be significant predictors of outcomes in the online medium. Covariates included in the study are: whether the student had a child (and age of youngest child); gender; race/ethnicity; age; work hours; income; parental education; developmental mathematics, English, and ESL course placement; marital/cohabitation status; immigration generational status; native speaker status; college level (two-year, four-year, or graduate); G.P.A; and number of credits/classes taken that semester. During preliminary analyses, different non-linear versions of variables were explored (e.g. converting credits to part-time/full-time status, squaring age), but these did not produce significantly different results. Also included in the study survey are scales measuring: motivation to complete the course; course enjoyment/engagement; academic integration (i.e. interaction with faculty/students outside class); self-directed learning skills; time management skills; preference for autonomy; and grit (i.e. perseverance and passion for long-term goals). As much as possible, these scales were based on previous instruments that had already been tested for reliability and validity (Duckworth, Peterson, Matthews, & Kelly, 2007; Macan, Shahani, Dipboye, & Phillips, 1990; Pintrich & de Groot, 1990; U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2009; Vallerand et al., 1992). However, they were shortened and modified for use in this study. Confirmatory factor analysis using structural equation modeling (SEM) was used on the full dataset to model items for each scale as predictors of a single latent construct. Error covariance terms were added between some individual items based on theory, prior to estimation. Some items from the motivation and grit scales were eliminated because of poor performance during SEM. For the final scales, average variance extracted (AVE) was 0.50 or greater, indicating convergent validity, and composite reliability (CR) ranged from 0.77 to 0.89, indicating good reliability (Hair, Anderson, Tatham, & Black, 1998); SRMR ranged from 0.000 to 0.059, supporting the operationalization of each scale as a single factor structure (Hu & Bentler, 1999).

**Analytical Approaches**

Courses for which valid grades did not exist (e.g. not submitted by instructor, course was audited) were dropped. Multivariate multiple imputation by chained equations was used to impute values for survey questions with missing responses, using all IVs chosen for subsequent analyses. Binomial, ordered, or multinomial logit models, or predictive mean matching on three nearest neighbors was used for imputation depending on variable type. A median of 2.6% of data were missing in each imputed variable in the dataset. After preliminary tests for stability of model estimates, 35 imputations were used.

Propensity scores, indicating the probability of online enrollment, were generated using logistic regression and included all of the IVs used in the subsequent analyses. The scores were averaged across imputed datasets. Because this approach yielded the best balance on covariates based on the standardized bias for each imputed variable averaged across imputations, matched datasets were generated using single nearest-neighbor matching with replacement. The median standardized bias across variables was 2.6%, showing a good balance on all covariates based on Rubin's (2001) rule of thumb. Distribution of propensity scores was evaluated both before and after matching, and there was significant overlap in the region of common support.

The imputed dataset was used to run multilevel mixed-effects logistic regression models with course as the first-level and student as the second-level factors, in order to control for unobserved heterogeneity between students. The KHB decomposition method (Kohler, Karlson, & Holm, 2011) was used to calculate direct and indirect effects, in order to explore the relationship between online course outcomes, student characteristics, and subsequent college persistence. Standard errors during KHB decomposition were computed using clustering by course, to account for the multi-level data structure.

## Results

Initial models were run on the whole dataset, including mathematics and non-mathematics courses, in order to look for baseline patterns. Both unmatched and matched datasets were analyzed. For the full dataset consisting of non-STEM, STEM non-mathematics, and mathematics courses, the most consistent predictor of both course retention and successful course completion was being foreign-born. Native-born students were at greater relative risk online compared to foreign-born students, and this was particularly true for native-born students for whom both parents were also native-born. Native-born students with one or no native-born parents were also at increased relative risk online, but the difference was less pronounced. Having a child under six years old was associated with higher risk of unsuccessful course completion or dropout online. No other factors tested were consistently significantly correlated with differential online-versus-face-to-face course retention or successful course completion across different models of the dataset.

The next step in this research project is to repeat these analyses with non-STEM courses, STEM courses, and mathematics courses specifically, to see if different patterns emerge for each group. If there are differences in the patterns observed between groups, then these differences will be tested for significance in an attempt to determine the extent to which different factors are relevant for online STEM and mathematics courses specifically. Further surveys and interviews of online mathematics and STEM students are also being conducted to explore other factors that may be relevant to online mathematics course-taking. Another round of data has recently been collected that explores the following additional constructions: computer and internet self-efficacy (Eastin & La Rose, 2000; Torkzadeh, Chang, & Demirhan, 2006); ethnic/gender identity and stigma consciousness (Picho & Brown, 2011); mathematics (and other STEM subject) self-efficacy and domain identification (May, 2009; Picho & Brown, 2011); sense of belonging (Osterman, 2000); and achievement orientation, e.g. fixed/growth mindset (Dweck, 2006). Interviews are also being conducted with online students; roughly 45 students have been interviewed so far.

## Implications

The results of this study will have strong practical applications. If specific factors can be identified that make students particularly at risk of dropout or failure when they take an online mathematics or STEM course, then these students could be targeted for additional interventions (e.g. tutoring, advising, technical assistance) when they enroll in an online mathematics or STEM course. Future research could test the efficacy of various interventions in improving course outcomes for these at-risk groups.

## Questions

There are several questions that we see as important as we move this research forward. Specifically:

- Are there particular factors that might be relevant to outcomes in online versus face-to-face STEM and mathematics courses specifically that we have not yet considered?
- We plan to run parallel analyses on non-STEM courses; STEM courses; and mathematics courses. But are there other analyses relevant to mathematics courses specifically that may not be parallel to analyses that would make sense in the context of non-mathematics courses?
- Is there anything else that we have overlooked in our choice of variables and analytical approaches, or in our overall study design, that has not yet been addressed?

## References


Allen, I. E., & Seaman, J. (2010). *Class differences: Online education in the United States, 2010.* (No. ED529952).Sloan Consortium. Retrieved from http://sloanconsortium.org/publications/survey/class_differences

Angiello, R. S. (2002). *Enrollment and success of Hispanic students in online courses.* (No. ED 469 358). Washington, D.C.: U.S. Department of Education, Office of Educational Research and Improvement Educational Resources Information Center (ERIC).

Aragon, S. R., & Johnson, E. S. (2008). Factors influencing completion and noncompletion of community college online courses. *American Journal of Distance Education, 22*(3), 146-158. doi:10.1080/08923640802239962

Ashby, J., Sadera, W. A., & McNary, S. W. (2011). Comparing student success between developmental math courses offered online, blended, and face-to-face. *Journal of Interactive Online Learning, 10*(3), 128-140. Retrieved from http://www.ncolr.org/jiol/issues/pdf/10.3.2.pdf

Atkins, S. (2013, ). Ambient Insight Whitepaper: The 2012 boom in learning technology investment. *USnews.Com*

Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovsk, E., Wade, A., Wozney, L., . . . Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research, 74*(3), 379-439. doi:10.3102/00346543074003379

Bernard, R. M., Brauer, A., Abrami, P. C., & Surkes, M. (2004). The development of a questionnaire for predicting online learning achievement. *Distance Education, 25*(1), 31-47. Retrieved from http://eric.ed.gov/?id=EJ680546

Bowen, W. G., & Lack, K. A. (2012). *Current status of research on online learning in postsecondary education.* ().Ithaka S+R. Retrieved from http://continuingstudies.wisc.edu/innovation/ithaka-sr-online-learning.pdf

Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2012). Interactive learning at public universities: Evidence from randomized trials. *Journal of Policy Analysis and Management, 33*(1), 94-111. doi:10.1002/pam.21728

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087-1101.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York: Random House.

Eastin, M., S, & La Rose, R. (2000). Internet self-efficacy and the psychology of the digital divide. *Journal of Computer-Mediated Communication, 6*(1) doi:10.1111/j.1083-6101.2000.tb00110.x

Hachey, A. C., Wladis, C. W., & Conway, K. M. (2012). Is the second time the charm? Investigating trends in online re-enrollment, retention and success. *Journal of Educators Online, 9*(1) Retrieved from http://www.thejeo.com/Archives/Volume9Number1/V9N1.htm#VOLUME_9,_NUMBER_1,_JANUARY_2012

Hair, J. F. J., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate Data Analysis* (5th ed.). Upper Saddle River, New Jersey: Prentice Hall.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55.

Jones, E. H. (2010). *Exploring common characteristics among community college students: Comparing online and traditional student success* (Ph.D.).

Kim, C., Park, S. W., & Cozart, J. (2014). Affective and motivational factors of learning in online mathematics courses. *British Journal of Educational Technology, 45*(1), 171-185. doi:10.1111/j.1467-8535.2012.01382.x

Kohler, U., Karlson, K. B., & Holm, A. (2011). Comparing Coefficients of Nested Nonlinear Probability Models. *Stata Journal, 11*(3), 420-38.

Macan, T. H., Shahani, C., Dipboye, R. L., & Phillips, A. P. (1990). College students' time management: Correlations with academic performance and stress. . *Journal of Educational Psychology, 82*(4), 760-768.

May, D. (2009). *Mathematics self-efficacy and anxiety questionnaire.* (Unpublished Ph.D.). University of Georgia, Athens, Georgia.

Nora, A., & Snyder, B. P. (2009). Technology and higher education: The impact of e-learning approaches on student academic achievement, perceptions and persistence. *Journal of College Student Retention: Research, Theory & Practice, 10*(1), 3-19. doi:10.2190/CS.10.1.b

Osterman, K. F. (2000). Students' Need for Belonging in the School Community. *Review of Educational Research, 70*(3), 323-67.

Patterson, B., & McFadden, C. (2009). Attrition in online and campus degree programs. *Online Journal of Distance Learning Administration, 12*(2) Retrieved from http://www.westga.edu/~distance/ojdla/summer122/patterson112.html;

Picho, K., & Brown, S. W. (2011). Can stereotype threat be measured? A validation of the social identities and attitudes scale (SIAS). *Journal of Advanced Academics, 22*, 374-411. doi:10.1177/1932202X1102200302

Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33-40. doi:10.1037/0022-0663.82.1.33

Shea, P., & Bidjerano, T. (2014). *Does online learning impede degree completion? A national study of community college students.. Computers & Education, 75*, 103-111.

Torkzadeh, G., Chang, J. C., & Demirhan, D. (2006). A contingency model of computer and internet self-efficacy. *Information & Management, 43*(4), 541-550. doi:10.1016/j.im.2006.02.001

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2009). *Beginning postsecondary students longitudinal study.* (). Washington, D.C.: U.S. Department of Education. Retrieved from http://nces.ed.gov/surveys/bps/

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics (NCES). (2008). *National postsecondary student aid study (NPSAS)*

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Valliéres, E. F. (1992). The academic motivation scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement, 52*, 1003-1017.

Waschull, S. B. (2005). Predicting success in online psychology courses: Self-discipline and motivation. *Teaching of Psychology, 32*(3), 190-192. doi:10.1207/s15328023top3203_11

Welsh, J. B. (2007). *Identifying factors that predict student success in a community college online distance learning course.*(Ed.D.). . (AAT 3300982).

Wladis, C. W., Hachey, A. C., & Conway, K. M. (2012). An analysis of the effect of the online environment on STEM student success. Paper presented at the *Proceedings of the 15th Annual Conference on Research in Undergraduate Mathematics Education, , 2*

Wladis, C., Conway, K. M., & Hachey, A. C. (2015). The Online STEM Classroom—Who Succeeds? An Exploration of the Impact of Ethnicity, Gender, and Non-traditional Student Characteristics in the Community College Context. *Community College Review,* doi:10.1177/0091552115571729

Wladis, C., Hachey, A. C., & Conway, K. M. (2015). The Representation of Minority, Female, and Non-Traditional STEM Majors in the Online Environment at Community Colleges: A Nationally Representative Study. *Community College Review, 43*(1), 89-114. doi:10.1177/0091552114555904

Xu, D., & Jaggars, S. S. (2011). The effectiveness of distance education across Virginia's community colleges: Evidence from introductory college-level math and English courses. *Educational Evaluation and Policy Analysis, 33*(3), 360-377. doi:10.3102/0162373711413814

Xu, D., & Jaggars, S. S. (2013). *Adaptability to online learning: Differences across types of students and academic subject areas.* (No. CCRC Working Paper No. 54).Community College Research Center, Columbia University. Retrieved from http://ccrc.tc.columbia.edu/media/k2/attachments/adaptability-to-online-learning.pdf