Course Taking Patterns and Pathways Through STEM: A Case Study Using Individual-Level Institutional Data for Program Assessment

Mary M. Walczak Kathryn Ziegler-Graham Emily Berry Bosak Katherine Henry
Kiegan Rice Marc Richards
St. Olaf College

Abstract

Motivated by assessment of externally awarded grants and internal programmatic practices, we used transcript and demographic data for students entering a highly selective residential liberal arts college between 2002-2015 to answer institutional-level questions about course taking patterns and pathways relevant to Science, Technology, Engineering, and Mathematics (STEM). We used transcript data to categorize a student's intention to major in STEM at three points: their first semester, at the end of sophomore year, and at graduation. Students took several paths toward graduation including complete persistence in STEM, early switching, late switching and double switching. Using this classification scheme, we investigated any potential gender effect or "new building effect" and compared participation in an integrated introductory chemistry and biology course sequence to traditional discipline-specific sequences. For students where self-reported data on intention existed, either from matriculation surveys or major declaration, we compared our transcript classification with self-reported interest in STEM. Relying solely on institutional level data provides an inclusive, unbiased analysis with minimal missing data.

Keywords: STEM Persistence; Institutional Data; STEM attrition; Assessment; Evaluation

1. Introduction

1.1 Purpose of the Study

Motivated by assessment of externally awarded grants and internal programmatic practices, we sought to identify a procedure using student level institutional data to answer several questions relating to pathways though STEM at St. Olaf College. Some of these questions translate directly to other institutions; others serve as a model for the types of institution specific questions this methodology can quantify. We explored course taking patterns and pathways relative to STEM to answer three important institutional-level questions: (1) Are there gender and racial/ethnic differences in persistence in STEM? (2) Exploring time trends in intention and graduation rates can we quantify the changes in the number of STEM-intentioned students following the completion of the new science building and investigate trends in incoming AP credits? and (3) What differences do we see in trajectories based on beginning chemistry courses? These are questions specific to our College, but using transcript data and our course taking methodology allows for this high-level examination of institutional inquiries.

Often when an institution undertakes strategies to increase persistence in STEM, analysis of the effectiveness of the intervention is tracked by aggregate measures (e.g., graduation rates) or student self-reported data (e.g., surveys). Aggregate measures do not account for individual-level variation. For example, consistent graduation rates might not account for large numbers of individuals moving into and out of STEM fields. Assessment strategies targeting individual-level responses that rely on multiple administrations of a survey (e.g., pre- and post-test) often suffer from very low response rates, especially when students are asked to repeat the same survey multiple times. These low response rates raise questions about the reliability of the data and increase the likelihood of a biased sample.

Studies relying on institution-level data are often conducted at large universities where students enroll in specific STEM programs (e.g., School of Engineering) and/or require a major declaration at matriculation. At our liberal arts institution, major declaration is fluid and students are not required to declare a major until the end of sophomore or beginning of junior year. In addition, changing a major or "undeclaring" a major is a low-barrier activity. In this case it is more difficult to classify students regarding their STEM or non-STEM major intention.

1.2 Relevant Literature

Nationally, 52% of students who show initial intention to major in STEM persist to graduate in a STEM field (Chen 2013; Watkins and Mazur 2013; Wilson et al. 2012; Ackerman et al. 2013). Concerns about the retention of US college students in STEM fields (Chen and Ho 2012) have led to a series of studies designed to investigate factors associated with persistence in STEM. Some studies have focused on the persistence of underrepresented student populations (Valerio et al. 2014). Many studies use retrospective data to examine the role of demographic variables (race, gender, high school characteristics, socioeconomic status), preparation for college level STEM (SAT or ACT scores, AP or IB courses taken in high school) or college variables (school, classroom peers) as predictors of graduating with a STEM major. Based on reports such as these, researchers have suggested some strategies for reversing the decline in STEM majors (Holdren and Lander 2012).

Efforts to improve STEM retention include curricular redesign, implementation of active learning pedagogies, and enhanced advising or other support services (see, for example Watkins and Mazur 2013; Wilson et al. 2012). Others have used student-level data to identify factors that lead to persistence in STEM. Such studies often rely on a declared major at matriculation as a measure of intention. Ackerman et al. (2013) studied the transcripts of over 26,000 Georgia Tech students and determined that receiving credit for AP Calculus and successful completion of three or more STEM AP courses in high school were important predictors of graduating with a STEM major.

Griffith (2010) analyzed data from two national longitudinal studies (National Longitudinal Survey of Freshman and the National Education Longitudinal Study of 1988) to identify factors that lead to STEM persistence, particularly for women and minorities. In addition to reporting their initial intended major, students reported their major in their sophomore and senior years. This study also found that the number of high school advanced placement courses in STEM is a predictor of STEM persistence in college. Furthermore, during the first two years of college, students are more likely to persist in STEM if the ratio of their STEM grades to their non-STEM grades increases. In a multivariate analysis, Griffith found that after adjusting for institutional characteristics, women and men have similar persistence patterns, as do majority and minority students. The relationship between the timing of major declaration and graduation major has also been studied (Kokkelenberg and Sinha 2010) revealing that students are more likely to switch from a STEM major to a non-STEM major than vice versa.

1.3 Conceptual Framework

Following the model of "early switching, late switching, and complete persistence" proposed by Ma (2011) we have categorized *STEM-intention* at three different points during students' academic career and have identified pathways that examine switching into and out of STEM.

1.4 Institutional Context

St. Olaf College is a highly selective residential liberal arts college with about 3000 students. The college is organized into five divisions; the Faculty (division) of Natural Sciences and Mathematics (FNSM) consists of the five departments of biology, chemistry, physics, psychology and mathematics, statistics, and computer science. Within the cohorts entering from 2002-2015, about 40% of those graduating had at least one STEM major. Biology, mathematics, psychology, and chemistry are consistently among the largest majors at the college. Other majors outside the FNSM (specifically nursing, exercise science and environmental studies) require multiple STEM courses so we consider them STEM-related majors.

In comparison with many other institutions that have been studied in trajectories research, we have a relatively high four-year graduation rate. Our four-year graduation rate is 83% and the six-year rate is 87%. In addition to courses required to complete a major, students are required to complete a suite of courses to fulfill general education requirements. Graduates must complete 26 general education attributes which correspond to 14-16 courses outside of a major, since some attributes can be satisfied within a major, and some courses carry more than one attribute.

Many institutions are building new academic facilities and anecdotal evidence suggests that enrollment surges upon building completion. The FNSM moved into new space in 2008. We observed more students majoring in STEM fields after the building opened, along with improvement in student perceptions of interdisciplinary learning, students' ability to stay focused, and the environment for student learning (Van Wylen et al. 2013; Van Wylen and Walczak 2011; Walczak and Van Wylen 2013, 2015). Throughout the planning process for new facilities we promoted interdisciplinary interactions. This core principle is manifested in faculty office arrangement in the building (by research interests rather than departmental affiliation) and through our creation and endorsement of interdisciplinary programs such as Neuroscience, Integrated Introduction to Chemistry and Biology (CH/BI), and Mathematical Biology.

All of the institution's interdisciplinary courses and programs were created with specific outcomes in mind. The CH/BI course sequence was designed for beginning students with strong interests at the interface of these two fields. Some of the intended outcomes of this sequence and other disciplinary-focused programs have been assessed by evaluating student work or attitudes (Van Wylen and Walczak 2011). However, other outcomes are better assessed by looking at student-level institutional data.

2. Methods

2.1 Data Sources

In this paper we combined several data sources. With assistance of the Director of Institutional Effectiveness and Assessment, we used two primary sources of institutional data. First, we compiled demographic and graduation data for each of the 10,883 students in the 2002-2015 incoming cohorts including gender, degree type, honors, GPA, class rank, cohort year, major(s) and concentration(s), race/ethnic group, and official graduation year. Second, we gathered transcript information by cohort year including information on every class for which they received credit (year, term, institution, department, class number, class name, whether a class was graded or otherwise, grade received, and number of credits). AP course credit is also included on the transcript. Generally, students with AP exam scores of 4 or 5 receive credit, although the requirements vary somewhat by department. We reduced the data to consider only specific courses, categories of courses (e.g. level 2 or level 3 courses), and number of STEM courses taken to analyze a much simpler data set that still provided all the necessary information.

We also obtained self-reported matriculation interest data from surveys taken by 2,784 students matriculating from 2009-2012. Incoming students report interest on a survey in which they are asked to list up to three possible majors. We collected major declaration data from any majors declared before the February of a student's junior year, as that was the posted date for a student to declare their major (though typically students declare by the end of their sophomore year). Major declaration data from a total of 2,603 students was available.

Table 1 summarizes the data used in our classification and validation methods. Students were classified at three points during their college careers, as shown in the Timing row. The courses taken first semester were used to classify beginning students as STEM or non-STEM. In our validation process we used information from the matriculation survey as an indicator of a student's intention of completing a STEM major as they started college. Similarly, Table 1 shows that classification at the end of sophomore year included transcripted courses and the validation process involved declared majors.

2.2. Data Coding Using Completed Coursework

Using transcript data alone, we analyzed coursework and identified criteria for categorizing students who appeared to be working towards a STEM major. We found that having 5 or more STEM courses by the end of sophomore year was a good predictor of someone working towards a STEM major, as 95% of students who graduate with a STEM major have at least 5 STEM courses by the end of sophomore year. To illustrate this, we examined course-taking behavior for different majors. Figure 1 shows the total number of STEM courses taken over five different students' college careers. The STEM majors, panels (a) and (d), have steadily increasing slopes throughout college; the double STEM major (d) accelerates STEM course taking in the last two years. The STEM-related majors, panels (b) and (e) begin college with a robust schedule of STEM courses, but as they move through college they take fewer STEM courses. The exercise science major takes STEM courses steadily through junior year; the nursing student abruptly ceases taking STEM courses after sophomore year when clinical courses begin. The history major (c) in this example takes the required three STEM courses, one each in their first, fourth and sixth semesters.

Using transcript data alone, we analyzed coursework and identified criteria for categorizing students who appeared to be working towards a STEM major. All courses taken by students in their first semester of their first year were available in the institutional data. Students typically take four full credit courses each semester. First semester students typically have one prescribed course (e.g., first year writing or first year religion) and three flexible courses. Students are advised to choose classes in their areas of potential majors or classes for general

Timing	First Semester	End of Sophomore Year	Graduation
Inputs: Student Transcript Data	Enrolled STEM course(s)	Completed STEM course(s)	Major(s) awarded
Classification	STEM or non-STEM	STEM or non-STEM	STEM, STEM-Related, or non-STEM
Validation Data	Matriculation Survey	Declared Major(s)	

Table 1. Summary of Data Inputs, Classification and Validation

education requirements.

We looked at all STEM courses that students took in their first semester from 2002-2015 to identify courses that were related to graduating with a STEM major. Students enrolled in courses highly associated with graduating with a STEM major in their first semester were classified as STEM-intentioned. For example, within our data we found that 80.3% of students who take a 200-level STEM course their first semester ended up graduating with a STEM major. We next calculated the percentage of students who took a 100-level STEM class first semester and graduated with a STEM or STEM-related major. The findings which informed our procedure are reported in Table 2. If the percentage of STEM plus STEM-related majors who started in a class was greater than 50%, we categorized students starting in that course as STEM-intended. Finally, we found that 71.8% of students taking 2 or more STEM classes their first semester major in STEM so we also categorized students taking two or more STEM classes as STEM-intentioned.

For the second categorization, we identified major requirement courses that a student would usually complete for each STEM major by the end of their second year (Table 3) and the total number of STEM courses students completed before the end of their sophomore year. The major-specific courses identified (Table 3) are not typically taken for general education credit; we found that over 75% of students who had taken each of these classes by the end of their sophomore year graduated with a STEM major.

2.3 Analysis Plan of Student Pathways/ General Course Taking Patterns

Following the model of "early switching, late switching, and complete persistence" proposed by Ma (2011) we categorized STEM-intention at three different points during students' academic careers: (1) during a student's first semester, (2) at the end of sophomore year, and (3) at graduation. Students were categorized as STEM or non-STEM at each of these three points. At graduation we included a third category "STEM-related," which included exercise science, environmental studies, and nursing majors.

The first semester marker indicates whether a student is interested in pursuing STEM at the time they start college. The second point, the end of sophomore year, aligns with institutional expectations that students declare a major by this stage. Moreover, it would be practically and logistically challenging to begin many STEM programs after this point and still graduate in four years. The third point uses students' actual majors at graduation to categorize the student as STEM, non-STEM or STEM-related. Approximately 34% of our students graduate with more than one major. For purposes of categorizing students as STEM, a student with *at least one* STEM major is considered a STEM major.

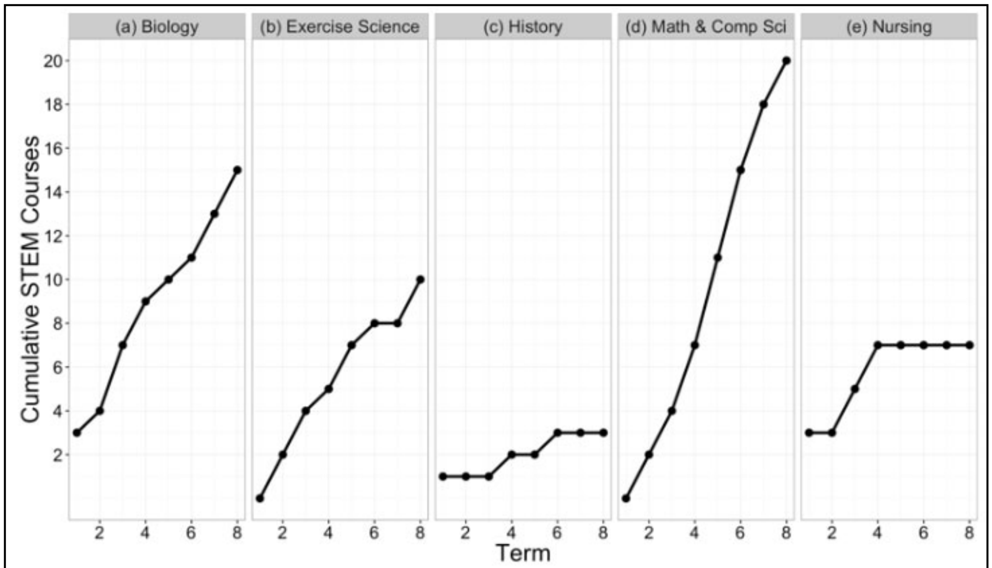

Figure 1. Cumulative number of STEM course for a sample of students/majors.

Starting STEM Course	% STEM Major at Graduation	% STEM-related Major at Graduation	Number of students
BIO 125: Cell Biology and Genetics	49.7	23.7	687
BIO 126: Evolution and Diversity	70	10	20
CHEM 121: General Chemistry	61.2	7.1	464
CHEM 125: Structural Chemistry and Equilibrium	79.1	4.3	979
CH/BI 125: Integrated Introductory Chemistry and Biology I	87.5	1.6	128
CSCI 125: Computer Science for Scientists and Mathematicians	83.3	0	12
MATH 126: Calculus II	61.8	4.3	508
MATH 128: Honors Calculus II	73.4	1.9	674
PHYS 124: Principles of Physics I	100	0	1
PHYS 126: Analytical Physics II*	81.2	1.3	149
PHYS 130: Analytical Physics II*	85.5	1.4	237
PSYCH 130: History and Methods of Psychology	100	0	4

*In 2008-09 the pedagogical approach for the Analytical Physics course sequence changed markedly and the course number changed as a consequence.

Table 2. First semester, first year courses used to categorize STEM-intention

Course	% STEM Major at Graduation	% STEM-Related Major at Graduation	Number of Students
BIO 126: Evolution and Diversity	79.7	5	1565
BIO 233: Intermediate Genetics	93.1	0.8	475
CHEM 248: Organic Chemistry II	93.4	0.8	865
CSCI 241: Hardware Design	80.8	0	73
MATH 244 or 252			
244: Real Analysis I	94.6	0.6	504
252: Abstract Algebra I			
PHYS 244: Modern Physics	100	0	183
PSYCH 230: Research Methods	92.5	2.1	93

Table 3. Courses taken by end of second year that signal STEM-intention

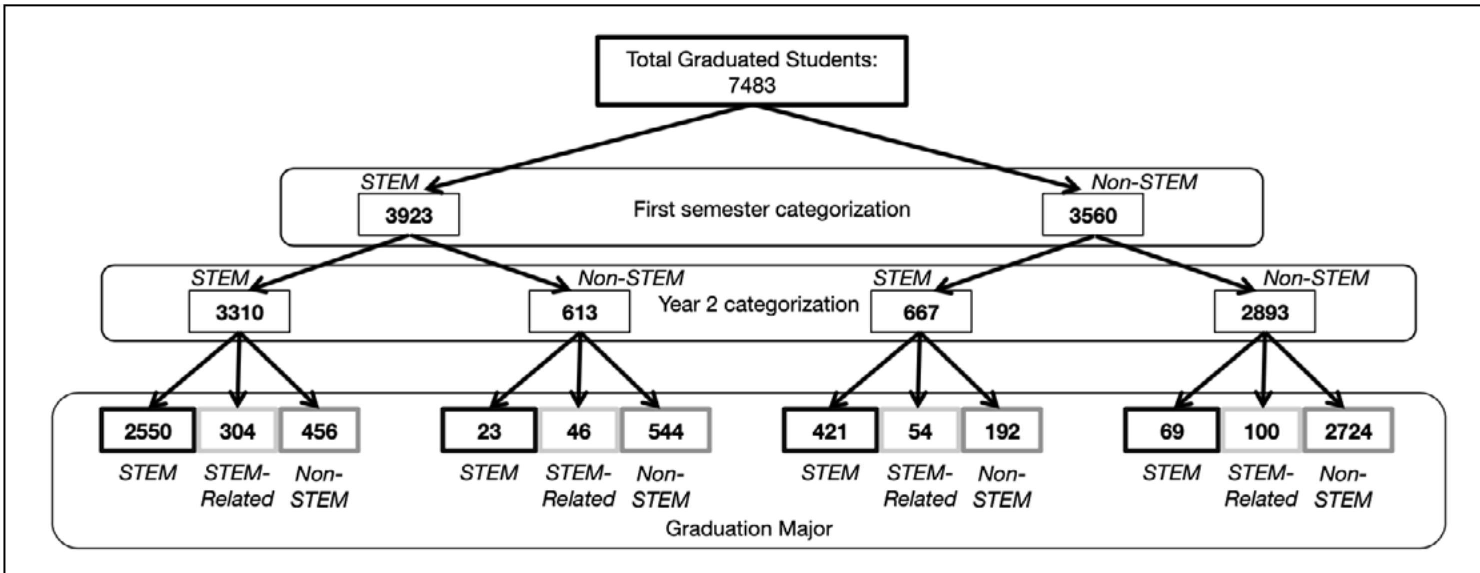

Figure 2. STEM-intention categorization for students graduating from college. Students are categorized as STEM or non-STEM based on the courses taken during the first semester of college, on the course completed by the end of sophomore year and on the majors(s) awarded at graduation.

Students categorized as STEM-intentioned during first semester but not at the end of the sophomore year can be characterized as "early switching." Students identified as STEM-intentioned both in the first semester and at the end of the sophomore year but do not graduate with a STEM major suggest evidence of "late switching." Some students are characterized differently at each point (STEM, non-STEM, STEM or non-STEM, STEM, non-STEM) and are described as "double-switchers." Finally those students who are categorized as STEM or non-STEM at all three markers are those who exhibit "complete persistence."

2.4 Validation

To validate our procedure, we looked at a) matriculation interest and b) major declaration data. Table 1 illustrates how these two sources of data were used. A student was coded STEM-intentioned if one or more of a student's potential majors entered on the matriculation interest survey was STEM. Any students who did not state their major interest at the time of the survey were not included in this analysis. Major declaration data was used in a similar way. Any student who declared at least one STEM major is considered a STEM student at the end of their sophomore year. Any students who did not declare a major at the end of sophomore year were not included in this analysis.

Self-reported matriculation interest data was compared with first-semester STEM categorization, and major declaration data was compared to end-of-sophomore-year STEM categorization. For both a student's first semester and the end of their sophomore year, we looked at the overall classification agreement, as well as agreement by gender and race. For each demographic group, we looked at the total number of students in that group, the total agreement of our procedure, which we defined as percent correct classification. We also report agreement for STEM and non-STEM classifications in addition to Cohen's Kappa.

2.5 Institutional- and Program-Level Assessment: Gender, Time Trends, Integrated Chemistry and Biology

Using our categorization scheme, we calculated the percentage of male and female students that are STEM-intentioned based on their coursework in the first semester, at the end of the sophomore year and at graduation. Since our students frequently switch into or out of STEM during college, we also examined the gender breakdown of all students indicating STEM-intention at each categorization point. This includes not only the complete persisters in STEM, but also those students classified as STEM either at the end of sophomore year or at graduation regardless of their path.

Institution-level data included self-reported gender of male or female and self-reported race/ethnicity, which was categorized into four groups: Asian, Domestic Multicultural, Hispanic, International and White. Domestic Multicultural includes the following groups: American Indian or Alaska Native, Black or African American, Multiracial, Native Hawaiian or Pacific Islander.

We explored our categorization of intention along with STEM graduation rates, and incoming AP STEM credits over time. We considered three groups of AP STEM students: those with zero AP STEM credits, those with one or two AP STEM credits and those with three or more AP STEM credits. We compared these trends before and after the opening of the new science building.

3. Results

3.1 Overall Rates, Trends and Trajectories

As shown in Figure 2, 7,483 students graduated from college among the cohorts matriculating between 2002 and 2012. The cohorts from 2013, 2014, and 2015 had not yet graduated from college at the time of this analysis so they were excluded. The first split between STEM and non-STEM (3,923 STEM and 3,560 non-STEM) is based on the first semester categorization; 52% of incoming students take courses that suggest they intend on a STEM major.

Of the students who indicate an early intention to major in STEM (3,923), 84% (3,310) persist in STEM through the end of sophomore year; the other 16% (613) are "early switching" students who switch out of STEM during their first two years of college. About three fourths (2,550) of the 3,310 students that persist as STEM-intended at the end of sophomore year graduate with a STEM major (complete persistence in STEM), although 9% (304) major in STEM-related fields and 14% (456) are "late switching" students who leave STEM majors in the last half of college. Five hundred forty four (89%) of the 613 early switching students graduate with non-STEM majors; 23 (4%) of these 613 students graduate with a STEM major and are classified as "double-switchers."

A minority of the 3,560 students that don't appear to have an interest in STEM when they begin college (19%) are categorized as STEM students at the end of sophomore year (667). Sixty three percent (421) of these students graduate with a STEM major and are considered "early switching" *into* STEM. Among this group of 667 students who switch into STEM after beginning college, 29% (192) are "double-switchers" who re-join the non-STEM students by the time they graduate.

Finally, 81% of the students who do not begin college as STEM students (3,560) persist as non-STEM students at the sophomore year (2,893). Nearly all of these students (94%) graduate with a major outside of STEM (complete persistence in non-STEM) (2,724). Only 69 (2%) of these 2,893 students are late-switchers into STEM. As mentioned, students who major in nursing, exercise science and environmental studies are required to take several STEM courses. Although the number of STEM-related majors in our dataset is relatively small (504 or 7%), 60% of these students (304) are late-switchers out of STEM. The early STEM course taking patterns for students with these majors (see Figure 1b and 1e) are consistent with categorization of late switching out of STEM.

3.2 Early-Switchers, Late-Switchers, and Double-Switchers

For each of the categories of switchers (early-, late- and double-), students can either switch out of or into STEM. Considering early switching students, for whom the STEM or non-STEM classification changes between first semester and end of sophomore year, we see fewer students switch out of STEM (16%; 613/3,923) as compared to those who switch into STEM (19%; 667/3,560). Among the late-switchers who change classification between end of sophomore year and graduation, only 2% of students who are categorized non-STEM at the end of sophomore year graduate with a STEM major (69/2,893) while 14% of those classified as STEM at the end of year two graduate with a non-STEM major (456/3,310).

For each of the six groups of switchers, we identified the awarded major(s) of the students and described some trajectories through common majors (Table 4). For comparison purposes, characteristics of complete persisters (either complete STEM or non-STEM persisters) are also included in Table 4. We will first consider those groups that majored in STEM.

3.3 Graduating with STEM Major(s)

There were 421 students who made an early switch from non-STEM to STEM. They were classified as non-STEM in their first semester, but went on to be classified as STEM at the end of sophomore year and graduate with a STEM major. Of these early-switchers into STEM, 237 (56%) are single majors and 184 (44%) were double majors. Among these students, second majors came from each of the five Faculties. Of the 421 early-switchers into STEM, 168 (40%) paired a STEM major with a major from Fine Arts (34; 8%), Humanities (45; 11%), Interdisciplinary and General Studies (27; 6%) and Social Sciences (62; 15%). Five of these students had triple majors (1 STEM and 2 non-STEM majors). Sixteen students (4%) had two STEM majors (e.g., Computer Science and Mathematics, see Figure 1d). Most frequently, mathematics was coupled with computer science (6 students). The most frequently awarded STEM majors among this group of early-switchers were psychology (63%), biology (15%) and mathematics (15%).

For comparison, information about the students that persisted in STEM at all three points of classification are also provided in Table 4. Note that the 2,550 persisters in STEM were slightly more likely to be single majors (61%) and the percentage of psychology majors among this group is much lower (10%), while the percentage of biology and mathematics majors are higher (42% and 22%, respectively). Of the STEM persisters, 968 (38%) paired a STEM major with one other major and 39 (2%) had three majors.

Complete STEM persisters were more likely to have multiple STEM majors than the early switcher *into* STEM group. While only 4% (16/421) of the early-switchers into STEM had two STEM majors, 15% (377/2,550) of the STEM persisters had two or three STEM majors.

The STEM persisters that had two majors paired a STEM major with majors from Fine Arts (84; 9%), Humanities (217; 22%), Interdisciplinary and General Studies (134; 14%) and Social Sciences (185; 19%). Three hundred forty-eight (36%) of the double majors combined two STEM majors. The most frequent combinations were mathematics and physics (122; 35% of the STEM-STEM double majors), chemistry and mathematics (65; 19%) and biology and chemistry (62; 18%).

Of the 39 students that persisted in STEM throughout college and completed three majors, 11 had three STEM majors, 18 had two STEM and one other major, and 10 had one STEM major and two other majors. The majors from non-STEM areas of the college were distributed similarly to the numbers cited above for the double majors. The triple STEM majors all included mathematics as one of the majors and the most frequent combination was computer science, mathematics and physics (7/11). Within each of the three different groups of triple majors (3 STEM, 2 STEM + 1 Other, 1 STEM + 2 Others) that persisted in STEM, mathematics was the most commonly included major.

Other students made late switches into STEM between the end of sophomore year and graduation. Sixty-nine students are characterized as non-STEM-intentioned at our first and second year points, but graduate with at least one STEM major (58% are single majors and all of the double majors have only one STEM major). Generally, these students start course-work in their STEM major during the junior year, and have at least one year of college in which no STEM major classes are taken. Forty nine of the 69 late-switchers into STEM are psychology majors.

The double-switchers are the final category of students switching into STEM. Twenty-three students are originally classified as STEM as first year students, are classified non-STEM as sophomores, yet graduate with a STEM major. Forty three percent of these students are single majors and all the double majors have only one STEM major. This group of double-switchers typically enter college with AP credit for Calculus, take no STEM courses one year (typically as sophomores) and tend to concentrate their STEM courses in the junior and senior years. The most frequent STEM majors among this group are mathematics (7) and psychology (7).

Considering all 513 (23+421+69) switchers who graduate with at least one major in STEM, 56% have only one major compared to 61% of complete persisters. This is perhaps expected since this group of students spends more time taking courses in other fields since they are not classified as STEM at some point in their college career. The majority of the double-switching STEM majors are psychology majors. Our psychology major is not nearly as linear as some other STEM majors (e.g., physics and chemistry), so it is possible for students to consolidate their psychology courses into fewer years than for these other majors. In fact, among all three categories of STEM major switchers (early-, late- and double-; a total of 529 *majors*), there are only five physics majors and twelve chemistry majors. Over 60% of the STEM major switchers are psychology majors and about 15% are each biology and mathematics majors. The remaining 10% are chemistry, computer science and physics majors.

	Early Switchers		Late Switchers		Double Switchers		Complete Persisters (Students)	
	n	%	n	%	n	%	n	%
STEM Majors	421		69		23		2550	
Single Majors	237	56	40	58	10	43	1543	61
Psychology Majors	266	63	49	71	7	30	267	10
Biology Majors	65	15	6	9	4	17	1072	42
Mathematics Majors	64	15	9	13	7	30	567	22
non- STEM Majors	544		456		192		2,724	
Single Majors	345	63	385	84	160	83	1809	66
Economics Majors	104	19	107	23	39	20	299	11
English Majors	78	14	46	10	22	11	374	14
Political Science Majors	38	7	25	5	12	6	188	7
Music	65	12	22	5	5	3	383	14
Sociology/Anthropology	24	4	30	7	12	6	263	10
Art	37	7	29	7	21	11	122	4

Table 4. Characteristics of students who switch between STEM and non-STEM classifications

3.4 Graduating with No STEM Major

The largest group of switchers that graduate in non-STEM fields are early-switchers out of STEM. These 544 students were identified as STEM in their first semester, were categorized as non-STEM at the end of their second year and continued to earn a degree with a major outside of STEM. Of these 544 students, 63% are single majors with 19% majoring in economics and 14% majoring in English. It should be noted that economics and English are frequently among the most common majors outside of STEM on campus.

The 456 late-switching students that graduate with non-STEM majors are often single majors (84%). Overall, 33% of these students major in the most frequent non-STEM majors, economics (23%) and English (10%). The 192 double-switchers that graduate with a non-STEM major are also typically single majors (83%). The most frequent majors in this group are economics (20%), English (11%) and art (11%).

For comparison, the students classified as non-STEM at all three points (non-STEM persisters) are also shown in Table 4. About two-thirds of these students are single majors. This is comparable to the early-switchers, but quite a bit lower than the late and double-switchers. The percentage of students majoring in economics among the persisters (11%) is also a little lower than in each of the switcher groups (19-23%). The percentage of English majors in the persister and switcher groups, on the other hand, are more comparable. The early-switcher English majors (14%) represent the same fraction of the population as the persisters; the late and double-switchers are a little less likely to be English majors (10% and 11%, respectively). Taken together, this suggests that economics majors that switch out of STEM tend to be late-switchers (STEM, STEM, non-STEM) and English majors switch out of STEM earlier (STEM, non-STEM, non-STEM).

	First Semester Total		First Semester Agreement		First Semester Kappa	End of Sophomore Year Total		End of Sophomore Year Agreement		End of Sophomore Year Kappa
Overall	2784		2115	(76%)	0.48	2603		2204	(85%)	0.67
		STEM	1334	(73%)			STEM	1156	(94%)	
		Non-STEM	781	(82%)			Non-STEM	1048	(76%)	
Gender										
Male	1192		927	(78%)	0.48	1079		911	(84%)	0.65
		STEM	620	(76%)			STEM	510	(94%)	
		Non-STEM	307	(81%)			Non-STEM	401	(75%)	
Female	1592		1188	(75%)	0.45	1524		1293	(85%)	0.67
		STEM	474	(70%)			STEM	646	(95%)	
		Non-STEM	714	(83%)			Non-STEM	647	(77%)	
Race/Ethnicity										
Asian	141		109	(77%)	0.38	128		111	(87%)	0.58
		STEM	73	(75%)			STEM	66	(97%)	
		Non-STEM	36	(82%)			Non-STEM	45	(73%)	
Domestic Multicultural	147		98	(67%)	0.22	135		111	(82%)	0.52
		STEM	53	(58%)			STEM	48	(94%)	
		Non-STEM	45	(82%)			Non-STEM	63	(75%)	
Hispanic	107		68	(64%)	0.17	98		79	(81%)	0.45
		STEM	38	(53%)			STEM	31	(86%)	
		Non-STEM	30	(86%)			Non-STEM	48	(77%)	
White	2250		1740	(77%)	0.5	2109		1843	(87%)	0.67
		STEM	1113	(75%)			STEM	1008	(94%)	
		Non-STEM	627	(82%)			Non-STEM	835	(76%)	

Table 5. Validation of course taking intention: student matriculation interest and major declaration data

3.5 Validation of Categorization Procedure with Self-Reported Data

Table 5 shows the classification accuracy of our procedure compared to student self-reported data at two points in time – the beginning of first semester and end of the sophomore year. For end of sophomore year classifications, the agreement percentages are higher than those at the end of the first semester. Our course taking categorization procedure better captures STEM students at the end of the sophomore year, as compared to first semester, but is not as good as classifying non-STEM students. The lowest agreement at the end of sophomore year of any demographic group was for Hispanic and Domestic Multicultural Students with 81% and 82% agreement, respectively, between the course-taking categorization and declared major.

3.6 Assessment

3.6.1 Gender differences among STEM majors.

We note in Table 6 that 52% (3,923/7,483) of students are STEM-intentioned at the first semester. Of the 3,923 students, 44% (1,742) are male and 56% (2,181) are female, similar to the gender demographics of the 7,483 graduates in our study (43% male; 57% female).

Considering the STEM persisters along the left side of Figure 2, the gender breakdown at the three categorization points is illustrated in Table 6. Of the 3,923 students classified as STEM-intentioned in the first semester, 1,832 of the 2,181 female students (84%) are subsequently classified as STEM at the end of the sophomore year. This is comparable to 85% of males (1,478/1,742). Seventy-four percent of females classified as STEM at the end of their sophomore year progress to graduate with a major in STEM (1,355/1,832), compared to 81% (1,195/1,478) of males. When we include STEM and STEM-related majors we see a similar rate of females and males persisting from STEM at the end of the sophomore year to graduation: 87% for females and 86% for males (males: 1,266/1,478; females 1,588/1,832).

Of the 667 students who are first classified as STEM-intentioned at the end of the sophomore year, 65% of females (257/394) continue to graduate with at least one major in STEM, compared to 60% of males (164/273). When we expand this to include STEM and STEM-related majors we see 74% of females and 67% of males continuing to graduate with at least one STEM major.

3.6.2 Trends in Intention, STEM Graduation, and AP Credits

Over the last decade, a growing percentage of students matriculated with transcripted STEM AP course credit as shown in Figure 3. The percentage of students with one or more AP STEM credits moves from under 30% in the early cohorts of our data to over 40% in 2011 and 2012. Additionally, 81% of students with three or more AP STEM credits graduate with at least one STEM major. This is in comparison to 56% of students with 1-2 AP STEM credits and 32% of students with zero AP STEM credits.

Our data set, which includes cohorts matriculating from 2002 to 2015, can also be used to illustrate the "new building effect." As mentioned, institutions anecdotally report an increase in student interest in fields upon completion of new facilities for those programs. Figure

	Complete STEM Persisters					Students first classified as STEM at the end of sophomore year				
	Male		Female		Total	Male		Female		Total
All Graduates	**3185**	**43%**	**4298**	**57%**	**7483**					
First Semester STEM	1742	55%	2181	51%	3923					
End of Sophomore Year STEM	1478	85%	1832	84%	3310	273	41%	394	59%	667
Graduation (Awarded Major)										
STEM	1195	81%	1355	74%	2550	164	60%	257	65%	421
STEM + STEM related	1266	86%	1588	87%	2854	182	67%	293	74%	475

Table 6. Percent of male and female students continuing from classification points

3 illustrates the percentage of students showing STEM-intention by cohort year. Data for categorization at first semester, end of second year and at graduation are shown. The highest percentages come from the first semester categorization, with 48% of incoming students in 2002 classified as STEM per our categorization scheme. This percentage grows over years 2002-2008 and levels off around 54%, where it holds until about 2013.

Similar trends are shown in Figure 3 for second year categorization and graduation. For the cohorts entering college from 2002 to 2008, the percentage of students who are classified as STEM-intended at the end of sophomore year increases from 48% to 58%. This level remained relatively constant until the 2012 cohort. Similarly, the percentage of students graduating with a STEM major increases from 37% in 2002 to around 45% for 2008-2012.

We can also use the data in Figure 3 to make comparisons within a cohort. The drop in percentage of STEM-intentioned students between the first semester and second year represents the overall effect of students moving out of STEM fields. For students beginning college in 2002-2007 this drop is between 2-8% (average 4%). After 2007, however, the two lines converge with very little drop off in STEM-intention between first and second years. In fact, for the class enrolling in 2008 (and later), more students were classified as STEM-intentioned at the end of sophomore year than during their first semester of college!

The data for the 2007 cohort year indicates a slight increase in STEM-intention at year 2, but a marked increase in STEM graduation. These students began their college STEM education in the old facilities but spent three years taking STEM classes in the new facilities. The 2008 cohort, who began college as the new building opened, show a large increase in STEM-intention at year 2 and graduation. The percentage of STEM majors at graduation is about 12% lower than STEM-intention at the end of sophomore year. This is consistent across all cohort years. Although there is an upward trend of percentage of STEM majors, it follows the STEM-intention of the students in each cohort. The right side of Figure 3 shows data for students who had not yet graduated at the time of our study. From the limited available data, it appears that the surge in STEM-intention that we experienced with the opening of the new building may be waning.

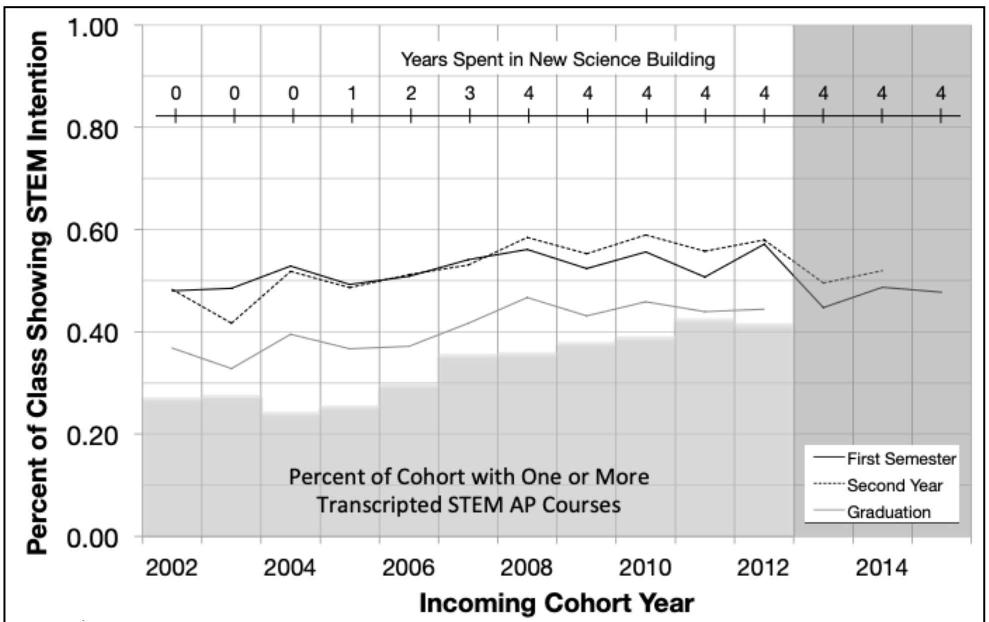

Figure 3. Illustration of building effect by incoming cohort.

3.6.3. Specific Programmatic Assessment: Integrated Chemistry and Biology

Individual-level transcript analysis can also be used in programmatic assessment. Our Integrated Introductory Chemistry and Biology (CH/BI) sequence was developed, in part, to make it possible for students to take more science classes earlier in their college careers (Van Wylen et al. 2013). This outcome lends itself well to transcript analysis.

Table 7 shows a comparison of the number of STEM courses taken by the end of sophomore year and upon graduation for students entering college between 2007-2012. The CH/BI sequence began in 2007, so comparison of earlier cohorts is not possible. Chemistry and biology students can begin by taking one of four science courses: BIO 125, CHEM 121, CHEM 125 or CH/BI 125. A chemistry placement exam is used to recommend students for CHEM 121, 125 or CH/BI 125. Students enrolling in one of the disciplinary sequences (BIO 125, CHEM 121 or CHEM 125) complete 7-8 STEM courses by the end of their sophomore year; CH/BI students complete 9-10 STEM courses in this same time. Thus, the goal of students taking more STEM early is achieved.

Interestingly, the early acceleration continues through the second half of college. Upon graduation, students that began in the disciplinary sequences complete 12-16 STEM courses while the CH/BI students complete 18. In addition, students beginning in the CH/BI sequence complete three 300-level STEM courses in college, but students beginning in the disciplinary sequences only complete one or two high level courses.

4. Discussion

Nationally, about 28% of students begin college intending on a STEM major (Chen 2013), while at our institution, 52% [3,923/7,483] of students show initial STEM-intention. This could be somewhat inflated due to general education requirements and the limitations of our categorization; nevertheless, this high percentage shows an overall student population which has strong interest in STEM fields.

Among college graduates nationally, 52% of students with initial STEM-intention graduate with STEM majors. At our institution 65% [2,550/3,923] of our graduates who are categorized as STEM-intentioned in their first semester persist in STEM through graduation. These characteristics also point toward a student body with higher than average STEM persistence. Others have shown that students are more likely to switch out of

Starting STEM Class	Students Enrolled (n)	Total Number of STEM courses Completed		
		By end of sophomore year Mean (Median)	Upon graduation Mean (Median)	At the 300 level Mean (Median)
BIO 125	329	7.02 (7)	12.10 (12)	1.47 (1)
CHEM 121	312	8.08 (8)	13.77 (15)	1.47 (1)
CHEM 125	416	7.97 (8)	15.53 (16)	2.44 (2)
CHEM/BIO 125	138	9.35 (9)	17.60 (18)	3.03 (3)

Table 7. STEM participation by introductory course for entering cohort years 2007-2012

STEM than into STEM (Griffith 2010; Ma 2011). Our early switching students, for whom the STEM or non-STEM classification changes between first semester and end of sophomore year, are less likely to switch out of STEM (16%; 613/3,923) than into STEM (19%; 667/3,560).

Persistence in STEM by gender in our cohort is somewhat consistent with other studies. Griffith (2010) found that overall persistence in STEM by women is lower than men, but this difference disappears if student preparation is taken into account. Ma (2011) found that women and men persist equally if the women had decided as high school students to pursue STEM majors. Our analysis shows some difference in persistence between men and women with 38% of men and 32% of women classified as complete persisters. Considering gender at each of the classification points we do not see meaningful gender differences. In fact, once students are classified as STEM we see relatively similar rates of "retention" to the next classification point between men and women. We see a higher percentage of women who are first classified as STEM-intentioned at the end of the sophomore year. This is largely attributed to our method that includes psychology as STEM, but excludes the introductory psychology course in the first semester as an indication of being STEM-intentioned (since many students take this course to complete general education requirements and will not continue in psychology). It is not clear from this level of inquiry whether this difference in gender retention is actionable. Having this data available, however, will allow for further exploration in the future. For example, the relationship of other demographic characteristics, such as race and ethnicity, to persistence could be examined.

Due to the low barriers toward changing majors and relaxed course sequencing within majors at our liberal arts school, categorizing students as STEM/non-STEM proves to be much more difficult and nuanced than it may be for other programs in which students must enroll directly into the school (or program) from which they hope to earn a degree. Liberal arts education also allows many more opportunities for students to lose (and/or gain) interest in a variety of fields, and more easily allows movement between fields. While our categorization cannot capture the complete fluidity of this process throughout all four years, nor capture what students are truly thinking, it gives some insight into what factors can affect a liberal arts student's chances of persisting in STEM and thus contributing to a diverse and robust workforce founded in a solid education of STEM principles.

The strength of our methodology arises from our use of institutional data rather than student self-report data. This approach prevents us from unintentionally adding bias that arise from student behaviors related to declaring majors. It also gives us the opportunity to easily link with other institutional data, such as SAT or ACT scores, high school GPA, AP or IB courses, and the record of major declaration for an individual student.

After looking at the classification accuracy of our method, we found that it correctly identifies students with their self-reported data a majority of the time. We found that, on the whole, demographic groups had similar classification accuracy. Within STEM and non-STEM majors, there was very little variation in accuracy. Classification accuracy was also consistent within demographic categories (e.g., race and gender). As shown in Table 5 for gender, the largest difference in the classification method accuracy was for STEM students at the end of their first semester. Males were correctly identified 76% of the time and females 70% of the time. For racial and demographic groups, the largest difference in the classification method's accuracy was at the end of their first semester STEM classification with Hispanic students correctly categorized 53% of the time and Asian and White students 75% of the time.

This suggests that the qualities used to classify students as STEM or non-STEM may be appropriate for all demographic groups. However, some demographic groups did have lower classification accuracy. The highest classification accuracy was 97% for Asian STEM students after their sophomore year. This would suggest that our classification method is a good predictor both overall and for individual demographic groups. However, further investigation is needed to explore why some student groups have lower classification accuracy than others.

The ability of our classification method to correctly identify STEM students increased from first semester to sophomore year. The overall agreement at the end of first semester was 73%, whereas the overall agreement at the end of sophomore year was 94%. The ability of our method to correctly classify non-STEM students, however, decreased from 82% to 76% from first semester to sophomore year.

4.1 Limitations

This approach to classifying students as STEM at various points in their college careers is useful for different assessment purposes. Although not relying on self-reported data provides some advantages, this approach does have some limitations. First, our definitions of STEM and STEM-related courses/majors are unique to St. Olaf College and may not apply at other institutions. Second, this method only employs data from student transcripts. We do not have records of individuals who dropped courses prior to the deadline. We suspect that we are not including a cadre of students who begin in STEM but drop early in the semester so the course does not appear on their transcripts. Third, we did not consider courses taken at other institutions (e.g., study abroad or transferred courses). Although many of our STEM majors do study abroad during college, few of these students take STEM courses off campus. Students who study abroad during a semester of their sophomore year, in particular, are potentially mis-categorized as non-STEM at the end of sophomore year. Those students can still graduate with a STEM major by doubling up on STEM courses in their last two years or by taking STEM courses in the summer.

4.2 Future Directions

In spite of these limitations, our categorization method for determining STEM-intention at various points during college demonstrates a useful approach for using institutional-level data to answer meaningful student-level questions. In addition, there is still much to learn from our data about the students that switch into and out of STEM. We anticipate that further analysis will provide insight into characteristics of the "switching" students. Through this investigation we hope to better understand the reasons behind switching into or out of STEM, determine which factors (e.g., STEM GPA, high school preparation) increase the likelihood of persistence, and ultimately identify ways in which we might better support students who might, in other circumstances, persist in STEM. Finally, we will investigate ways in which our process can be updated to include criteria to identify students with demonstrated capacity to work in STEM and STEM-related fields following graduation. Our institution attracts a large number of STEM-intentioned students, and although some of those do not graduate with a STEM major we hypothesize that many have both depth and breath in STEM for opportunities in the STEM workforce.

4.3 Conclusions

Our approach to transcript analysis illustrates a new method for categorizing students' STEM-intention without relying on self-report or institutional major declaration. Once we created our categorization for students at three different points in their college career, we were able to analyze and answer some of the important questions about specific groups of students in STEM.

Thirty four percent (2,550/7,483) of our graduates are complete persisters in STEM and 36% (2,724/7,483) are never classified by our categorization method as STEM-intentioned (non-STEM persisters). Thirty-eight percent of our graduates who are classified as STEM at some point in their college careers do not graduate with a STEM major [1- (2,550+304+456+23+46+544+421+54+192)/ 7,483].

The students that switch between STEM-intentioned and non-STEM-intentioned during their college career are particularly interesting. Each of the six groups of switchers has unique characteristics. Those who eventually graduate with a STEM major are likely psychology or mathematics majors; non-STEM graduates that switch out of STEM are most likely economics or English majors. While consistent with institutional perceptions that chemistry, computer science and physics are linear, prescriptive majors, our analysis of the switching patterns of students by major reinforces that psychology, mathematics and biology are more flexible majors. This same methodology can be applied to future students in assessing course taking trends and pathways through STEM.

By examining student course taking patterns, we are able to correctly categorize students as STEM-intentioned or non-STEM-intentioned at three points during their college career. This approach relies only on transcript data sources, removing any bias introduced by relying on student self-reported data.

Bibliography

Ackerman, P. L., Kanfer, R., & Calderwood, C. (2013). High School Advanced Placement and Student Performance in College: STEM Majors, non-STEM Majors, and Gender Differences. *Teachers College Record, 115*(10).

Chen, X. (2013). STEM Attrition: College Students' Paths into and out of STEM Fields. Statistical Analysis Report. NCES 2014-001. *National Center for Education Statistics.*

Chen, X., & Ho, P. (2012). STEM in postsecondary education: entrance, attrition, and coursetaking among 2003-2004 beginning postsecondary students. NCES 2013-152. National Center for Education Statistics.

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, doi:Doi 10.1016/J.Econedurev.2010.06.010.

Holdren, J. P., & Lander, E. (2012). Report to the President—Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. *President's Council of Advisors on Science and Technology.*

Kokkelenberg, E. C., & Sinha, E. (2010). Who succeeds in STEM studies? An analysis of Binghamton University undergraduate students. *Economics of Education Review, 29*(6), 935-946, doi:Doi 10.1016/J. Econedurev.2010.06.016.

Ma, Y. Y. (2011). Gender Differences in the Paths Leading to a STEM Baccalaureate. *Social Science Quarterly, 92*(5), 1169-1190, doi:Doi 10.1111/J.1540-6237.2011.00813.X.

Valerio, J., Chen, X., Gonzalez, H. B., & Kuenzi, J. J. (2014). *Attrition in science, technology, engineering, and mathematics (STEM) education data and analysis* (Education in a Competitive and Globalizing World). New York: Nova Publishers.

Van Wylen, D. G., Abdella, B. R. J., Dickinson, S. D., Engbrecht, J. J., & Vandiver, R. (2013). Interdisciplinarity: The Right People, a Supportive Place, and a Program Emerges. *Cbe-Life Sciences Education, 12*(2), 140-143, doi:Doi 10.1187/ Cbe.13-01-0001.

Van Wylen, D. G., & Walczak, M. M. (2011). Connectedness by design: the teaching laboratories in St. Olaf College's Regents Hall of Natural and Mathematical Sciences. *J Coll Sci Teach, 41*, 44-52.

Walczak, M. M., & Van Wylen, D. G. (2013). Tiered Classrooms at St. Olaf College: Faculty and Student Perceptions of Three Different Designs. *Journal of Learning Spaces, 2*(2).

Walczak, M. M., & Van Wylen, D. G. (2015). Are "New Buliding" Learning Gains Sustainable? Revisiting Our Goals After Five Years. *Journal of College Science Teaching, 44*(6), 17-23.

Watkins, J., & Mazur, E. (2013). Retaining students in science, technology, engineering, and mathematics (STEM) majors. *J Coll Sci Teach, 42*(5), 36-41.

Wilson, Z. S., Holmes, L., Degravelles, K., Sylvain, M. R., Batiste, L., Johnson, M., et al. (2012). Hierarchical Mentoring: A Transformative Strategy for Improving Diversity and Retention in Undergraduate STEM Disciplines. *Journal of Science Education and Technology, 21*(1), 148-156, doi: 10.1007/S10956-011-9292-

Emily Berry Bosak is a data analyst at the National Marrow Donor Program | Be The Match. She received an undergraduate degree in biology from St. Olaf. College with concentrations in statistics, mathematical biology, and film studies. She has also completed graduate courses in statistics at Iowa State University. She is passionate about R programming and creating beautiful and informative data visualizations. Emily loves being able to use her background in both statistics and biology in her work at Be The Match, improving the process for facilitating bone marrow transplants and other cellular therapies to help save lives.

Katherine Henry is a Digital Services Assistant at Northbrook Public Library and a Teen Services Advisor at Arlington Heights Memorial Library. She received her undergraduate degree in Mathematics and Studio Art from St. Olaf College with a concentration in Statistics. From designing interactive displays on ciphers and cryptography and teaching patrons how to use a laser cutter, to creating crochet masterpieces and looking at ways that libraries can use data to show their impact in their communities, Katherine uses her love of art, mathematics, and technology to continue to promote STEAM concepts in the library world.

Kiegan Rice is a Ph.D. Candidate in Statistics at Iowa State University. She received an undergraduate degree in Mathematics and Spanish from St. Olaf College with a concentration in Statistics, and a Master of Science degree in Statistics at Iowa State University. Her research has included identifying reasons for STEM attrition, development of fair mathematics assessments, and the communication of quantitative information to non-quantitative audiences. She is currently developing materials to teach statistical concepts in forensic science to the general public.

Marc Richards is a PhD student in Statistics at the University of Pittsburgh. He received his undergraduate degree from St. Olaf College in Mathematics and Economics. His research interests are in statistical learning with applications in Sports, Medicine, and STEM Education.

Mary M. Walczak (walczak@stolaf.edu) is Edolph A. Larson and Truman E. Anderson, Sr. Professor of Chemistry at St. Olaf College. She received her Ph.D. and completed a postdoctoral fellowship at Iowa State University. She has taught a variety of chemistry courses at different levels and held many leadership roles. She currently serves as the Associate Dean for Natural Sciences and Mathematics. Mary enthusiastically promotes active learning in its many forms and evidence-based academic decision making. Her research program focuses on understanding the barriers to persistence in STEM and how data can be used to help improve outcomes.

Kathryn Ziegler-Graham is Associate Professor of Statistics at St Olaf College. She received her PhD in Biostatistics at the Johns Hopkins Bloomberg School of Public Health. Katie teaches across the statistics curriculum at St Olaf and is an active contributor to the Center for Interdisciplinary Research. Her research ranges from forecasting Public Health challenges, most notably limb loss and Alzheimer's disease, to undergraduate STEM persistence.