Journal of the Scholarship of Teaching and Learning, Vol. 20, No. 1, April 2020, pp.105-129.
doi: 10.14434/josotl.v20i1.24825

Exploring Organic Chemistry I Students' Responses to an Exam Wrappers Intervention

Evan Grandoit
Northwestern University

Rose Bergdoll
Brooklyn College

Edith Rosales
Brooklyn College

David Turbeville
Brooklyn College

Shoshana Mayer
Brooklyn College

Gail Horowitz
Brooklyn College

Abstract: Research has demonstrated that academically successful students are effective, self-regulated learners. Moreover, exam wrapper interventions have been shown to foster the development of self-regulated learning behaviors on the part of college students. In this naturalistic, qualitative, and exploratory study, an exam wrapper intervention was implemented in a key, gatekeeping STEM course at a diverse, public university. Student responses to a series of four exam wrappers were collected and analyzed. Results indicated that while many students were able to look critically at their study behaviors and course performance, these behaviors did not necessarily pay off, especially for weaker students. Notably, transfer and/or non-matriculated students were at greatest risk of withdrawal and failure. However, all students, both weak and strong, showed a lack of attention towards checking their answers and learning from their mistakes. Overall, the exam wrappers provided useful information regarding the self-regulated learning processes of these STEM students.

Keywords: Organic Chemistry, Self-Regulated Learning, STEM, Underrepresented Minorities

Introduction

The loss and attrition of qualified undergraduates from STEM majors is no longer an unfamiliar phenomenon. Unfortunately, it has become a well-known area of study, research, and exploration (Chang, Sharkness, Hurtado, & Newman, 2014; Hunter, 2016; Malcom & Feder, 2016; Seymour & Hewitt, 1997). Research indicates that among the most recent generation of college students, more than half of all students who enter college intending to major in STEM leave STEM or fail to graduate altogether (Eagan, Hurtado, Figueroa, & Hughes, 2014). For certain ethnic minority groups, attrition rates are even higher. Eagan and colleagues (2014) report that 76% of UREM (underrepresented minorities) who enter planning to major in STEM do not complete their degrees within six years.

The loss of women and UREM from STEM majors results in the underrepresentation of these populations in STEM fields (for example in chemical industry, biological research, engineering, etc.) (National Science Foundation, 2013), and also leads to their underrepresentation in vital health science fields such as Medicine and Dentistry (Association of American Medical Colleges, 2010; Smedley, Butler, & Bristow, 2004). Significantly, researchers have found that Organic Chemistry is a key "gatekeeping" course that deters students from moving forward in careers such as medicine (Barr, Matsui, Wanat, & Gonzalez, 2010).

Research indicates that problems with undergraduate science teaching, especially in introductory, so-called gatekeeping courses, are among the primary reasons why students switch out of STEM (Seymour & Hewitt, 1997). Problems cited include large lectures, where teaching is impersonal and primarily in a transmissive mode (Gasiewski, Eagan, Garcia, Hurtado, & Chang, 2012; Hunter, 2016), classroom climates that are competitive, chilly, and unwelcoming (Gasiewski et al., 2012; Rosser, 1997) and course content that is divorced from everyday life (Hunter, 2016).

Reform efforts such as Process Oriented Guided-Inquiry Learning (Moog & Farrell, 2006), Peer-Led Team Learning (Gosser, 2011), Team Based Learning (Michaelsen, Knight, & Fink, 2004), Flipped Classrooms (Talbert, 2017), and Peer Instruction (Watkins & Mazur, 2013) all attempt to create active, student centered learning environments where student cooperation is encouraged. These reforms have shown excellent promise in improving retention rates for all students, including women and UREM (Deslauriers, Schelew, & Wieman, 2011; Eberlein et al., 2008; Freeman et al., 2014; Hall, Curtin-Soydan, & Canelas, 2014; Lewis, 2011; Mooring, Mitchell, & Burrows, 2016). Unfortunately, their implementation has not been widespread for a number of reasons, including cost (e.g. for revamping classrooms, hiring additional instructors, professional development, release time, etc.), ideological resistance (to change in general and/or to student centered pedagogies in particular) (Burd et al, 2016; Chaskes & Anttonen, 2015; Kezar & Holcombe, 2016), deficit views of underrepresented groups (Castro, 2014), and/or simple inertia (Wieman, 2016).

Recently, researchers have begun to focus their attention on first-generation college students, another underrepresented population of students (Engle & Tinto, 2008; Jehangir, 2010; Stephens, Hamedani, & Destin, 2014; Stuber, 2012; Terenzini, Springer, Yaeger, Pascarella, & Nora, 1995; Uche, 2015). First-generation students are defined as either the first-generation in their family to earn a Bachelor's degree or as the first-generation in their family to enroll in college (Davis, 2010). National data show that first-generation college students are disproportionally from groups underrepresented in STEM, including women and UREMs, but are also disproportionally low-income students, non-traditional college age students and students with dependents (Engle & Tinto, 2008; Lohfink & Paulsen, 2005; Terenzini et al., 1995). A study by Saenz, Hurtado, Barreta, Wolf, & Yeung (2007) found that 87% of first-generation college students were students of color.

Disturbingly, data demonstrate that retention rates for first-generation college students in STEM are lower than those of continuing generation college students (Chen, 2013; Doerschuk et al., 2016) especially in key gatekeeping courses like General Biology and General Chemistry (Eddy & Hogan, 2014; Gregg-Jolly et al., 2016; Harackiewicz et al., 2014). Auspiciously however, interventions involving student centered pedagogies have shown excellent promise in combating these trends (Crimmins, 2017; Eddy & Hogan, 2014; Haak, HilleRisLambers, Pitre, & Freeman, 2011).

Background

Self-regulated learning (SRL) has been described as the regulatory processes that learners engage in while monitoring their own learning. These regulatory processes include: setting goals, selecting strategies and behaviors, monitoring one's performance, and adapting in response to feedback

(Zimmerman, 1990). Students who are more effective self-regulated learners are more successful in general (Plant, Ericsson, Hill, & Asberg, 2005; B. Zimmerman & Martinez-Pons, 1990), and in STEM in particular (Karabenick, 2003; Nandagopal & Ericsson, 2012; Szu et al., 2011; VanderStoep, Pintrich, & Fagerlin, 1996). Notably, research demonstrates that one of the key reasons why students struggle in key gatekeeping STEM courses is due to a lack of sophisticated study skills (Bunce et al., 2017; DiBenedetto & Bembenutty, 2013; Sebesta & Bray Speth, 2017).

Nonetheless, self-regulation is a learnable and trainable skill (Phillips, Clemmer, McCallum, & Zachariah, 2017; Weinstein & Acee, 2013). Discipline-specific training in metacognition and/or self-regulation has been shown to be effective in helping students improve in STEM subjects such as Chemistry (Zhao, Wardeska, McGuire, & Cook, 2014), Mathematics (Hudesman et al., 2014; Olszewski, 2016), and Biology (Azevedo & Cromley, 2004; Bernacki, Vosicka & Utz & 2017).

Exam wrappers (designed by Dr. Marsha Lovett of Carnegie Mellon University) are a specific pedagogical tool designed to encourage students to cultivate and improve their self-regulated learning skills (Lovett, 2013). Exam wrappers are typically written exercises that students complete (inside or outside of class) after receiving back a graded exam. These exercises encourage students to monitor their performance and to set goals and modify their behaviors because they invite students to reflect on their study behaviors and exam performance, to consider what strategies were or were not successful, and to state what behaviors they might want to initiate in preparing for the next exam (Lovett, 2013). Exam wrappers are sometimes also administered in a pre-test fashion to encourage students to reflect as they prepare for an upcoming exam. Students may be asked to assess their level of preparedness, to analyze the study strategies they have been employing and/or to consider making changes to their study strategies (Lovett, 2013).

Exam wrappers have been used in a variety of classroom settings (e.g. Achacoso, 2005; Soicher & Gurung, 2017; Thompson, 2012), but their use has been most frequently reported in STEM classrooms. Experimental studies have shown statistically significant improvements in student performance when exam wrapper assignments have been incorporated into Statistics and Engineering classrooms (Chen, Chaves, Ong, & Gunderson, 2017; Chew, Chen, Rieken, Turpin, & Sheppard, 2016). Similarly, studies in Biology and Physics classrooms have shown that when students are incentivized to correct their mistakes on exams these students show statistically significant improvements in their course performance (Brown, Mason, & Singh, 2016; Rozell, Johnson, Sexten, & Rhodes, 2017). Correlational or descriptive studies involving exam wrappers in Biology classrooms have shown weaker students have more limited SRL strategies and that these students are less able to implement changes in their SRL strategies (Stanton, Neider, Gallegos, & Clark, 2015).

In the naturalistic, qualitative study described below, exam wrappers were utilized in an exploratory fashion to investigate whether, with minimal intrusion onto the flow of a fairly traditional Organic Chemistry I lecture course, exam wrappers could encourage students to reflect upon their own learning and encourage them to make strategic and effective changes in study habits and behaviors. (We also conducted a separate, quasi-experimental intervention study utilizing exam wrappers in a different Organic Chemistry I course and hope to report on this study in the near future.)

Organic Chemistry I was chosen as the subject area for this study, in part, because of the crucial gatekeeping function that it plays for many STEM and pre-health science students. (It is also a course that has particularly low success rates at our institution.) Additionally, prior research at our institution (AUTHOR AND COLLEAGUES, 2013) had shown that students who were at risk for not passing Organic Chemistry I could overcome the odds and pass if they engaged in a particular self-regulated learning behavior: help-seeking -- specifically the making use of resources such as office hours and supplementary problem sessions. We hypothesized therefore, that if Organic

Chemistry students at our institution, at which there are large numbers of first-generation college students (2014 UNIVERSITY NAME STUDENT EXPERIENCE SURVEY), could be encouraged to be more self-reflective about their study habits, they might improve in their self-regulated learning abilities, modify their study behaviors and improve in their overall course performance. Thus, we chose to introduce exam wrappers into the Organic Chemistry I course because of their potential to impact the reflection and self-regulated learning behaviors of our students. We also chose exam wrappers as our pedagogical tool because they are a reform method that is fairly conventional, non-controversial, and easy to implement.

Method

Participants and Course Context

This IRB-approved study was conducted at a large, urban, public university located in the Northeastern United States. The undergraduate population at this institution is highly diverse. Approximately 40% of students are from ethnic groups underrepresented in STEM, about 50% are low income, about 30% are first-generation college students (neither parent has any college education) and approximately 40% speak English as a second or third language (2014 COLLEGE NAME STUDENT PROFILE; 2014 UNIVERSITY NAME STUDENT EXPERIENCE SURVEY).

This study was conducted in an Organic Chemistry I classroom in the spring of 2017. Including students who withdrew, the course enrolled a total of 176 students. Students met in a single classroom for lecture for a 75-minute class period twice a week (lectures were taught by AUTHOR). Additionally, students met in one of six smaller groups (of approximately 30 students each) for one 50-minute recitation period per week.

Student learning in the course was assessed via a combination of quizzes and exams. Five quizzes were administered over the course of the semester in recitation. Quizzes took approximately 20 minutes to complete and were standardized across all sections. Students also completed two 75-minute midterm exams and one 120-minute final exam (exams were administered in two large classrooms). Five percent of students' course grades were allotted towards completion of four assignments that were termed self-assessments. These self-assessments were actually two pre-exam and two post-exam exam wrappers. Students' course grades were determined using a mastery-based scheme, rather than a norm-referenced scheme (Popham, 1971).

Exam Wrappers

Exam wrapper 1 (see Appendix 1) was completed in recitation immediately before students took their first quiz (at approximately week three of the semester). Exam wrapper 2 (identical in content to exam wrapper 1) was administered similarly in that it was handed out and completed in recitation immediately before students took their second quiz (at approximately week four of the semester). One hundred sixty-six students (94%) completed exam wrapper 1 and one hundred sixty-three students (93%) completed exam wrapper 2.

Exam wrapper 3 (see Appendix 2) was completed online using the course management system available through the university. Exam wrapper 3 was only made available to students during week five of the semester (before the first midterm). Exam wrapper 4 (identical to exam wrapper 3) was similarly completed online and only available during week seven of the semester (in between the first and second midterm examinations). One hundred eleven students (63%) completed exam

wrapper 3 and one hundred thirty-seven students (78%) completed exam wrapper 4. (Figure 1 shows the timing and sequence of the exam wrappers).

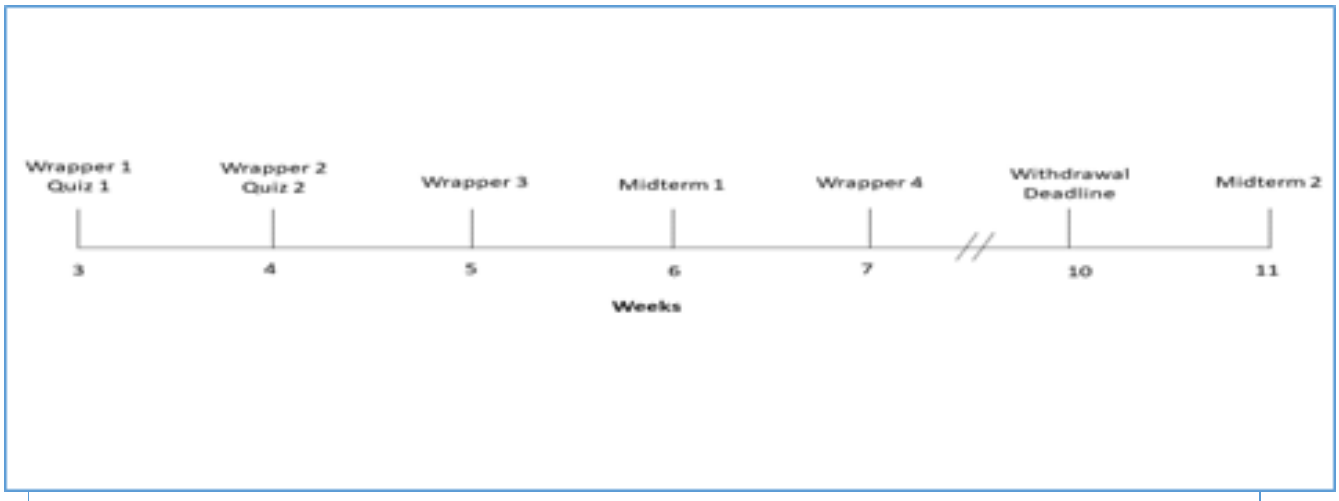


Figure 1. Timing of quizzes, midterm exams and exam wrappers

Exam wrappers 1 & 2 asked students to indicate how they felt about their competence in two areas: conceptual understanding and problem-solving ability.[1] These questions were intended to trigger self-reflection on the part of students (Zimmerman, 2002) and make them consider their degree of preparedness (Yuen-Reed & Reed, 2015). Students selected from among three choices (*strong*, *ok*, and *weak*) in two Likert-style questions. Additionally, students were asked to choose the extent to which they felt they could easily access help with the course when they needed it. This question was intended to lead students to consider whether or not they were seeking out and utilizing the resources available to them. Students selected from among three choices (*agree*, *unsure* and *disagree*) in a Likert-style question. Exam wrappers 1 and 2 were administered to students immediately before they took their first two quizzes (respectively) so that students who were underprepared but unaware of it would hopefully be hit with some dissonance if they put their expectations down on paper and then either struggled during the quiz or found out they were not in good shape when they received their quiz score.

Exam wrappers 3 and 4 asked students to report the extent to which they felt satisfied with their performance on their most recent assessment (quiz or midterm)[2]. Students selected from among five choices (*strongly agree, agree, neither agree nor disagree, disagree,* or *strongly disagree*) in a Likert-style question. Students were also asked to report how many hours per week they devoted to

[1] In the course syllabus, students were assigned readings (focused on conceptual understanding) and homework assignments (focused on problem solving) for each individual class period. The distinction between these two separate ideas or skills was also discussed in class. Thus, it was expected that students would understand and be able to differentiate between the terms conceptual understanding and problem solving ability.

[2] Typically within 24 hours of taking an exam, students' exam scores, along with a copy of the answer key, were posted online. However, typically they did not receive their graded exam papers back until their next recitation period. Quizzes took up to two weeks to be graded and were also returned to students in recitation. Quiz grades were also posted online. However, answer keys for quizzes were not provided.

studying and how many hours they spent problem-solving (with boxes provided for numerical responses.) Lastly, students were asked what they would do differently in the future to improve their course performance. (For this question, students wrote narrative responses of any length they wished and were able include as many study strategies as they wished.)

The first three questions asked in exam wrappers 3 and 4 were designed to encourage students to directly confront their feelings about how they were doing in the course, to determine how much time they were putting in overall, and to consider whether or not they were devoting sufficient time to problem-solving. The final question was intended to encourage students to self-reflect regarding their study behaviors, to consider what they ought to do differently, to set goals of changing their behaviors and to hopefully commit to those changes by asking them to put in writing what they intended to do differently from then on (Zimmerman, 2002).

In this study, the exam wrappers were utilized both as an intervention, intending to trigger students to reflect and possibly change their behavior, and also as a source of data, as a means for the researchers to examine and learn from the reflections and reported behaviors of students. Thus, with IRB permission, after the course was over the responses that students gave to the exam wrapper questions were collected, stripped of identifying information and utilized as data.

Course Syllabus with SRL Supplement

One significant modification was made to the course syllabus to reinforce the idea that students reflecting upon their own learning was an iterative process and that making decisions, being strategic, utilizing resources, and seeking help were all important aspects of being successful in the course. This modification took the form of an extra handout entitled: "How to Study for this Course". It contained a graphic display of a recommended iterative process for a student to follow and two subsections of guidance or advice entitled, "Ways to Assess Yourself" and "Help Seeking Guide". (See Appendix 3.) This handout was in addition to the usual guidance and information provided in the syllabus regarding how to be successful (e.g. spend 75% of your study time doing problem solving, come to class, get help right away if you get stuck, etc.) (See Appendix 4.) During the first day of lecture, AUTHOR reviewed the syllabus with students and explained this handout.

Website with SRL Emphasis

One factor relevant to this study is that AUTHOR maintains an extensive webpage (URL) devoted to providing students with additional resources for the Organic I course. The majority of these resources are supplementary problem sets. This is not unique as other Organic Chemistry instructors at her institution and elsewhere (Cortes, 2017; Reusch, 2017) also provide supplementary problem sets to students through webpages and/or course management software. However, the particular presentation of AUTHOR'S Organic I homepage gives a strong visual impression to students that problem solving is an extremely important aspect of the course. Furthermore, the homepage is organized such that students can quickly find either introductory or advanced practice problems (with answer keys) corresponding to each topic they learn about in lecture. Thus, students who utilize the website are encouraged to think about what topics they need to work on and are guided towards thinking about practicing in a strategic way where they ideally try to master simpler problems before they move to more challenging ones.

Another resource provided through the Organic I homepage is links to old exams (with answer keys) written by AUTHOR. The old exams are provided as additional practice problems for students. But they also illustrate what the format of exams will look like, what the level of difficulty

of exams will be, and in what ways they will be expected to answer questions that require them to think across multiple chapters from their textbook while taking the exam.

One aspect of the old exams found on the website is that next to each exam question is listed the number of minutes a student should allocate toward completing it. This is a feature of AUTHOR'S exams that she includes to help students learn how to effectively pace themselves during exams, a skill that she finds is particularly challenging for many students at her institution.

Overall, there are a number of features in AUTHOR'S website that attempt to encourage students to strengthen and improve their self-regulation. However, these features have been in place on her website in a consistent fashion for a number of years. Therefore, for the purposes of this study, no modifications were made to the website.

Course Performance

Student scores on all assessments (quizzes, midterm examinations, and the final) were collected, stripped of identifying information, and utilized as data.

Prior Performance in Chemistry

Previously, AUTHOR and her colleagues (2013) had established that for students at her institution, the letter grades that they received in General Chemistry II (the immediate prerequisite course for Organic Chemistry I) were a good predictor of performance in Organic Chemistry I. Specifically if General Chemistry II was taken at our institution, 49% of the variability in performance in Organic Chemistry I was explained by performance in General Chemistry II. Therefore, for the purposes of the current study, we obtained transcripts from all students and collected General Chemistry II letter grades for students who completed General Chemistry II at our institution. (Letter grades for students who had taken General Chemistry II outside of our institution were not collected.) These letter grades were combined with the exam wrapper and course performance data described above and then all identifying information was removed.

Data Analysis

Exam wrappers 1 & 2. On exam wrappers 1 and 2, students were prompted to report the degree to which they felt competent (*strong, ok, weak*) in two areas: conceptual understanding and problem-solving ability. Students' self-perceptions were then compared to their actual quiz scores and then classified as either accurate, overestimates, or underestimates (see table 1).

A self-rating of *weak* on exam wrapper 1 was determined to be an underestimation if the student's quiz grade was above 45% and accurate if the student's grade was less than or equal to 45%. (An exam average of 45 was the approximate cutoff for passing the course.) A self-rating of *ok* was judged to be an overestimation if the student's grade was less than 46% and an underestimation if their grade was 80% or higher (and accurate if the grade was 46-79%). A self-rating of *strong* was gauged as an overestimation if the student's grade was less than 80% and deemed accurate if their grade was equal to or greater than 80%.

Because the subject matter of quiz 2 was more difficult than that of quiz 1 (and the student average on quiz 2 was 10 points lower than on quiz 1), more lenient criteria were utilized to define what was considered an overestimate, an underestimate or an accurate self-assessment for quiz 2. A self-rating of *weak* on exam wrapper 2 was gauged as an underestimation if the student's grade was above 35% and accurate if the student's grade was less than or equal to 35%. A self-rating of *ok* was categorized as an overestimation if the student's grade was less than 36% and an underestimation if their grade was 70% or higher (and accurate if the grade was 36-69%).

A self-rating of strong was gauged as an overestimation if the student's grade was less than 70% and deemed accurate if their grade was equal to or greater than 70%.

Table 1. Gauging students' accuracy in their assessments of their confidence in conceptual knowledge and problem-solving ability

		Overestimate	Accurate	Underestimate
Quiz 1				
	Weak	NA	≤ 45 %	> 45 %
	Ok	≤ 45 %	46-79 %	≥ 80%
	Strong	< 80%	≥ 80%	NA
Quiz 2				
	Weak	NA	≤ 35 %	> 35 %
	Ok	≤ 35 %	36-69 %	≥ 70%
	Strong	< 70%	≥ 70%	NA

Exam wrappers 3 & 4: Study time and problem-solving time. Questions on exam wrappers 3 and 4 prompted students to reflect on the number of hours they spent studying as well as the amount of time they spent practicing problems. The sample size, mean, standard deviation, and range of these reported hours were calculated. Outliers, participants whose reported hours were more than one standard deviation from the mean, were highlighted for possible further analysis. Differences between students' reported hours at the time of exam wrapper 3 and exam wrapper 4 were calculated.

Exam wrappers 3 & 4: Future plans. The final question of exam wrappers 3 and 4 asked students to describe what they would change or do differently in their future studying. Student responses were text-based and averaged approximately 30 words in length. Outliers, participants with word counts more than one standard deviation from the mean, were highlighted for possible further analysis.

Student written responses (to the question regarding what they planned to change in their study habits) were coded according to the following procedure. Four coders independently reviewed four different subsets of the student responses. (In total, one third of the student entries were reviewed.) In keeping with the explanatory nature of the study, the coders did not approach coding with preconceived or a-priori ideas of what the codes should be. Rather, we allowed themes to emerge from the data. Preliminary lists of themes and categories observed in the data were generated independently by each of the four coders. The coders then met and compared their preliminary lists. The preliminary lists were organized and collapsed into four codes, each of which contained a number of sub-codes. (See table 14 for a listing of the codes and sub-codes utilized in this study.) The four codes which emerged from the data were *Study Behaviors, Strategic Behaviors & Decisions, Help Seeking* and *What's Going on With Me.*

The code *Study Behaviors* referred to the many types of behaviors that students stated they planned to engage in, for example reading the textbook, working on practice problems, or reviewing their lecture notes. Each specific *Study Behavior* described by a student was given a separate sub-code. For example, the behaviors just described were assigned the sub-codes *Textbook, Problem Solve* and

Lecture Notes, respectively. Study behaviors were coded as *neutral, plus,* or *minus* depending if the student said he or she would *do it*, would spend *more time* doing it or would spend *less time* doing it, respectively.

The code *Strategic Behaviors & Decisions* was created to capture behaviors that students described which could best be characterized as being strategic about their studying. For example, choosing to do a little bit of problem solving every day, rather than saving it all for the weekend or choosing to read the textbook before lecture rather than after, were both coded under the code *Strategic Behaviors and Decisions* and under the sub-code *Timing Specific*, whereas choosing to work on advanced problems rather than simple problems was also coded under *Strategic Behaviors and Decisions,* but under the sub-code of *Which Problems*.

The code *Help Seeking* referred to the different ways in which students described how they would try to get help in the course. Sub-codes were created for the seeking of human help, *HHelp* – as in help from an instructor, TA, or tutor, electronic help, *EHelp* – as in help from an online resource like a tutorial or video, or help from a physical resource, *PHelp* - like a review book or molecular model set.

A code or category was created called *What's Going on With Me* to capture descriptions that students gave that did not fit under categories of study strategies or behaviors, but rather described emotional or psychological states. For example, a few of the sub-codes in this category were *Anxious, Unsure, Confident, Careless* and *Overwhelmed*.

A code book listing all the codes, sub-codes and their definitions was created. After this, the data entries were divided up into three equal portions. Each third of the data was coded independently by two coders (Six coders in total were utilized.) Afterwards, all six coders met as a group and went through each data entry one by one comparing the two sets of codes from the pair of coders against one another and against each data entry. Together, the group of six coders came to agreement on what the most complete and accurate codes should be for each data entry. Often the consensus or *agreed upon codes* matched the original codes assigned by the two coders, but occasionally errors or oversights were caught through this process. Therefore, it was determined that this method of meeting as a group of six and going through each data entry one by one was useful as it allowed for the most thorough, complete, and detailed analysis, without resulting in the coders reaching consensus prematurely. Occasionally as a result of this process, a few new sub-codes arose and had to be defined and created, and a few clarifications or refinements of existing codes had to be made. The code book was updated accordingly and modifications to the already coded data were made. After a complete listing of all the sub codes for each data entry was compiled, tallies were taken to determine the number of times each sub code was cited.

Exam 2 - 1. Student scores on exam 1 were subtracted from the scores on exam 2. Participants were then listed into four categories based on that difference: improved, worsened, no change, or NA (did not take exam 2).

Prior performance in chemistry. Students were coded as *at risk* for not succeeding in Organic Chemistry I if they had scored a grade of C plus or lower in General Chemistry II at our institution and as *not at risk* if they had scored a B minus or higher. Students who had not taken General Chemistry II at our institution (non-matriculated and transfer students) were coded as *unknown risk*.

Course performance. Students were grouped into categories based on their performance in the Organic course. For students entering the Organic course *not at risk*, satisfactory performance was

defined as completing Organic with a grade of B minus or above.[3] For students entering the course *at risk* or with *unknown risk,* satisfactory performance was defined as completing the course with a C minus or above. Students who did not meet these criteria were categorized having unsatisfactory performance. (Students who withdrew from the course were categorized separately.)

Student categories. Nine categories of students were differentiated based on a) the level of risk of the students entering the course and b) their actual performance in the course (see table 12). These categories were analyzed and compared across areas of interest, such as confidence in conceptual understanding and problem-solving ability, accuracy of self-ratings compared to subsequent quiz performance, hours reported studying and practicing problems, and planned changes to study behaviors.

Overall Results

Exam Wrappers 1 & 2

Students were asked to describe their confidence in their conceptual understanding and problem-solving ability. In both exam wrappers 1 and 2 (see table 2), approximately 70% of students reported feeling *ok* about their understanding and ability. Additionally, the percentages of students who characterized their *conceptual understandings* as *weak* or *strong* changed only minimally from exam wrapper 1 to 2. However, with regards to confidence in *problem solving ability*, there was a notable increase from exam wrapper 1 to 2 in the percentage of students who felt they were *weak,* as well as a sizeable decrease in students who felt they were *strong*.

Table 2. Students' confidence in their conceptual understanding & problem-solving ability

	Wrapper 1				Wrapper 2			
	Concept		Problem		Concept		Problem	
	N	%	n	%	n	%	N	%
Weak	5	3	16	10	9	5	28	17
Ok	123	74	116	70	120	74	117	72
Strong	38	23	34	21	34	21	18	11

When comparing students' self-assessments to their actual quiz performances (see table 3), student accuracy was low (not on target), ranging from approximately thirty to forty percent. Student accuracy also decreased somewhat from exam wrapper 1 to 2. Furthermore, weak students (who scored 45 or below on quiz 1, or 35 or below on quiz 2) were highly likely to overestimate their abilities, while strong students (who scored 80 or above on quiz 1 or 70 or above on quiz 2) were highly likely to underestimate their abilities.

[3] We defined satisfactory as B- or better for *not at risk* students because we make the assumption that a C+ or worse will hinder future progress for these students (Hrabowski, 2016).

Table 3. Students' accuracy re their conceptual understanding & problem-solving ability

	Wrapper 1				Wrapper 2			
	Concept		Problem		Concept		Problem	
	n	%	n	%	n	%	n	%
Total Students who Are Accurate	67	40%	56	34%	51	31%	48	29%
Total Students who Overestimate	23	14%	25	15%	34	21%	24	15%
Total Students who Underestimate	76	46%	84	51%	78	48%	91	56%
Weak Students who Overestimate	12	100%	10	83%	26	93%	19	68%
Strong Students who Underestimate	74	73%	79	78%	72	71%	84	83%

The final question on exam wrappers 1 and 2 surveyed students' feelings about how easily they felt they could obtain help with the course material when needed (see table 4). Students were prompted to select a response of either *agree, unsure,* or *disagree* from a 3-point Likert-type scale. The most frequently selected response for both exam wrappers 3 (80%, n=166) and 4 (79%, n=163) was *agree*. Only 3% of participants chose *disagree* as their response to this question.

Table 4. Student responses to the statement that they are easily able to obtain help

	Wrapper 1		Wrapper 2	
	Help		Help	
	n	%	N	%
Agree	133	80	128	79
Unsure	28	17	30	18
Disagree	5	3	5	3

Exam Wrappers 3 & 4

In exam wrapper 3, student reports of satisfaction with their course performance (see table 5) spread in a bell-shaped distribution, with the majority of students reporting a neutral, mildly positive, or mildly negative attitude. (At this point in the semester, students had only received grades back on two quizzes, which had a combined average of 72 and counted only as 5-10% of their final course average.) However, by the time of exam wrapper 4, there was a large shift in student satisfaction with nearly 70% of students reporting dissatisfaction with their course performance. (Students filled out exam wrapper 4 shortly after receiving back their scores on exam 1 which had an average of 57% and counted as 20% of their final course average.)

Table 5. Student agreement that they are satisfied with their course performance

	Wrapper 3		Wrapper 4	
	n	%	n	%
Strongly Agree	16	14	2	1
Agree	28	25	20	15
Neither Agree nor Disagree	24	22	21	15
Disagree	27	24	40	29
Strongly Disagree	16	14.4	54	39

Table 6 shows student responses on exam wrappers 3 and 4, indicating number of hours spent studying and practicing problems. The changes in reported study and problem-solving time from wrapper 3 to 4 showed that on average, students only increased their study time by 0.2 hours and their problem-solving time by 0.5 hours (see table 7).

Table 6. Student reported study times and problem-solving times

	n	M	SD	Highest	Lowest
Wrapper 3					
Study Time (Hours)	111	8.2	3.5	25	1
Problem Solve Time (Hours)	111	5.7	3.5	25	1
Wrapper 4					
Study Time (Hours)	137	8.4	4.7	35	2
Problem Solve Time (Hours)	136	6.2	3.9	21	1

Table 7. Increase in reported study hours from exam wrapper 3 to 4?

	n	M	SD	Highest	Lowest
Study Time (Hours)	96	0.2	3.28	10	-11
Problem Solve Time (Hours)	96	0.5	3.13	14	-10

The length of participant responses to the question "What are you going to do differently from now on?" are reported in table 8. The average response was about 30 words in length. However, 21% and 23% of respondents (in wrappers 3 and 4, respectively) had word counts of less than 10 words.

Table 8. Length of student responses - what are you going to do differently?

	M	SD	Lowest	Highest
Wrapper 3				
Word Count	29	28	2	197
Wrapper 4				
Word Count	32	40	2	248

Because exam wrappers 3 and 4 asked students to report their overall study times and problem-solving times, students who reported an intention (on wrapper 3) to change their behavior by increasing their study time or problem-solving time were checked to see if they followed through on their intentions. Only 22% of students indicated that they would increase their overall study time and only about half of these students fulfilled their intention. Fifty percent of students indicated that they would increase their problem-solving time. Similarly, only about half of those students followed through on their intention (see table 9).

Table 9. Evidence (wrapper 4) of follow through of intended study plans (from wrapper 3)?

		n	%
Study Time			
	Did Follow Through	12	11
	Did Not Follow Through	11	11
	Not Applicable	82	78
Problem Solve Time			
	Did Follow Through	26	25
	Did Not Follow Through	26	25
	Not Applicable	49	47
	Unclear	3	4

Exam 2-Exam 1

Student performance on exams 1 and 2[4] was compared and differences were calculated (see table 10). Approximately half improved their scores and one-third worsened. An additional 15% did not take exam 2. (All students who did not take exam 2 also did not take the final exam.)

[4] Exam 2 covered more advanced and more challenging material than exam 1.

Table 10. Performance change from exam 1 to exam 2

	n	%
Improved	91	52
Worsened	57	32
No Change	2	1
NA (didn't take exam 2)	26	15

Prior Performance in Chemistry

Based on their prior performance in General Chemistry II, students were grouped into three categories. Categories indicated whether or not they were at risk of not succeeding in the Organic course. Each category contained approximately one-third of all students (see table 11).

Table 11. Students' at risk of not succeeding in Organic

	n	%
At Risk	54	31
Not At Risk	56	32
Unknown Risk	66	37

Course Performance

Overall, 63% of students performed satisfactorily in the Organic course, 29% performed unsatisfactorily, and 8% withdrew. Students of *unknown risk* were least likely to perform satisfactorily with only 55% of them successful, while 74% of *not at risk* and 71% of *at-risk* students were successful. (See table 12.)

Student Categories

Students were grouped into nine categories based on their risk when entering the course and their satisfactory performance in the course or lack thereof (see table 12). Seven of these nine categories were subjected to further analysis. Two categories were excluded (at risk & withdrawal, not at risk & withdrawal) because they each comprised only 1% of the student population.

	n	%

Table 12. Student categories*

	n	%
At Risk & Satisfactory	36	20
At Risk & Unsatisfactory	14	8
Not At Risk & Satisfactory	41	23
Not At Risk & Unsatisfactory	14	8
Unknown Risk & Satisfactory	31	18
Unknown Risk & Unsatisfactory	23	13
Unknown Risk & Withdrawal	11	6
At Risk & Withdrawal	2	1
Not At Risk & Withdrawal	2	1

***** Two students, with grades of incomplete, were excluded and were not categorized.

Results for Student Categories

Exam Performance

Despite the fact that the second exam covered more challenging material than the first exam, the majority of *satisfactorily performing* students improved their performance from exam 1 to exam 2.

Estimation of Abilities

While results consistent with a Kruger-Dunning effect (1999) were observed for the overall population (see table 3), unsatisfactorily performing students, overall, did not tend to overestimate their abilities (see table 13). Satisfactorily performing students, however, were found to underestimate their abilities ($c^2(1) = 26.0867$, $p < 0.001$), while unknown risk students who withdrew were somewhat likely to overestimate their abilities ($c^2(1) = 16.264$, $p < 0.000055$, see table 13).

Reported Study Times

Satisfactorily performing students did not necessarily put in more study time or problem-solving time than unsatisfactorily performing ones. Furthermore, when unsatisfactorily performing students did increase their problem-solving time, this increase did not result in success. However, not at-risk students were somewhat more likely to indicate an intention (at wrapper 3) to increase the amount of time they were going to devote to problem solving. (See table 13.)

Exam Wrappers 3 & 4 and Unknown Risk Students

Unknown risk students differed from the other categories of students in a number of ways. Unknown risk, unsatisfactorily performing students was the only category to decrease their rate of completion of the exam wrappers from wrapper 3 to wrapper 4. At the time of wrapper 3, they were the most dissatisfied with their course performance as compared to the other categories. They reported the lowest average hours spent studying and doing practice problems. They also had the highest percentage of students who did not take either exam 2 nor the final exam.

The unknown risk, withdrawal students had the lowest completion rate of exam wrappers 3 and 4. Only three students (27%) completed wrappers 3 and 4. These students reported the greatest decrease in time spent studying and doing practice problems from wrapper 3 to 4, yet the greatest number of hours spent studying at the time of wrapper 3.

Planned Study Behaviors

Problem solving plans. At the time of wrapper 3, large numbers of all students in all categories indicated that they intended to devote more time to problem solving. However, by wrapper 4, almost none of the not at risk, satisfactorily performing students indicated that they needed to devote *additional* time to problem solving. Yet, over 70% of the not at risk, unsatisfactorily performing respondents reported that they *still* intended (and needed) to devote additional time to problem solving (see table 14).

Behaviors not reported. Plans to adopt behaviors such as joining a study group, attending office hours, reviewing lecture notes, checking one's answers against a key, or learning from one's mistakes were rarely (or never) reported by students of any category (see Appendix 1).

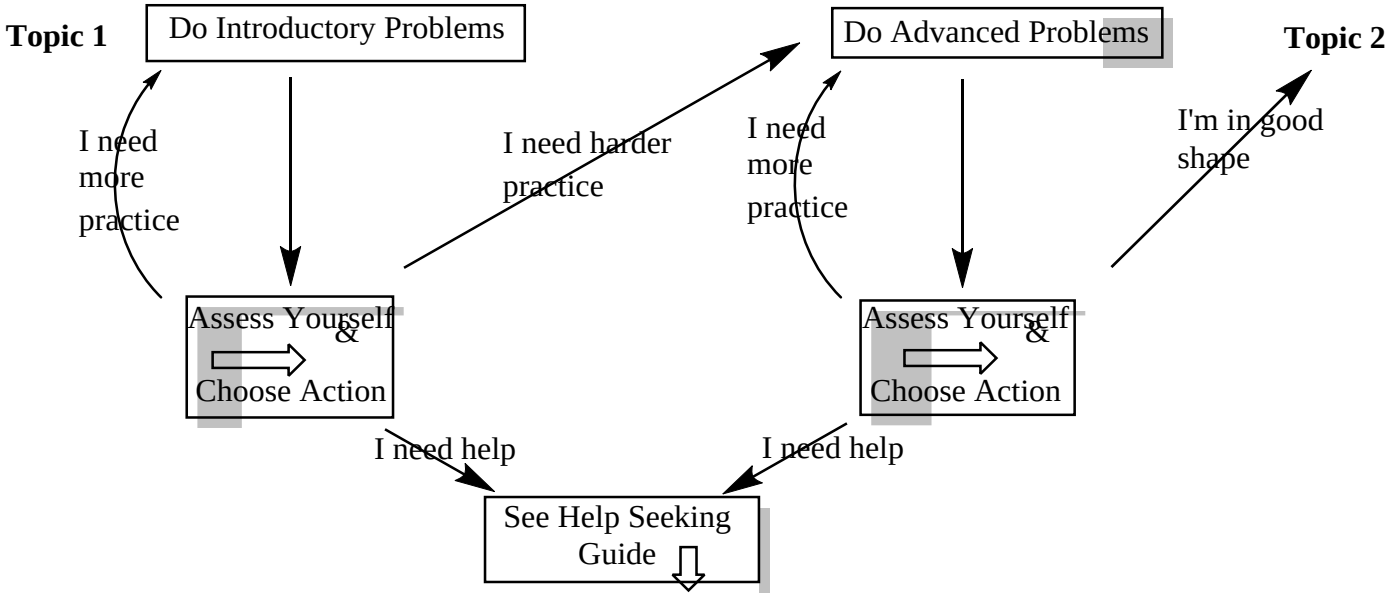
Appendix

Appendix 1: Comparison of Self-Assessments, Study Times, Follow Thru, Exam Performance Across the 7 Categories

	At Risk Satisfactory	At Risk Unsatis-factory	Not at Risk Satisfactory	Not at Risk Un-satisfactory	Unknown Risk Satisfactory	Unknown Risk Un-satisfactory	Unknown Risk Withdrawal
n in Category	36	14	41	14	31	23	11
% of Overall Population	20%	8%	23%	8%	18%	13%	6%
Self-Assessments of Conceptual & Problem Solving Abilities							
% Who were Accurate	64%	38%	31%	23%	41%	36%	42%
% Who Underestimate	7%	31%	66%	50%	51%	22%	11%
% Who Overestimate	29%	31%	2%	12%	8%	42%	47%
% Who Completed Wrapper 3	64%	57%	73%	64%	61%	57%	27%
% Who Completed Wrapper 4	86%	71%	95%	100%	81%	43%	27%
% Dissatisfied at Wrapper 3	30%	38%	23%	22%	53%	77%	33%
% Dissatisfied at Wrapper 4	71%	90%	63%	64%	52%	90%	100%
Mean Study Time (Hours/Wk)							
Reported in Wrapper 3	7.2	9.8	9.3	7.8	7.7	6.8	10.5
Reported in Wrapper 4	8.2	9.3	9.4	8.1	8.1	6.9	8.5
Wrapper 4 – 3	+1.0	-0.6	+0.1	+0.3	+0.4	+0.1	-2.0
Mean ProblemSolveTime (Hrs/Wk)							
Reported in Wrapper 3	5.0	7.8	6.1	5.9	4.9	4.7	8.7
Reported in Wrapper 4	5.6	8.0	6.3	6.3	6.1	5.3	5.5
Wrapper 4 – 3	+0.6	+1.0	+0.2	+0.4	+1.1	+0.6	-3.2
Average Word Count Wrapper 3	19	32	32	26	34	26	36

Average Word Count Wrapper 4	24	23	49	27	28	29	45
% of Category Who Planned to							
Increase Study Time	3%	29%	15%	7%	16%	13%	9%
Increase Problem Solve Time	28%	29%	41%	43%	29%	17%	9%
% of *Category* Who Followed Thru							
Increased Study Time	3%	14%	5%	7%	6%	4%	9%
Increased Problem Solve Time	8%	21%	19%	21%	23%	4%	9%
% Who Improve from Exam 1-2	67%	18%	90%	31%	68%	13%	0%
% Who Didn't Take Exam2 & Final	0%	14%	0%	7%	0%	30%	NA

Appendix 2: How to Study for Organic Chemistry



Ways to Assess Yourself:

- Check your work against an answer key.
- Try explaining your answers out loud to a friend or classmate.
- Try to convince a friend or classmate that your point of view is right (and theirs is wrong).
- Take the assigned question and modify it slightly but in a way that you think will matter and see if you can still answer it.

Help Seeking Guide:

1. Are you feeling lost about a specific topic or topics?
   a. A good source for simple explanations of Organic Chemistry concepts is Organic Chemistry as a Second Language by Klein.
   b. Sometimes a short, online video tutorial (no more than 10 minutes!) can help. See URLs listed above.
   c. Don't waste hours searching for or watching videos. If you don't find what you need quickly, ask a classmate or the course instructor for suggestions.
2. Are you getting some of the content, but missing bits and pieces, like parts of the HW you get right, parts you get wrong? Or you are not always sure why you get things right or wrong? In these kinds of situations (where you need quick, short explanations), it can be very helpful to
   a. Go to the tutors in the learning center.
   b. Ask questions of your recitation instructor, e.g. before or after class or during office hours.
   c. Ask questions of the lecture instructor, e.g. before or after class or during office hours.
   d. Ask questions of your lab instructor during quieter times in the lab (when there are waiting periods or when lab ends early).
3. Are you feeling completely lost in the course?
   This is a time to see someone like the course instructor (who is an expert at helping students succeed in Organic Chemistry) or another mentor/advisor that you know well and trust.

Appendix 3: How to Succeed In Organic Chemistry

How to Succeed In Organic Chemistry:

1. Set aside 10 hours per week of study time for this course.
2. Skim the textbook before class.
3. Attend class religiously and come on time to class.
4. Take notes in lecture.
5. Review your lecture notes as soon as you can after each class meeting.
6. Think of this class like a math class - problem solving is the most important thing.
   Spend the majority (at least 75%) of your study time doing problems, not reading!
   Do all the assigned homework problems (textbook and internet).
   Practice each topic until you have mastered it. Don't just stop when you have completed the assigned problems. Make sure you really understand what you are doing.
7. Study with a partner or in a group.
8. Don't be afraid to ask for help. Get help immediately if you get stuck.

References

Achacoso, M. V. (2004). Post-test analysis: A tool for developing students' metacognitive awareness and self-regulation. *New Directions for Teaching and Learning, 100*, 115-119.

Association of American Medical Colleges. (2010). *Diversity in the physician workforce: Facts and figures 2010*. Washington, D.C.: Association of American Medical Colleges.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*(3), 523-535. http://psycnet.apa.org/doi/10.1037/0022-0663.96.3.523

Barr, D. A., Matsui, J., Wanat, S. F., & Gonzalez, M. E. (2010). Chemistry courses as the turning point for premedical students. *Advances in Health Sciences Education, 15(1)*, 45-54. https://doi.org/10.1007/s10459-009-9165-3

Bell, P., & Volckmann, D. (2011). Knowledge surveys in General Chemistry: Confidence, overconfidence, and performance. *Journal of Chemical Education, 88*, 1469-1476. https://pubs.acs.org/doi/pdf/10.1021/ed100328c

Bernacki, M. L., Vosicka, L., & Utz, J. (2017). *Web-delivered training to improve learning and achievement for underrepresented and first-generation STEM learners*. Paper presented at the Annual Conference of the American Educational Research Association, San Antonio, Texas.

Bragg, D. D. (2012). Two-year college mathematics and student progression in STEM programs of study. In S. Olson & J. B. Labov (Eds.), *Community colleges in the evolving STEM education landscape* (pp. 80–101). Washington, DC: The National Academies Press.

Brown, B. R., Mason, A., & Singh, C. (2016). Improving performance in quantum mechanics with explicit incentives to correct mistakes. *Physical Review Physics Education Research, 12*(1), 1-20. https://doi.org/10.1103/PhysRevPhysEducRes.12.010121

Bunce, D. M., Komperda, R., Schroeder, M. J., Dillner, D. K., Lin, S., Teichert, M. A., & Hartman, J. R. (2017). Differential use of study approaches by students of different achievement levels. *Journal of Chemical Education, 94*(10), 1415-1424. https://pubs.acs.org/doi/pdf/10.1021/acs.jchemed.7b00202

Burd, G. D., Tomanek, D., Blowers, P., Bolger, M., Cox, J., Elfring, L., . . . Wallace, C. (2016). Developing faculty cultures for evidence-based teaching practices in STEM: A program report. In G. C. Weaver, W. D. Burgess, A. L. Childress, & L. Slakey (Eds.), *Transforming*

institutions: Undergraduate STEM education for the 21st century (pp. 90-102). West Lafayette, Indiana: Purdue University Press.

Castro, E. L. (2014). "Underprepared" and "at-risk": Disrupting deficit discourses in undergraduate STEM recruitment and retention programming. *Journal of Student Affairs Research and Practice, 51*(4), 407-419. https://doi.org/10.1515/jsarp-2014-0041

Chang, M. J., Sharkness, J., Hurtado, S., & Newman, C. B. (2014). What matters in college for retaining aspiring scientists and engineers from underrepresented racial groups. *Journal of Research in Science Teaching, 51*(5), 555-580. https://doi.org/10.1002/tea.21146

Chaskes, J., & Anttonen, R. G. (2015). Advocating for first-year students. In M. L. Upcraft, J. N. Gardner, & B. O. Barefoot (Eds.). *The freshman year experience: Helping students survive and succeed in college* (pp. 191-203). San Francisco, California: Jossey Bass Publishers.

Chen, P., Chaves, O., Ong, D. C., & Gunderson, B. (2017). Strategic resource use for learning: A self-administered intervention that guides self-reflection on effective resource use enhances academic performance. *Psychological Science, 28*(6), 774-785. https://doi.org/10.1177/0956797617696456

Chen, X. (2013). STEM Attrition: College students' paths into and out of STEM fields (NCES 2014-001). Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Chew, K. J., Chen, H. L., Rieken, B., Turpin, A., & Sheppard, S. (2016). *Improving students' learning in statics skills: Using homework and exam wrappers to strengthen self-regulated learning.* Paper presented at the American Society for Engineering Education (ASEE) 123rd Annual Conference, New Orleans, LA.

Christianson, K., Mestre, J. P., & Luke, S. G. (2012). Practice makes (nearly) perfect: Solving 'students-and-professors'-type algebra word problems. *Applied Cognitive Psychology, 26*, 810-822. https://doi.org/10.1002/acp.2863

Cortes, S. (2017). Dr. Sergio Cortes' Organic Chemistry page. Retrieved from http://www.utdallas.edu/~scortes/ochem/

Crimmins, M. T. (2017). [Personal communication].

Davis, J. (2010). *The first-generation student experience: Implications for campus practice, and strategies for improving persistence and success.* Sterling, VA: Stylus Publishing, LLC.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment Physics class. *Science, 332(6031)*, 862-864. DOI: 10.1126/science.1201783

DiBenedetto, M. K., & Bembenutty, H. (2013). Within the pipeline: Self-regulated learning, self-efficacy, and socialization among college students in science courses. *Learning and Individual Differences, 23*, 218-224. https://doi.org/10.1016/j.lindif.2012.09.015

Doerschuk, P., Bahrim, C., Daniel, J., Kruger, J., Mann, J., & Martin, C. (2016). Closing the gaps and filling the STEM pipeline: A multidisciplinary approach. *Journal of Science Education and Technology, 25*(4), 682-695. https://doi.org/10.1007/s10956-016-9622-8

Dougherty, K. J. (1992). Community colleges and baccalaureate attainment. *Journal of Higher Education, 63*(2), 188-214. https://doi.org/10.1080/00221546.1992.11778349

Eagan, K., Hurtado, S., Figueroa, T., & Hughes, B. (2014). Examining STEM pathways among students who begin college at four-year institutions. Retrieved from http://sites.nationalacademies.org

Eberlein, T., Kampmeier, J., Minderhout, V., Moog, R. S., Platt, T., Varma-Nelson, P., & White, H. B. (2008). Pedagogies of engagement in science: A comparison of PBL, POGIL, and PLTL. *Biochemistry and Molecular Biology Education, 38*(4), 262-273. https://doi.org/10.1002/bmb.20204

Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? *CBE-Life Sciences Education, 13(3)*, 453-468. https://doi.org/10.1187/cbe.14-03-0050

Engle, J., & Tinto, V. (2008). *Moving beyond access: College success for low-income: First-generation students*. Retrieved from https://eric.ed.gov/?id=ED504448

Ertmer, P.A., & Newby, T.J. (1996). The expert learner: Strategic, self-regulated, and reflective. *Instructional Science, 24*, 1-24. https://doi.org/10.1007/BF00156001

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410-8415. https://doi.org/10.1073/pnas.1319030111

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research in Higher Education, 53(2)*, 229-261. https://doi.org/10.1007/s11162-011-9247-y

Gosser, D. K. (2011). The PLTL boost: A critical review of research. *Progressions, 14*(1).

Gregg-Jolly, L., Swartz, J., Iverson, E., Stern, J., Brown, N., & Lopatto, D. (2016). Situating second-year success: Understanding second-year STEM experiences at a liberal arts college. *CBE-Life Sciences Education, 15*(3). https://doi.org/10.1187/cbe.16-01-0044

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science, 332*(6304), 1213-1216. https://doi.org/10.1126/science.1204820

Hall, D. M., Curtin-Soydan, A. J., & Canelas, D. A. (2014). The science advancement through group engagement program: Leveling the playing field and increasing retention in science. *Journal of Chemical Education, 91*(1), 37-47. https://pubs.acs.org/doi/ipdf/10.1021/ed400075n

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Giffen, C. J., Blair, S. S., Rouse, D. I., & Hyde, J. S. (2014). Closing the social class achievement gap for first-generation students in undergraduate biology. *Journal of Educational Psychology, 106*(2), 375-389. http://psycnet.apa.org/doi/10.1037/a0034679

Hrabowski III, F. A. (2016). Transcript of keynote address by Freeman Hrabowski III at the 2014 Transforming Institutions Conference. In G. C. Weaver, W. D. Burgess, A. L. Childress, & L. Slakey (Eds.), *Transforming institutions: Undergraduate STEM education for the 21st century* (pp. 501-513). West Lafayette, Indiana: Purdue.

Hudesman, J., Crosby, S., Ziehmke, N., Everson, H., Isaac, S., Flugman, B., ... & Moylan, A. (2014). Using formative assessment and self-regulated learning to help developmental mathematics students achieve: A multi-campus program. *Journal on Excellence in College Teaching, 25*(2), 107-130.

Hunter, A.-B. (2016). *Factors contributing to undergraduate switching from STEM majors: Has anything changed in the last two decades?* Paper presented at the Annual Conference of the American Educational Research Association, Washington, D.C.

Jehangir, R. R. (2010). *Higher education and first-generation students: Cultivating community, voice, and place for the new majority*. New York, NY: Palgrave Macmillan.

Karabenick, S. A. (2003). Seeking help in large college classes: A person-centered approach. *Contemporary Educational Psychology, 28*(1), 37-58. https://doi.org/10.1016/S0361-476X(02)00012-7

Keeley, E.J. III & House, J.D. (1993). Transfer shock revisited: A longitudinal study of transfer academic performance. Paper presented at Association for Institutional Research Annual Forum, May 17, 1993, Chicago, Illinois.

Kezar, A., & Holcombe, E. (2016). Institutional transformation in STEM: Insights from change research and the Keck-PKAL project. In G. C. Weaver, W. D. Burgess, A. L. Childress, & L. Slakey (Eds.), *Transforming institutions: Undergraduate STEM education for the 21st century* (pp. 35-47). West Lafayette, Indiana: Purdue University Press.

Kobrak, M. (2006). *Surviving science at Brooklyn College: What every transfer student needs to know*. Brooklyn College. Brooklyn, New York.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134. http://psycnet.apa.org/doi/10.1037/0022-3514.77.6.1121

Lewis, S. E. (2011). Retention and reform: An evaluation of peer-led team learning. *Journal of Chemical Education, 88(6)*, 703-707. https://pubs.acs.org/doi/ipdf/10.1021/ed100689m

Lindsey, B. A., & Nagel, M. L. (2015). Do students know what they know? Exploring the accuracy of students' self-assessments. *Physical Review Special Topics - Physics Education Research, 11*(2), 020103-020101. https://doi.org/10.1103/PhysRevSTPER.11.020103

Lohfink, M. M., & Paulsen, M. B. (2005). Comparing the determinants of persistence for first-generation and continuing-generation students. *Journal of College Student Development, 46*(4), 409-428. https://doi.org/10.1353/csd.2005.0040

Lopez, E., Nandagopal, K., Shavelson, R., Szu, E., & Penn, J. H. (2013). Self-regulated learning study strategies and academic performance in undergraduate organic chemistry: An investigation examining ethnically diverse students. *Journal of Research in Science Teaching, 50*(6), 660-676. https://doi.org/10.1002/tea.21095

Lovett, M. C. (2013). Making exams worth more than the grade. In M. Kaplan, N. Silver, D. LaVaque-Manty, & D. Meizlish (Eds.), *Using reflection and metacognition to improve student learning* (pp. 18-52). Sterling, VA: Stylus.

Mäkinen, J., Olkinuora, E., & Lonka, K. (2004). Students at risk: Students' general study orientations and abandoning/prolonging the course of studies. *Higher Education, 48(2)*, 173-188. https://doi.org/10.1023/B:HIGH.0000034312.79289.ab

Malcom, S., & Feder, M. (Eds.). (2016). *Barriers and opportunities for 2-year and 4-year STEM degrees: Systemic change to support students' diverse pathways*. Washington, D.C.: National Academies Press.

Mason, A., & Singh, C. (2010). Do advanced physics students learn from their mistakes without explicit intervention. *American Journal of Physics, 78*(760). doi: 10.1119/1.3318805

Michaelsen, L. K., Knight, A. B., & Fink, L. D. (Eds.). (2004). *Team-based learning: A transformative use of small groups*. Sterling, VA: Stylus Publishing.

Moog, R. S., & Farrell, J. J. (2006). *Chemistry: A Guided Inquiry* (3rd ed.). New York, NY: John Wiley and Sons.

Mooring, S. R., Mitchell, C. E., & Burrows, N. L. (2016). Evaluation of a flipped, large-enrollment organic chemistry course on student attitude and achievement. *Journal of Chemical Education, 93*(12), 1972-1983. https://pubs.acs.org/doi/ipdf/10.1021/acs.jchemed.6b00367

Nandagopal, K., & Ericsson, K. A. (2012). An expert performance approach to the study of individual differences in self-regulated learning activities in upper-level college students. *Learning and Individual Differences, 22(5)*, 597-609. https://doi.org/10.1016/j.lindif.2011.11.018

National Science Foundation. (2013). *Women, minorities, and person with disabilities in science and engineering: 2013*

Nuñez, A. M., & Yoshimi, J. (2017). A phenomenology of transfer: Students' experiences at a receiving institution. *Innovative Higher Education, 42(2)*, 173-187. https://doi: 10.1007//s10755-016-9374-7

Olszewski, P. (2016). Teaching millennials how to study under the 21st century sky. *Pyrex Journal of Educational Research and Reviews, 2*(1), 1-9.

Pazicni, S., & Bauer, C. F. (2014). Characterizing illusions of competence in introductory chemistry students. *Chemistry Education Research and Practice, 15*, 24-34. https://doi.org/10.1039/C3RP00106G

Phillips, J. A., Clemmer, K. W., McCallum, J. E. B., & Zachariah, T. M. (2017). A problem-solving framework to assist students and teachers in STEM courses. *Journal of College Science Teaching, 46*(4), 33-39.

Plant, E. A., Ericsson, K. A., Hill, L., & Asberg, K. (2005). Why study time does not predict grade points average across college students: Implications of deliberate practice for academic performance. *Contemporary Educational Psychology, 30(1)*, 96-116. https://doi.org/10.1016/j.cedpsych.2004.06.001

Popham, W. J. (Ed.) (1971). *Criterion-referenced measurement: An introduction*. Englewood Cliffs, NJ: Educational Technology Publications.

Reusch, W. (2017). Organic Chemistry practice problems at Michigan State University. Retrieved from https://www2.chemistry.msu.edu/faculty/reusch/virttxtjml/Questions/problems.htm

Rosser, S. V. (1997). *Re-Engineering female friendly science*. New York, NY: Teachers College Press.

Rozell, T. G., Johnson, J., Sexten, A., & Rhodes, A. E. (2017). Correcting missed exam questions as a learning tool in a physiology course. *Journal of College Science Teaching, 46*(5), 56-63.

Saenz, V. B., Hurtado, S., Barreta, D., Wolf, D. S., & Yeung, F. (2007). *First in my family: A profile of first-generation college students at four-year institutions since 1971*. Los Angeles: Reagents of the University of California.

Sebesta, A. J., & Bray Speth, E. B. (2017). How should I study for the exam? Self-regulated learning strategies and achievement in introductory biology. *CBE-Life Sciences Education, 16*(2). https://doi.org/10.1187/cbe.16-09-0269

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*: HarperCollins Publishers, Inc.

Smedley, B. D., Butler, A. S., & Bristow, L. R. (Eds.). (2004). *In the nation's compelling interest: Ensuring diversity in the health-care workforce*. Washington, D.C.: National Academies Press.

Soicher, R. N., & Gurung, R. A. R. (2017). Do exam wrappers increase metacognition and performance? A single course intervention. *Psychology Learning and Teaching, 16*(1), 64-73. https://doi.org/10.1177/1475725716661872

Stanton, J. D., Neider, X. N., Gallegos, I. J., & Clark, N. C. (2015). Differences in metacognitive regulation in introductory biology students: When prompts are not enough. *CBE - Life Sciences Education, 15*, 1-12. https://doi.org/10.1187/cbe.14-08-0135

Stephens, N. M., Hamedani, M. G., & Destin, M. (2014). Closing the social-class achievement gap: A difference-education intervention improves first-generation students' academic performance and all students' college transition. *Psychological Science, 25*(4), 943-953. https://doi.org/10.1177/0956797613518349

Stuber, J. M. (2012). *Inside the college gates: How class and culture matter in higher education*. New York, NY: Rowman & Littlefield Publishers, Inc.

Szu, E., Nandagopal, K., Shavelson, R., Lopez, E. J., Penn, J. H., Scharberg, M., & Hill, G. W. (2011). Understanding academic performance in Organic Chemistry. *Journal of Chemical Education, 88*(9), 1238-1242. https://pubs.acs.org/doi/ipdf/10.1021/ed900067m

Talbert, R. (2017). *Flipped learning*. Sterling, VA: Stylus.

Terenzini, P., Springer, L., Yaeger, P. M., Pascarella, E. T., & Nora, A. (1995). *First-generation college students: Characteristics, experiences, and cognitive development*. Paper presented at the Meeting of the Association for Institutional Research, Boston, MA.

Thompson, D. R. (2012). Promoting metacognitive skills in intermediate Spanish: Report of a classroom research project. *Foreign Language Annals, 45*(3), 447-462. https://doi.org/10.1111/j.1944-9720.2012.01199.x

Tipton, R.L., & Bender, P. (2006). From failure to success: Working with under-prepared transfer students. *Reference Services Review, 34*(3), 389-404. https:// doi.org/ 10.1108/00907320610685337

Townsend, B. K. (1995). Community college transfer students: A case study of survival. *The Review of Higher Education*, *18*(2), 175-193. https://doi: 10.1353/rhe.1995.0022

Townsend, B.K., & Wilson, K.B. (2006). The transfer mission: Tried and true, but troubled? *New Directions for Community Colleges, 2006*(136), 33-41. https://doi.org/10.1002/cc.257

Uche, A. R. (2015). *The retention of first-generation college students in STEM: An extension of Tinto's longitudinal model.* (Ph.D.), University of North Carolina, Charlotte.

VanderStoep, S. W., Pintrich, P. R., & Fagerlin, A. (1996). Disciplinary differences in self-regulated learning in college students. *Contemporary Educational Psychology, 21*(4), 345-362. https:// doi.org/10.1006/ceps.1996.0026

West, C., & Sadoski, M. (2011). Do study strategies predict academic performance in medical school? *Medical Education, 45*(7), 696-703. https://doi:10.1111/j.1365-2923.2011.03929.x

Wieman, C. E. (2016). Foreword. In G. C. Weaver, W. D. Burgess, A. L. Childress, & L. Slakey (Eds.), *Transforming institutions: Undergraduate STEM education for the 21st century* (pp. ix-xiv). West Lafayette, Indiana: Purdue University Press.

Watkins, J., & Mazur, E. (2013). Retaining students in science, technology, engineering, and mathematics (STEM) majors. *Journal of College Science Teaching, 42*(5), 36-41. http://www.jstor.org/stable/43631580

Weinstein, C. E., & Acee, T. W. (2013). Helping college students become more strategic and self-regulated learners. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines* (pp. 197-236). Charlotte, NC: Information Age Publishing.

Yuen-Reed, G., & Reed, K. B. (2015). Engineering student self-assessment through confidence-based scoring. *Advances in Engineering Education*, *4*(4), 1-23.

Zhao, N., Wardeska, J. G., McGuire, S. Y., & Cook, E. (2014). Metacognition: An effective tool to promote success in college science learning. *Journal of College Science Teaching, 43*(4), 48-54. http://www.jstor.org/stable/43632012

Zimmerman, B.J., & Martinez-Pons, M. M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23*(4), 614-628. https://doi.org/10.3102/00028312023004614

Zimmerman, B.J., Moylan, A., Hudesman, J., White, N., & Flugman, B. (2011). Enhancing self-reflection and mathematics achievement of at-risk urban technical college students. *Psychological Test and Assessment Modeling, 53*(1), 108-127

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3-17. https://doi.org/10.1207/s15326985ep2501_2

Zimmerman, B.J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*(1), 51-59. http://psycnet.apa.org/doi/10.1037/0022-0663.82.1.51