Pearson

Technical Report

Quasi-experimental Study of Mastering Chemistry, Ohio State University: Final Technical Report

SOW No. 2
SRI Project No. 23282

Prepared for:
Emily Lai

Authors:
SRI International
Rebecca Griffiths
Jared Boyce
Shuai Wang
Tallie Wetzel

Other contributors:
Robert Murphy

Pearson Global Product Organization
Efficacy & Research
Impact Evaluation

Table of Contents

List of Tables

List of Figures

Executive Summary

- In the United States General Chemistry is a required gateway course for most STEM majors. A growing number of instructors have adopted adaptive learning systems as homework supplements and to support classroom instruction. As use of these systems becomes more widespread, it is important that we understand their efficacy, how they are implemented, and what can be done to best help promote student learning. This need is especially acute in postsecondary chemistry courses, which have been the subject of very few rigorous studies. There is a pressing need to better understand how various adaptive learning products can best help students in General Chemistry courses.
- In order to advance this field of research, Pearson contracted with SRI Education, a non-profit research and development organization, to conduct an evaluation of Mastering Chemistry, an online homework, tutorial, and assessment system designed for use in postsecondary General Chemistry courses. To support independent study by students as well as their classroom activities, the system includes features enabling personalized learning to address students' individual needs. Assessments include quizzes and homework problems aligned with Pearson textbooks.
- SRI conducted a study of Mastering Chemistry with Ohio State University (OSU) during fall semester 2016 and spring semester 2017 in General Chemistry I (Chem 1210) and General Chemistry II (Chem 1220). (Another study of Mastering Chemistry was conducted with a second selective research university and is described in a separate report.) This study examined the implementation of Mastering Chemistry, including how the system was used in the course, the instructor and student experience, and the association between usage of the system and course outcomes. In addition, we conducted a quasi-experimental study comparing outcomes of students in sections that used Mastering Chemistry with other sections that used another online homework system called Sapling. We also compared costs for the two versions of the course.
- Data collections for the study included pre- and post-surveys of students, an instructor survey, qualitative data from a site visit, course outcomes, background information about students, use data provided by Pearson, and cost data. Student and instructor surveys were administered during fall 2016. The site visit to OSU also took place during fall 2016 and involved interviews with instructors and the department chair, a student focus group, and class observations. We gathered and analyzed data from the student pre-survey, instructor survey, and site visit activities for sections that used Sapling (the comparison condition) as well as those that used Mastering Chemistry. For the cost analysis we collected information through the instructor survey, site visit, and phone interviews.

- Six instructors participated in the study, though only five participated in the impact analyses of Mastering Chemistry users vs. nonusers. The final *fall* analytic sample included in the comparative impact analysis consisted of 1,700 students (742 Mastering Chemistry users, 958 nonusers) for the analyses of most course outcomes; of the 742 Mastering Chemistry users, 623 students had sufficient data to be included in the use analyses. The final *spring* analytic sample included in the impact analysis consisted of 1,137 students (782 Mastering Chemistry users, 355 nonusers) for the analyses of most course outcomes. Of the 782 Mastering Chemistry users, 774 Mastering Chemistry users were included in the use analyses. A substantial number of students were eliminated from the study during the process of matching datasets in both the fall and spring studies.
- For both fall 2016 and spring 2017 we used a non-equivalent comparison group quasi-experimental design to estimate effects on student outcomes of the use of Mastering Chemistry in General Chemistry courses. We examined three outcomes: final scores on a common exam based on an assessment from the American Chemical Society, total course grades, and whether students passed the course. In addition, we looked for interaction effects with background characteristics and examined the relationship between the level of student use of Mastering Chemistry (specifically, the number of attempted problems) and achievement outcomes.
- We used linear regression models and, when warranted, hierarchical regression models (i.e., students clustered within sections) to estimate the impact of using Mastering Chemistry on student outcomes relative to the comparison condition and to analyze usage data. The models controlled for student background characteristics, including age, gender, and baseline measures of academic preparation (e.g., SAT or ACT score). Because the number of instructors was too small to be included in the statistical model, we examined possible differences in instructor-related factors between conditions qualitatively.
- In order to address implementation research questions, we analyzed data from surveys and site visits. Notes from the interviews and focus groups were reviewed for key themes across respondents. For the cost analysis, we took an "ingredients" approach, which identifies all inputs regardless of who bears the costs. Our analyses primarily focused on costs of providing instruction, primarily staffing, as well as access costs for students. Total costs in each condition were divided by the number of students to calculate a cost per student, and the difference in cost per student was then divided by the effect size to generate a cost per unit of improvement (or decrement) relative to the comparison condition.

Results

Instructors at OSU used Mastering Chemistry in General Chemistry I and II primarily as a homework supplement and to enable active classroom activities. Some instructors used Mastering Chemistry to help students prepare for the lecture by assigning readings, videos, and Mastering Chemistry problems in advance of lectures and, by requiring problems to be completed the night before class. This approach freed up class time for scaffolded problem sessions and discussion, and also allowed the instructor to monitor students' preparation.

Instructors reported that the system allowed them to give students more frequent assignments with immediate feedback and that students were better prepared for class as a result. Students also reported that Mastering Chemistry helped them to prepare for class as well as tests and quizzes. They rated Wrong Answer Feedback as by far the most frequently used feature, followed by Hints and Videos. They were less likely to report that Mastering Chemistry enabled them to receive more personalized feedback or that it increased their enjoyment of the course. The primary challenge cited by both instructors and students was the system's inflexibility with regards to the format it accepted for inputting symbols as part of solutions.

In fall 2016, students who used Mastering Chemistry performed significantly better on all three outcome measures examined than non-Mastering Chemistry users, when controlling for student backgrounds and accounting for section-level effects. Similarly, for spring 2017, for two of the three outcome measures examined—course grades and binary grades—students using Mastering Chemistry significantly outperformed non-users. Students in the comparison condition were assigned to use Sapling solely as a homework supplement. Finally, our analysis of usage data indicated that the number of problems attempted had a statistically significant positive relationship with all three student course outcome measures (final scores, total scores, and binary grades).

We found that implementation of Mastering Chemistry took marginally more time for instructors than use of Sapling, while the license costs for students were the same. Differences in instructor time yielded a modest difference of $.67 per student per semester for use of Mastering Chemistry.

While we found that use of Mastering Chemistry was associated with many positive outcomes, the results cannot be interpreted as evidence of a causal effect due to limitations of the study. These included substantial attrition of students (exceeding 30% in both fall 2016 and spring 2017) due to issues linking various datasets. We do not have reason to believe there were systematic biases in student attrition, and differential attrition between Mastering Chemistry and comparison groups was much lower.

Another, perhaps more significant, limitation is that we cannot rule out the possibility that differences in student performance were caused by differences among instructors rather than use of Mastering Chemistry. Instructors using Sapling had somewhat higher average years of experience teaching at the college level but fewer years of experience using the product than instructors using Mastering

Chemistry. Future studies could increase the confidence of these findings by increasing the number of instructors involved.

Introduction

Background

In the United States General Chemistry is a required gateway course for most STEM majors. A growing number of instructors have adopted adaptive learning systems as homework supplements and to support classroom instruction. In many cases instructors use these systems to "flip" their classes, assigning students to review materials ahead of class so that the time can be used for active problem solving. As use of adaptive learning becomes more widespread, it is important that we understand how these systems impact student learning outcomes.

Although prior research has enumerated a multitude of reasons why, theoretically, adaptive software products will help learners, very few high quality efficacy studies have been conducted to test these propositions in General Chemistry, especially at the postsecondary level. Almost all studies have been in the fields of computer science, math, and physics. For example, out of seven high quality meta-analyses or meta-analytic reviews conducted on adaptive learning technologies in the last decade (Steenbergen-Hu and Cooper 2013, Steenbergen-Hu and Cooper 2014, Ma et al 2014, Kulik and Fletcher 2015, VanLehn 2011, Nesbit et al 2014, Durlach and Ray 2011), only around three articles out of over one hundred focused on chemistry. Out of these three (Mclaren and Isotani 2011, Mclaren et al 2011, Adamson et al 2014), only one focused on college-level students. There is thus a pressing need to better understand how different adaptive learning can best help students in chemistry courses.

In order to advance this field of research, Pearson has contracted with SRI Education (https://www.sri.com/about/organization/education), a non-profit research and development organization in Menlo Park, CA, to conduct an evaluation of Mastering Chemistry, an adaptive learning system designed for use in postsecondary General Chemistry courses.

This is the final report on the quasi-experimental study of Mastering Chemistry in General Chemistry I (Chem 1210) and General Chemistry II (Chem 1220) at Ohio State University (OSU). It includes results from data that SRI collected and analyzed for fall semester 2016 and spring semester 2017, including surveys of students and instructors, qualitative data from a site visit, course outcomes, background information about students, use data provided by Pearson, and cost data. Another study of Mastering Chemistry was conducted with a second selective research university and is available in a separate report.

Description of Mastering Chemistry

Mastering Chemistry is an online homework, tutorial, and assessment system for postsecondary introductory General Chemistry courses. The system is designed to improve results and increase student engagement before, during, and after class. To support independent study by students as well as their classroom activities it includes features that provide personalized learning that addresses students' individual needs. Assessments include quizzes and homework problems aligned with Pearson textbooks. Instructional supports include:

- hints, with targeted scaffolding for specific problems
- instructional videos
- wrong answer feedback
- Dynamic Study Modules, which help students study on their own by continuously assessing their activity and performance in real time and providing feedback
- Adaptive Follow-up Assignments, which are based on each student's past performance on coursework and which provide additional coaching and targeted practice as needed
- Learning Catalytics (LC), a tool for instructors to generate class discussion, customize lectures, and promote peer-to-peer learning using students' smartphones, tablets, or laptops to engage them in more interactive tasks and thinking.

Instructors have the option of using or disabling some features. They can also customize items in some features, such as Dynamic Study Modules, for better alignment with their own syllabi.

The Present Study

The present study addressed the following primary research questions regarding Mastering Chemistry:

Implementation and experience

- What was the intended role of the product within the instructional system to support teaching and learning?
- What was the intended role of the instructors and what practices were used to integrate the product?
- To what extent were the products used as intended?
- What institutional, human, and technology factors facilitated or hindered product implementation?
- To what extent were instructors and students satisfied with their experience using the product and with the training and support provided?

Outcomes

- Have students who used Mastering Chemistry achieved better completion and progression outcomes than have similar students who did not use the product?

Moderators

- Do the measured effects vary for students in regard to prior achievement, age, enrollment status, or other policy-relevant student characteristics (such as English language learner [ELL] status, Pell grant status, gender, enrollment status)?
- Do the effects vary in regard to different instructor characteristics or practices?

Effects of implementation on student outcomes

- What are the relationships among incoming student proficiency, product use variables (such as frequency, units completed, features used), and student outcomes?

Cost-effectiveness

- How does the use of Mastering Chemistry in General Chemistry courses affect the cost structure of the courses compared with alternative approaches?
- What is the cost per unit of desired outcome for Mastering Chemistry use?

Methods

Participants

Table 1 shows the number of instructors and students associated with the fall 2016 and spring 2017 General Chemistry classes included in the study. Table 2 provides descriptive statistics for the student samples for fall 2016/spring 2017.

Table 1: Number of instructors and students in 2016–17 General Chemistry classes

Term	Instructors		Students	
	Mastering Chemistry	Nonusers	Mastering Chemistry	Nonusers
Fall	2*	4	1,063	1,633
Spring	3	1	1,326	678

**The cost analysis uses data from three instructors for fall, including one who taught a General Chemistry II course that is not included in the impact analysis.*

Table 2: Characteristics of the student sample by term

Characteristic	Fall 2016 Full Sample (N = 2,666)	Fall 2016 Analytic Sample (N = 1,700)	Spring 2017 Full Sample (N = 2,004)	Spring 2017 Analytic Sample (N = 1,187)
Gender				
Male	1,083 (41%)	676 (40%)	868 (43%)	489 (41%)
Female	1,527 (57%)	1,012 (60%)	1,134 (57%)	698 (59%)
Age (Median)	19.0	19.0	NA	19.0
Enrollment status				
Part-time	52 (2%)	17 (1%)	NA	12 (1%)
Full-time	2,528 (95%)	1,203 (71%)	NA	951 (80%)
Declared major (Yes)	2,257 (85%)	1,499 (88%)	NA	1,062 (89%)
Work status				
Not working	1,870 (70%)	915 (54%)	NA	747 (63%)
Part-time	715 (27%)	312 (18%)	NA	225 (19%)
Full-time	34 (1%)	8 (0.5%)	NA	4 (<1%)

Parent college attendance (Yes)	2,263 (85%)	1,509 (89%)	NA	1,048 (88%)
English spoken as primary language in the home (Yes)	2,324 (87%)	1,501 (88%)	NA	1,019 (86%)

Note: Not all percentages add to 100% because some students did not reply to all questions. The majority of demographic variables were collected from the student survey, which was administered only in the fall semester. Therefore, we provide two sets of demographics for the spring 2017 semester. The full sample characteristics reflect the characteristics of all students for whom we received outcome data. The analytic sample characteristics reflect the characteristics of the students who were also present in the fall semester and completed the survey questions.

Analytical Samples for Estimating Impacts

The final analytic sample used for estimating impacts (Mastering Chemistry users vs. nonusers, and the relationship between use and outcomes among users) represents a significantly smaller subset than the full sample of all students who completed the fall and spring General Chemistry courses and received a course grade.

Fall sample: The data file for OSU students with course outcomes (grades and final exam scores) included 2,696 individual student records (1,063 Mastering Chemistry users and 1,633 nonusers). This data file was merged with the post-survey file, which contained demographic variables (such as gender) needed for the impact analysis. OSU provided outcome data for 2,696 General Chemistry I students. The post-survey contained 1,878 students. Of the 1,878 students, we were able to merge 1,734 students between the two files. A further 34 students were excluded from the sample because of missing data for the key variables used in the models. The final fall analytic sample included in the comparative impact analysis (Mastering Chemistry users vs. nonusers) consisted of 1,700 students (742 Mastering Chemistry users, 958 nonusers) for the analyses of most course outcomes.

With respect to the usage analytic sample, students had to accurately recall and enter their student survey ID numbers from the pre-survey when completing the post-survey in order to merge the student outcome data with the usage data. This limitation was introduced when the institution switched from a paper/pencil survey implementation to an online survey implementation. Of the 1,878 observations in the post survey, 1,600 students entered a unique complete student survey ID. Of the 1,600 students, 692 students were Mastering Chemistry users. Of the 692, 677 were successfully merged to the usage data. The difference was likely due to merging on student name – OSU may have provided one variation of the name whereas Pearson may have provided another variation of the name. Of the 677, 634 were successfully merged to outcome data. And then finally, of the 634, 623 students had sufficient data to be included in the use analyses.

Spring sample: The data file for OSU students with course outcomes (grades and final exam scores) included 2,004 individual student records (1,326 Mastering Chemistry users and 678 nonusers). The age variable, which was required for impact analyses, was merged in from the fall student post-survey. Of the 2,004 General Chemistry II students in the OSU spring data file, 1,652 were successfully linked to a student survey. Of those 1,652 students, 1,187 had age information. A further 50 students were excluded from the sample because of missing data for the key variables used in the models. The final spring analytic sample included in the comparative impact analysis (Mastering Chemistry users vs. nonusers) consisted of 1,137 students (782 Mastering Chemistry users, 355 nonusers) for the analyses of most course outcomes. Of the 782 Mastering Chemistry users, 774 Mastering Chemistry users were included in the use analyses.

With respect to the usage analytic sample, of the original 1,326 treatment students from the full sample size, 1,311 students were successfully linked to the Pearson data file. The 15 students we were not able to match were a result of merging using student email. Of the 1,311 students, approximately 795 had non-missing age information. (Recall from above the age information had to be pulled in from the student survey.) Of the remaining 795 students, 774 had complete information to be included in the use analyses.

Appendix B provides additional details regarding the data files, data cleaning, and data merging, including dataset linking issues. For both Fall 2016 and Spring 2017 datasets, SRI performed statistical tests to determine whether or not there were systematic biases in the missingness of student data. No systematic biases were detected.

Outcome Measures

For the impact analyses we examined three outcomes: final exam scores, total course grades, and whether students passed the course. The latter used a binary outcome variable of Pass for students who received a C- or above, or Fail for students who failed or withdrew from the course.

The final exam score outcome measures are assessments required for all students in General Chemistry I and II. The exams are derived from an assessment provided by the American Chemical Society (ACS). ACS exams are nationally normed and produced by committees of subject matter experts.

The composition of course grades varies by instructor. Final exams account for 25-27.5% and homework accounts for 7.5-15%.

Data Collections

Surveys

Student surveys. Student pre-surveys were distributed in the first week of fall semester for both students in sections that used Mastering Chemistry and those that did not use Mastering Chemistry. Student post-surveys were distributed in the last week of classes during the fall 2016 semester. Surveys were administered to all students 18 years or older in General Chemistry courses. As per the study plan, only post-surveys from students using Mastering Chemistry were analyzed. The pre-survey was administered in print using TeleForm by Cardiff Software, which allows completed surveys to be scanned with the survey data captured electronically. The post-survey was administered electronically at OSU as requested by the site coordinator.

Instructor surveys. Instructor surveys, delivered online, were distributed three to five weeks before course completion. Instructor surveys were administered in fall 2016 term to instructors who both used and did not use Mastering Chemistry.

Site Visits

SRI staff conducted site visits during fall 2016 to a sample of instructors and their students in the Mastering Chemistry sections and comparison section. Site visits consisted of student focus groups, classroom observations, and in-person interviews with the department chair and instructors. The visits/interviews were conducted three to five weeks before course completion. This was to enable capture of the experience of the students and instructors toward the end of the course, and to avoid interfering with finals week and finals preparation.

System Use Log Files

SRI coordinated with Pearson data scientists to access backend data to monitor the use of the products and for information on product implementation. These data were obtained for each semester during academic year (AY) 2016-17.

Collecting Information on Instruction in the Comparison Group

In addition to conducting surveys and site visit data collections with Mastering Chemistry students and instructors, we also did so for concurrent comparison groups.

- Pre-surveys were administered to students using Mastering Chemistry and comparison group students. While post-surveys were administered to both Mastering Chemistry and comparison group students due to a miscommunication, only Mastering Chemistry-using post-surveys were analyzed.
- Instructors not using Mastering Chemistry were surveyed at the same time as instructors using Mastering Chemistry.

- During the site visits, we interviewed one instructor who did not use Mastering Chemistry, and we conducted a classroom observation in one non-Mastering Chemistry section. The department chair interview included questions about instruction in sections taught both with and without Mastering Chemistry.

Analytic Approaches

Impact Analysis

For both fall 2016 and spring 2017 we used a non-equivalent comparison group quasi-experimental research design to estimate effects on student outcomes of the use of Mastering Chemistry in General Chemistry courses. We used linear regression models and, when warranted, hierarchical regression models (that is, students clustered within sections) to estimate the impact of using Mastering Chemistry on student outcomes in terms of achievement, completion, and progression.

Because we did not randomly assign students to conditions, the baseline composition of students in courses using Mastering Chemistry may differ from students in courses that did not (selection bias). To establish the degree of equivalence, we compared the composition of the two groups on a set of baseline covariates, including gender, age, and baseline measures of academic preparation (for example, SAT or ACT score). Relying on What Works Clearinghouse standards, we considered groups within .25 standard deviations in age, gender, and academic preparation to be acceptably equivalent. We then used baseline student-level covariates in the hierarchical models to adjust for any baseline nonequivalence. If the groups differed in composition beyond these standards, we intended to use propensity score matching and baseline covariates in the hierarchical regression models to adjust for baseline nonequivalence between the Mastering Chemistry and comparison groups.

Because the number of instructors was too small to be included in the statistical model, we examined possible differences between conditions in instructor-related factors qualitatively.

Subgroup analyses. To examine the efficacy of the intervention for students with less prior academic preparation than that of the comparison group and of other students using Mastering Chemistry, we augmented the primary hierarchical analytical model with an interaction term of the student-level treatment variable and of whether the student's prior academic preparation measure (such as ACT or SAT score) was below the median for the sample. We used a similar model to test for differences in the estimated effects of age and gender. In addition, we conducted exploratory correlational analyses as described below.

Relationships among student factors, use of Mastering Chemistry, and student outcomes. We used linear regression models (hierarchical regression models when appropriate) to examine the relationship between the level of student use of Mastering Chemistry and achievement outcomes, controlling for students' baseline characteristics. Measures of use included the number of attempted problems.

Examination of implementation, practices, and student and instructor experience. Data from surveys and site visits were analyzed to understand how Mastering Chemistry was used, the facilitating factors and challenges associated with its use, and degree of student and instructor satisfaction with the courseware. To address these questions, descriptive statistics survey data were supplemented with a narrative analysis of the qualitative data from instructor interviews and student focus groups. Notes from the interviews and focus groups were reviewed for key themes that emerged across respondents.

Results

Implementation and Experience

Use Models

- Mastering Chemistry was used by OSU instructors to assign, support, and grade homework. Students' performance in Mastering Chemistry also contributed to their course grades, accounting for between 7.5% and 15% of course grades.
- Some instructors used Mastering Chemistry to help students prepare for the lecture and held them accountable for undertaking the assigned readings, videos, and Mastering Chemistry problems related to the next lecture, and by requiring the problems to be completed the night before the lecture. This approach freed more time during the lecture for scaffolded problem sessions and discussion, and also allowed the instructor to determine the extent to which students were preparing for lectures.
- One instructor used Learning Catalytics (LC) regularly (in addition to other apps) to support his "flipped" classroom (in other words, lectures that focused solely on group and individual problem-solving sessions).
- Instructors reported that they rarely used the dashboards to monitor student performance.

Benefits of Use (as reported by instructors and students)

- Homework was automatically graded, providing students with immediate feedback. Students found problem sets that included hints very helpful because they could learn from their mistakes and continue to make progress. Before the use of online homework tools like Mastering Chemistry, students at OSU did not have their homework graded.
- Instructors reported that use of Mastering Chemistry saved time on mechanical tasks such as entering grades but that Mastering Chemistry did not save time for teaching assistants (TAs).
- Instructors found that Mastering Chemistry allowed them to make more frequent assignments aligned with the lectures and, for some, to make sure that students were prepared for the lectures and labs. Before they used Mastering Chemistry, most instructors assigned only weekly problem sets.
- Because Mastering Chemistry is integrated with the textbook, the instructors felt confident that the problems assigned were aligned with the content, skills, and procedures covered in the readings. When using e-textbooks, they indicated they could easily indicate the relevant section in the textbook and accompanying videos (if available).

- Two instructors reported reviewing information the system provided on the time students took to answer different types of problems. They used this information to create problem sets that students should be able to complete in a reasonable amount of time.

Challenges to Use

- Many instructors and students complained about Mastering Chemistry's inflexibility in regard to the format it accepted for inputting symbols as part of solutions. Many students expressed frustration with getting answerss incorrect because of formatting issues. In some cases, instructors chose not to assign problems of a particular type with known solution format issues.
- Students were unhappy with instructors setting up Mastering Chemistry to deduct points for each failed attempt at a problem solution. However, many instructors commented they believed this aspect increased student motivation.
- Some instructors and students commented that the time required to load individual questions was too long. Some commented that this issue was likely to be even more problematic for students with slower computers and internet connectivity.
- One challenge to LC use was that it required students to have access to their laptops during lectures. One instructor mentioned that the time lost when students needed to log into their computers to use this feature was his main reason for not using LC. Indeed, our observations indicated most students took notes in their paper notebooks.
- Some students mentioned that they wished it were possible to use Mastering Chemistry to prepare for tests and exams by being able to review problem types they answered incorrectly in past assignments and have Mastering Chemistry point them to similar problems for practice. Students were not aware that Mastering Chemistry has a feature that enables this type of review.

Student Pre-survey

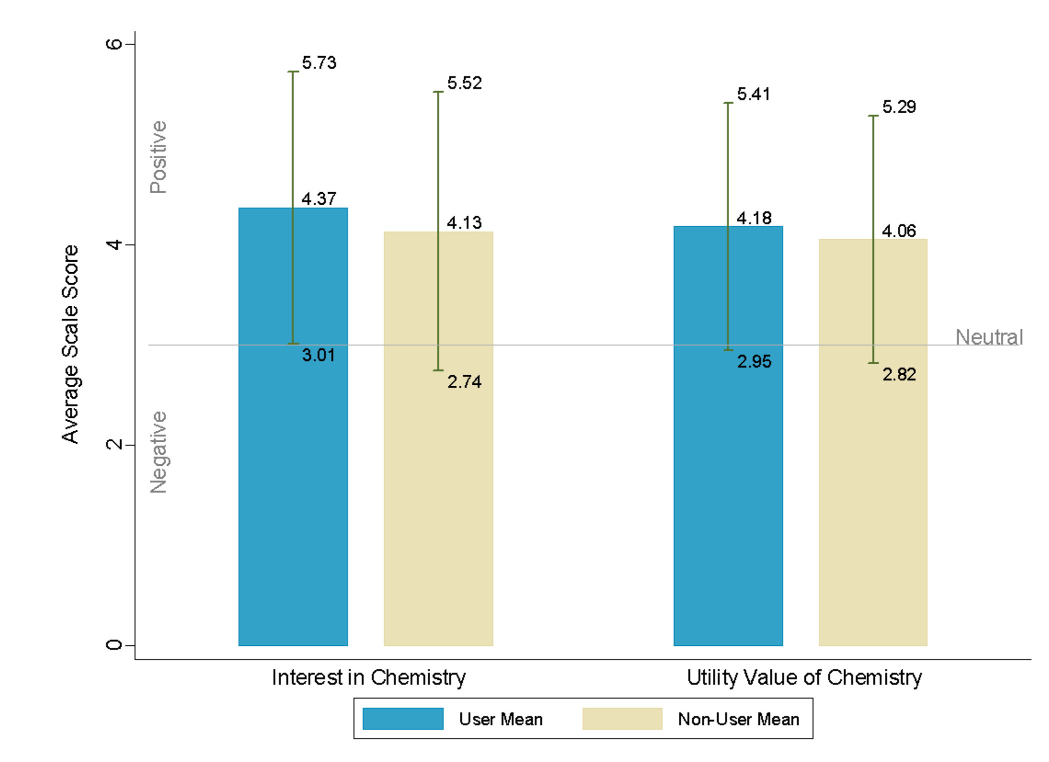
The student pre-survey was used to capture students' baseline attitudes toward chemistry, including their interest in the subject area and their beliefs about its relevance to their lives (see Figure 1).[1] (Details on the specific statements are provided as a note under Figure 1 and subsequent figures. The distribution of students' baseline responses for each statement by institution are shown under Question 5 in Appendix C.)

---

[1] *The two scales – Interest in Domain and Utility Value of Domain – were adapted from Hulleman, C. S., & Harackiewicz, J. M. (2009). Making education relevant: Increasing interest and performance in high school science classes. Science, 326, 1410-1412.*

As shown in Figure 1, students on average had positive views about their baseline interest in chemistry and on the utility of what they were learning.

Figure 1: Students' interest in chemistry and their beliefs in its utility



Note: The Interest in Chemistry scale is a three-item scale based on data collected from the student pre-survey. Survey items were coded 0-6, running from "Not at all true" to "Very True" with "Neutral" coded as 3. Students were asked to report the extent to which the following statements described themselves in their chemistry class: "I think the field of chemistry is interesting." "To be honest, I just don't find chemistry interesting." "I think what we will be learning in class will be interesting." Similarly, the Utility Value of Chemistry scale is a three-item scale based on data collected from the student pre-survey. Survey items were coded 0-6, running from "Not at all true" to "Very True" with "Neutral" coded as 3. Students were asked to report the extent to which the following statements described themselves in their chemistry class: "I can apply what we are learning in chemistry class to real life." "I think what we are studying in chemistry class is useful to know." "I can see how what I learn from chemistry applies to life." The values shown on the graph represent (1) the mean value for the institution on the scale (value at the top of the bar), and (2) the values that correspond to plus and minus 1 standard deviation (SD) from the mean value.

Student Post-survey

Figures 2 through 7 show the distribution of frequency responses for key implementation questions from the student post-survey. Appendices C and D provide frequency tables for the pre- and post-survey questions, respectively.

As shown in Figure 2, Students at OSU rated wrong answer feedback as by far the most useful feature, with 90% of respondents saying they "often" or "sometimes" used this feature. Hints followed, with 77% of students reporting use of this feature, followed by videos and LC. Dynamic Study Modules ranked lowest, with only 46% of students reporting significant use of this feature.

As shown in Figure 3, students generally gave very high ratings for the features that they reported using. (Note that students were only asked to rate the helpfulness of features that they reported using.) The top three most useful features were hints, tutorials, and wrong answer feedback. Dynamic Modules, Adaptive Follow-up Homework, and LC received slightly lower ratings of helpfulness, but even these were rated as "somewhat" or "very helpful" by over 90% of respondents.

In terms of the benefits of Mastering Chemistry, students were most likely to report that they felt better prepared and had access to a greater variety of learning materials, and that Mastering Chemistry helped them prepare for tests/quizzes. They were less likely to report that Mastering Chemistry enabled them to receive more personalized feedback and that it increased their enjoyment of the course, though over 70% of students still cited these as benefits (sometimes/mostly/very true), as shown in Figure 4.

As shown in Figure 5, students overall found Mastering Chemistry easy to use. Over 90% of students reported that most Mastering Chemistry functions were easy to use. The function with the lowest reported ease of use was "finding the relevant information in the digital textbook", though a majority of students still reported that this was easy to do (62%).

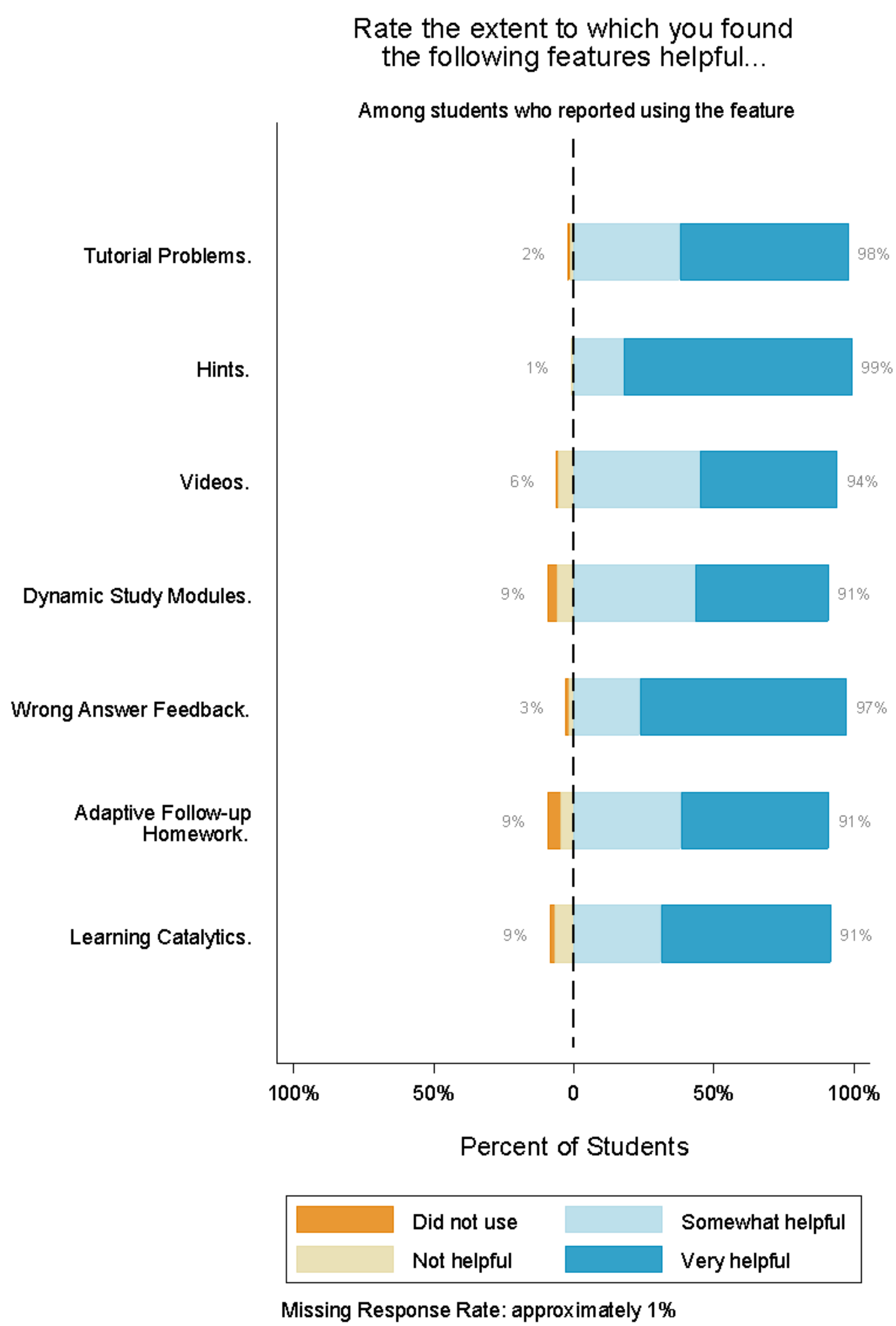
A majority of students reported some degree of negative impact on their use of Mastering Chemistry due to insufficient academic supports, as shown in Figure 6. Each of the specific academic supports asked in the survey—peer study groups, individual tutoring sessions with TAs or instructors, group tutorial sessions/recitations—was reported by over 50% of students to be insufficient to the point of affecting how they used Mastering Chemistry.

With respect to usability and technology support, Figure 8 shows that most students did not report negative impacts on how they used Mastering Chemistry for most of the specific issues asked in the survey. The exception to this is that 82% of students reported questions were marked as incorrect due to answers not being in the right format and that this negatively affected their use of Mastering Chemistry. To the extent that there were negative impacts due to other usability and technology support issues, the impacts were more often reported to be 'small' rather than either 'moderate' or 'significant'.

Figure 2: Use of different Mastering Chemistry features

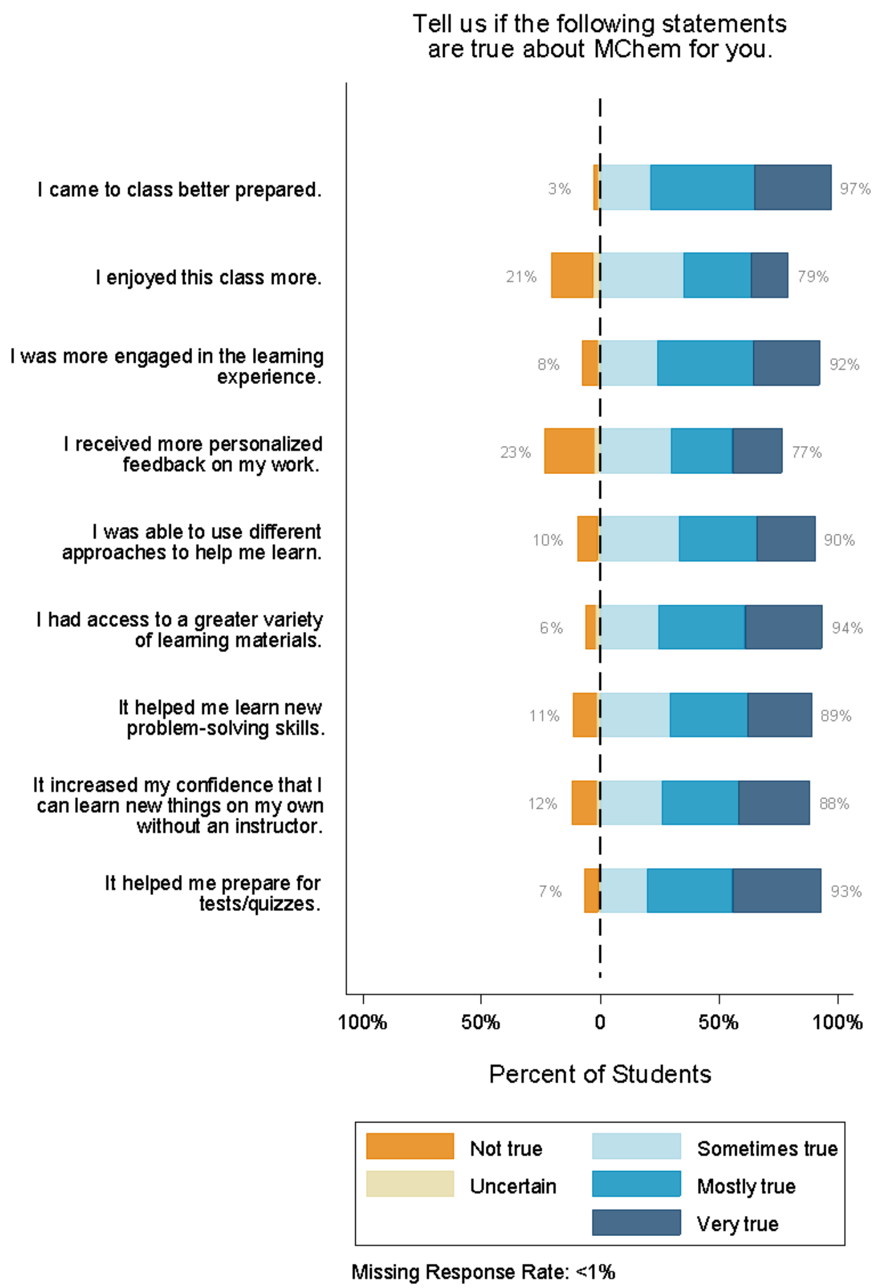
Note: The percentages shown to the left of 0% indicate not applicable, no, or less use of the features; the percentages to the right of 0% indicate greater use of the features.

Figure 3: Students' self-reported experiences with Mastering Chemistry features



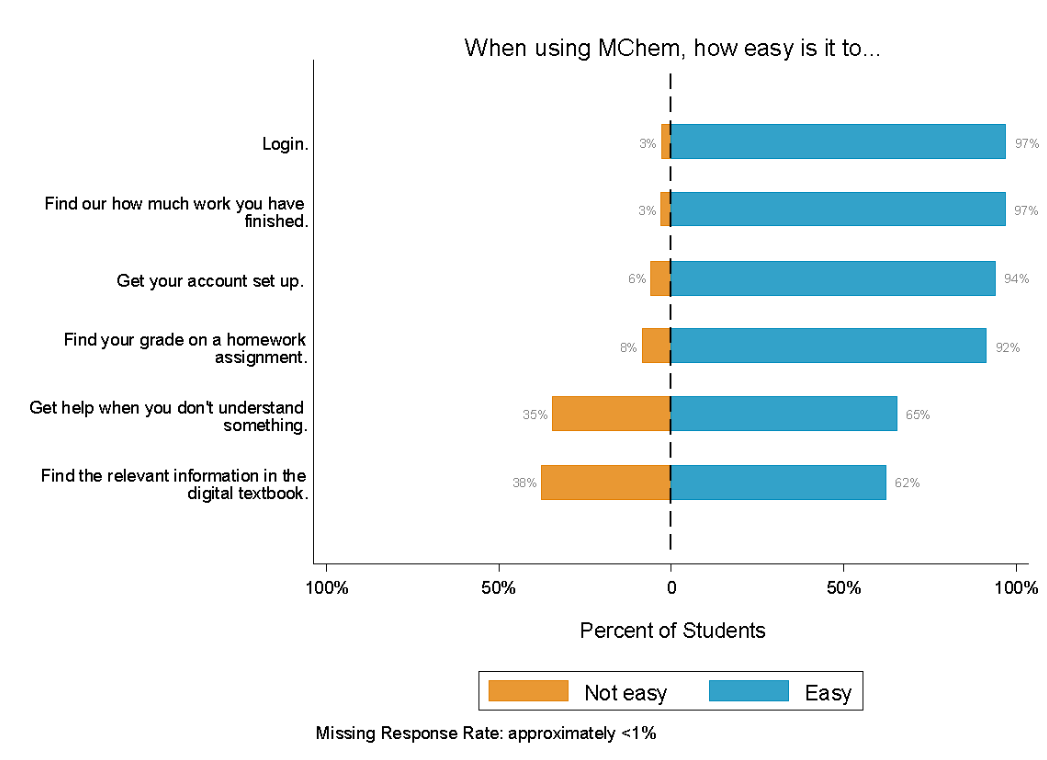
Note: The percentages shown to the left of 0% indicate the students did not use the feature or did not find the feature helpful; the percentages to the right of 0% indicate the feature was at least somewhat helpful.

Figure 4: Students' self-reported overall experiences with Mastering Chemistry



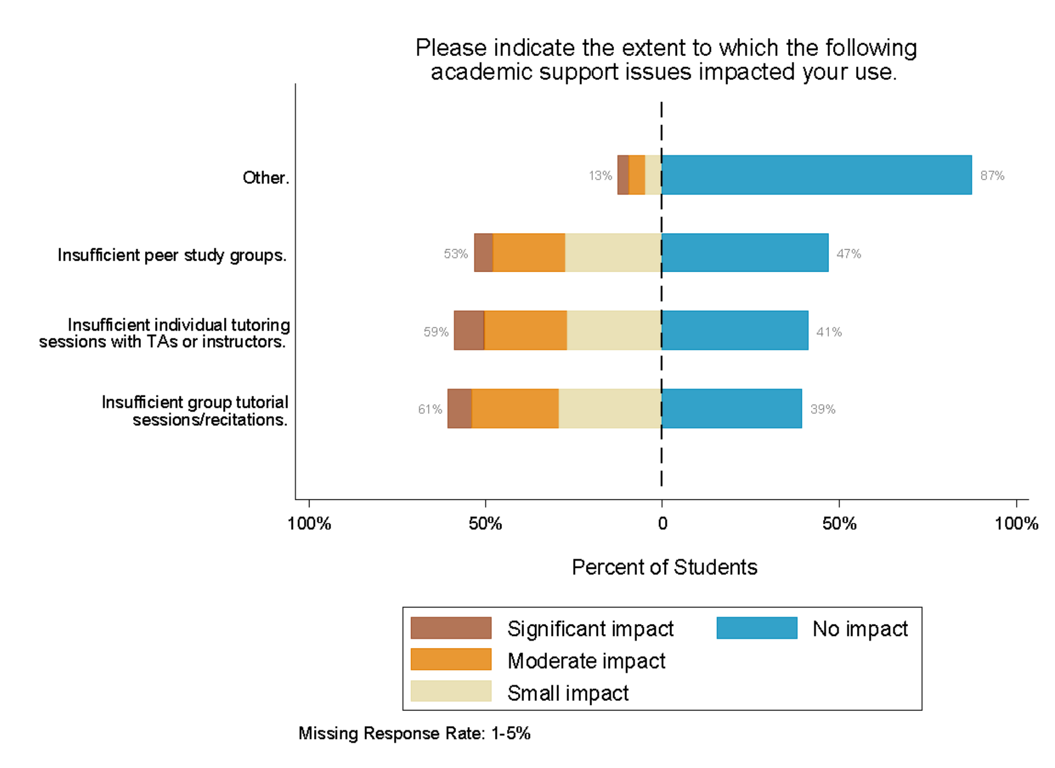
Note: The percentages shown to the left of 0% indicate that students were uncertain about how they felt about the statement of the potential benefit of Mastering Chemistry use or disagreed with the statement; the percentages to the *right of 0% indicate that students agreed with the statement of the potential benefit of Mastering Chemistry use at least "sometimes".*

Figure 5: Ease of use of different Mastering Chemistry functions



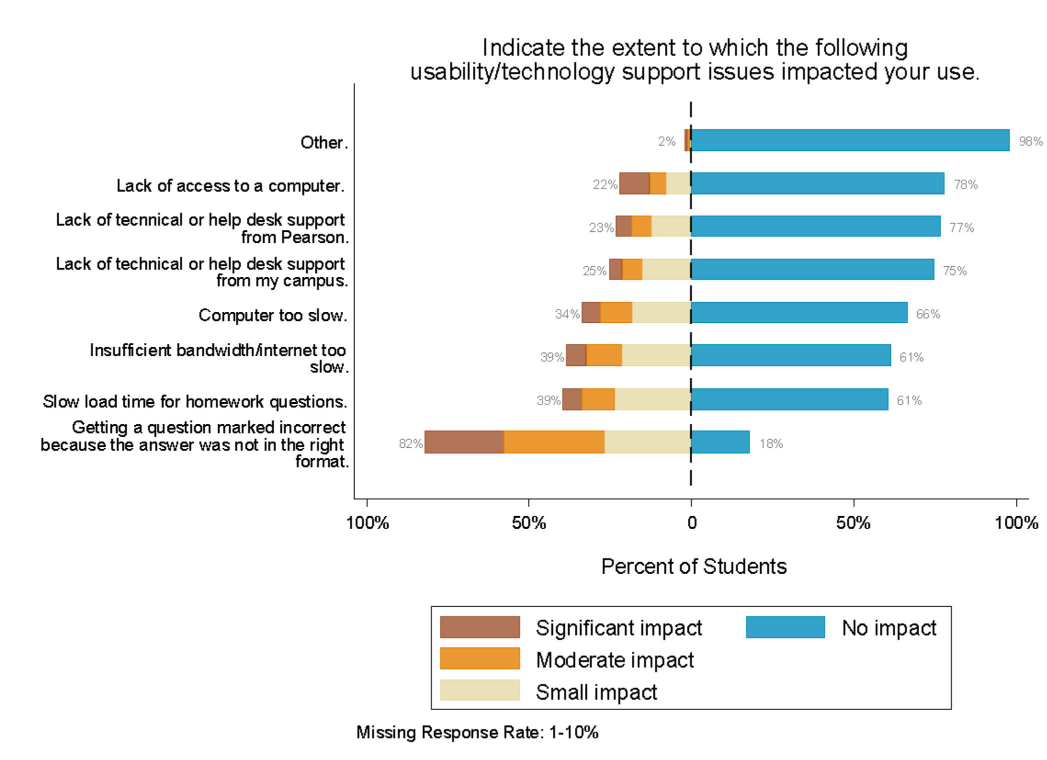
Note: The percentages shown to the left of 0% indicate that students found that particular action not easy to complete; the percentages to the right of 0% indicate that students found that particular action easy to complete.

Figure 6: Effects of different academic support issues on limiting Mastering Chemistry use



Note: The percentages shown to the left of 0% indicate that students found the factor to have some degree of impact on their use of Mastering Chemistry; the percentages to the right of 0% indicate that students found the factor to have no impact on their use of Mastering Chemistry.

Figure 7: Effects of usability/technology issues on Mastering Chemistry use



Note: The percentages shown to the left of 0% indicate that students found the factor to have some degree of impact on their use of Mastering Chemistry; the percentages to the right of 0% indicate that students found the factor to have no impact on their use of Mastering Chemistry.

Impact Analysis

Table 3 shows the simple summary statistics regarding Mastering Chemistry use. Appendix E presents additional use information. Students completed problems for homework, extra practice, tests, and adaptive learning, and used Mastering Chemistry on about half of the available days. Students spent approximately 30 hours on the system and attempted about 535 and 452 unique problems for fall and spring respectively (potentially more than once) on average. The Time Spent on Mastering Chemistry variable includes both the time students were actively engaged in the courseware and student idle time (for instance, when students were not present at their computer or were conversing with another student while still logged in to Mastering Chemistry). Therefore, this variable cannot be directly interpreted as a *time-on-task* measure. The study provided no evidence to suggest that the amount of idle time was either large or small relative to the amount of active engagement time.

Table 3: summary statistics for Mastering Chemistry use

	Fall 2016 (n = 677)		Spring 2017 (n = 1311)	
Variable	**Mean**	***SD***	**Mean**	***SD***
Unique Days	50.81	13.67	41.86	12.41
Percentage of Class Days Used	52%	14%	43%	13%
Time Spent on Mastering Chemistry (hours)	31.16	11.38	28.90	11.87
Unique Mastering Chemistry Assignments	71.84	26.26	60.22	29.37
Unique Mastering Chemistry Problems	533.71	155.83	452.10	151.04

Comparative Impact Analysis

The fall semester impact analytic sample consisted of 1,700 students (742 treatment students and 958 comparison students). The spring semester impact analytic sample consisted of 1,137 students (782 treatment students and 355 control students). For students included in the analysis, complete information was available for prior achievement in Math (ACT or SAT scores), treatment condition (treatment or comparison), age, gender, and section ID (some instructors taught more than one section of the course). The analytic sample for OSU spring semester was limited to students who were also enrolled in General Chemistry I because we required linking to fall student survey data to obtain student age information. We conducted a missingness analysis and found no significant source of bias between the analytic sample for spring 2017 (1,137 students total) and the maximal enrollment sample of spring 2017 students (2,004 students total).

Two-level regression models were used in the analyses, with students clustered within sections. Student-level covariates in the model included prior achievement, age, and gender. We examined the effect of Mastering Chemistry use on three dependent variables in the analyses: final exam scores, course grades, and a binary course pass/fail (P/F) indicator. Final exam scores were the scores students received from their instructor on their final exam. Course grades were also assigned by the instructor and represented combined scores for students' coursework (including homework, midterms, and final exams). We excluded a student's course grade from our analysis if the student did not take the final exam. The binary course P/F indicator had two possible values—"pass" or "fail/withdraw"—with "pass" defined as a course grade of C– or higher.

To place the dependent variables in context, SRI reviewed selected course syllabi of both treatment and comparison classrooms. Performance in Mastering Chemistry and Macmillan Learning's Sapling chemistry courseware – that is, "online homework" – regularly comprised between 7.5% and 15% of students' course grades. Similarly, students' final exam scores comprised between 25% and 27.5% of students' course grades. Sapling was used as a homework supplement outside of class time.

Treatment and Comparison Conditions

Separate analyses were conducted for students from fall semester 2016 and spring semester 2017. All fall students in the treatment (Mastering Chemistry users) and comparison conditions were enrolled in the General Chemistry course during fall semester 2016. All spring students in the treatment (Mastering Chemistry users) and comparison conditions were enrolled in the General Chemistry II course during spring semester 2017. Fall 2016 students in the treatment condition were in sections taught by two instructors, whereas students in the comparison condition were in sections taught by four instructors. Spring 2017 students in the treatment condition were in sections taught by three instructors, whereas students in the comparison condition were in sections taught by one instructor. The comparison condition consisted of students using Sapling.

Table 4 provides a summary of descriptive statistics for a series of student-level variables for students in the treatment and comparison groups.

Table 4: Descriptive statistics for students in the treatment and comparison groups

	Fall 2016			Spring 2017		
Variables	**Mastering Chemistry Mean (*SD*)**	**Non-Mastering Chemistry Mean (*SD*)**	**Absolute Effect Size (|ES|)**	**Mastering Chemistry Mean (*SD*)**	**Non-Mastering Chemistry Mean (*SD*)**	**Absolute Effect Size (|ES|)**
Sample size	742	958		782	355	
Mean ACT	29.33 (3.20)	29.06 (3.19)	0.084	29.60 (2.96)	29.81 (2.92)	0.071
Mean Age	19.61 (0.97)	19.60 (0.92)	0.013	19.49 (0.81)	19.43 (0.73)	0.071
Percentage Female	55%	63%	0.173	62%	55%	0.181

Note: The Mastering Chemistry user/non-Mastering Chemistry user statistics were considered "not substantively different" if the effect size was smaller than 0.25. All effect sizes were computed as Hedges' g.

As shown in Table 4, none of the prior achievement or student background characteristics differed substantively between the non-Mastering Chemistry students and Mastering Chemistry students for fall 2016 or spring 2017 sample. Thus the between-group baseline equivalence was sufficient to conduct the comparative impact analysis without the use of statistical matching techniques.[2]

Model Specification

Three dependent variables were modeled separately using multilevel models, with students at level 1 nested within course sections at level 2. We used a Hierarchical Linear Model (HLM), when the dependent variable was continuous (final exam scores and course grades); and a Hierarchical Generalized Linear Model (HGLM), when the dependent variable was dichotomous (binary grades). In both models, we included a set of baseline covariates to account for group nonequivalence (student's prior achievement, age, and gender) and improve the precision of the estimated impact. Note that for the spring 2017 sample, a convergence issue arose with the hierarchical model in examining binary grades; therefore, we performed a standard single-level logistic regression instead.

Results

Tables 5 through 10 show the impact model results for the three dependent variables: final exam scores, total course grade, and binary pass/fail. For the fall 2016 analysis with final exam scores as the dependent variable, the effect size for the treatment versus control comparison was 0.17, indicating a small positive treatment effect. An effect size of 0.17 is associated with an improvement index of 7%[3], indicating that the intervention would have led to a 7% increase in percentile rank for average students in the comparison group if those students were in sections using Mastering Chemistry; correspondingly, 57% of the students in sections using Mastering Chemistry scored above the mean score in the comparison group. For the spring 2017 analysis with final exam scores as the dependent variable, the effect size for the treatment versus control comparison was 0.10, indicating a small positive treatment effect. An effect size of 0.10 is associated with an improvement index of 4%, indicating that the

[2] Following What Work Clearinghouse guidelines, baseline group equivalence is acceptable (that is, does not require statistical matching to reduce differences) if the difference is less than an effect size of 0.25 for key baseline characteristics. If the difference for one or more baseline characteristics is greater than an effect size of 0.05, the model must include a covariate control for the baseline characteristic.

[3] We provide effect sizes as both Hedges' *g* and improvement index in alignment with What Works Clearinghouse v4.0 recommendations. For more information regarding the interpretation and calculation of effect sizes, please see the What Works Clearinghouse v4.0 Procedures Handbook, Appendices A and E: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_handbook_v4.pdf

intervention would have led to a 4% increase in percentile rank for average students in the comparison group if those students were in sections using Mastering Chemistry; correspondingly, 54% of the students in sections using Mastering Chemistry scored above the mean score in the comparison group.

For the fall 2016 analysis, when the total course grade was the dependent variable in the model, the estimated effect size was 0.12, again indicating a small positive treatment effect and an improvement index of 5%. For the spring 2017 analysis, when the total course grade was the dependent variable in the model, the estimated effect size was 0.53, indicating a substantial positive treatment effect and an improvement index of 20%; correspondingly, 70% of the students in sections using Mastering Chemistry scored above the mean score in the comparison group.

Finally, for the fall 2016 analysis for binary P/F grade as the dependent variable, the estimated effect size was 0.06 and the improvement index was 2%. For the spring 2017 analysis for binary P/F grade as the dependent variable, the estimated effect size was 0.43 and the improvement index was 17%; correspondingly, 67% of the students in sections using Mastering Chemistry scored above the mean score in the comparison group.

In addition, given that the effect of the intervention was assessed on multiple outcome measures (final exam scores, course grades, binary grades), Benjamini-Hochberg (BH) method was also conducted by SRI to correct for multiple comparisons (control for Type I error). The significance results are identical to the ones without multiple comparison adjustment reported in the tables.

Pearson

Table 5: Fall 2016 HLM results when comparing Mastering Chemistry students and non-Mastering Chemistry students for final exam scores (on a 100-point scale)

Solution for Fixed Effects						
Effect	*B*	*SE*	*p*-value	Effect size	95% Confidence Interval [min, max]	Improvement Index
Fixed effects						
Condition: treatment	**2.52**	**0.63**	**< .0001**	**0.17**	**[0.08, 0.27]**	**7%**
Condition: control	**0.00**	**-**	**-**			
Prior achievement (ACT)	2.23	0.10	< .0001			
Age	-0.97	0.33	.0030			
Gender female	-2.11	0.62	.0006			
Gender other	2.39	3.59	.5055			
Gender male	0.00	-	-			
Random effects						
Level 1 intercept	34.30	7.69	< .0001			
Level 2 intercept	0.73	1.38	.2974			

*Note: n = 1,700 in the HLM analysis. All effect sizes were on Hedges' g. The intraclass correlation coefficient (ICC) was 4.18%, and the R-squared at the student level was 28.09%. SRI also conducted an analysis based on a larger sample (n = 2,475) without controlling for student demographic information; that analysis yielded very similar results to the ones shown in Table 5. Interaction terms were added to the statistical model one by one (Condition*ACT, Condition*Age, and Condition*Gender), but none of them was significant. Accordingly, SRI removed the interaction terms from the final model.*

Table 6: Spring 2017 HLM results when comparing Mastering Chemistry students and non-Mastering Chemistry students for final exam scores (on a 100-point scale)

Solution for Fixed Effects						
Effect	*B*	*SE*	*p*-value	Effect Size	95% Confidence Interval [min, max]	Improvement Index
Fixed effects						
Condition: treatment	**1.36**	**0.83**	**.1007**	**0.10**	**[-0.02, 0.23]**	**4%**
Condition: control	**0.00**	**-**	**-**			
Prior achievement (ACT)	1.77	0.12	< .0001			
Age	-0.75	0.45	.0959			
Female	-3.87	0.71	< .0001			
Male	0.00	-	-			
Random effects						
Level 1 intercept	43.48	10.26	< .0001			
Level 2 intercept	2.29	1.97	.1232			

*Note: n = 1,137 in the HLM analysis. All effect sizes were on Hedges g. The ICC was 3.64%, and the R-squared at the student level was 19.35%. SRI also conducted an analysis based on a larger sample (n = 1,890) without controlling for student demographic information; that analysis yielded very similar results to the ones shown in Table 6. Interaction terms were added to the statistical model one by one (Condition*ACT, Condition* Age, and Condition*Gender), but none was statistically significant. Accordingly, SRI removed the interaction terms from the final model.*

Table 7: Fall 2016 HLM results when comparing Mastering Chemistry students and non-Mastering Chemistry students for course grades (on a 100-point scale)

Solution for Fixed Effects						
Effect	*B*	*SE*	*p*-value	Effect size	95% Confidence Interval [min, max]	Improvement Index
Fixed effects						
Condition: treatment	**1.33**	**0.57**	**.0188**	**0.12**	**[0.03, 0.22]**	**5%**
Condition: control	**0.00**	**-**	**-**			
Prior achievement (ACT)	1.63	0.07	<.0001			
Age	-0.99	0.25	<.0001			
Gender female	-0.67	0.46	.1462			
Gender other	2.34	2.68	.3823			
Gender male	0.00	-	-			
Random effects						
Level 1 intercept	51.72	5.78	<.0001			
Level 2 intercept	2.85	1.14	.0061			

*Note: n = 1,700 in the HLM analysis. All effect sizes were on Hedge's g. The ICC was 6.00%, and the R-squared at the student level was 27.12%. SRI also conducted analysis based on a larger sample (n = 2,475), without controlling for student demographic information; that analysis yielded very similar results to the ones in Table 7. Interaction terms were added to the statistical model one by one (Condition*ACT, Condition*Age, and Condition*Gender), but none of them was significant. Accordingly, SRI removed the interaction terms from the final model.*

Table 8: Spring 2017 HLM results when comparing Mastering Chemistry students and non-Mastering Chemistry students for course grades (on a 100-point scale)

Solution for Fixed Effects						
Effect	*B*	*SE*	*p*-value	Effect Size	95% Confidence Interval [min, max]	Improvement Index
Fixed effects						
Condition: treatment	**5.33**	**0.64**	**< .0001**	**0.53**	**[0.41, 0.66]**	**20%**
Condition: control	**0.00**	**-**	**-**			
Prior achievement (ACT)	1.33	0.09	< .0001			
Age	-0.83	0.35	.0185			
Gender female	-1.37	0.55	.0136			
Gender male	0.00	-	-			
Random effects						
Level 1 intercept	54.26	7.97	< .0001			
Level 2 intercept	1.44	1.23	.1193			

*Note: n = 1,137 in the HLM analysis. All effect sizes were on Hedge's g. The ICC was 7.82%, and the R-squared at the student level was 21.38%. SRI also conducted an analysis based on a larger sample (n = 1,890) without controlling for student demographic information; that analysis yielded very similar results to the ones shown in Table 8. Interaction terms were added to the statistical model one by one (Condition*ACT, Condition* Age, and Condition*Gender), but none was statistically significant. Accordingly, SRI removed the interaction terms from the final model.*

Table 9: Fall 2016 HGLM results when comparing Mastering Chemistry students and non-Mastering Chemistry students for binary grades (P/F/withdrawal) obtained for the course

Solution for Fixed Effects						
Effect	**B**	***SE***	***p*-value**	**Effect size**	**95% Confidence Interval [min, max]**	**Improvement Index**
Fixed effects						
Condition: treatment	**7.53**	**3.07**	**.0142**	**0.06**	**[-0.03, 0.16]**	**2%**
Condition: control	**0.00**	**-**	**-**			
Prior achievement (ACT)	0.29	0.03	< .0001			
Age	-0.05	0.11	.6517			
Condition*Age	-0.38	0.15	.0138			
Random effects						
Level 1 intercept	-4.68	2.45	.0583			
Level 2 intercept	0.04	0.13	.3752			

*Note: n = 1,706 in the HGLM analysis. All effect sizes were on Hedges' g. The model did not converge when prior academic achievement and student demographics were included in the model. There was no significant difference for gender when age was excluded from the model. However, there was a significant difference for age when gender was excluded from the model. Thus, gender was excluded from the final model. SRI also conducted an analysis of a larger sample (n = 2,684), without controlling for student demographic information, and the analysis yielded very similar results. Interaction terms were added to the statistical model one by one (Condition*ACT and Condition*Age), but only Condition*Age was significant. Accordingly, SRI included Condition*Age in the final model.*

Table 10: Spring 2017 logistic regression results when comparing OSU Mastering Chemistry students and non-Mastering Chemistry students for binary grades (P/F/withdrawal) obtained for the course

Estimates						
Effect	*B*	*SE*	*p*-value	Effect Size	95% Confidence Interval [min, max]	Improvement Index
Condition: treatment	**0.36**	**0.13**	**.0059**	**0.43**	**[0.31, 0.56]**	**17%**
Condition: control	**0.00**	**-**	**-**			
Prior achievement (ACT)	0.20	0.04	< .0001			
Age	-0.26	0.12	.0257			
Female	-0.05	0.13	.6893			
Male	0.00	-	-			
Intercept	2.19	2.83	.4382			

*Note: There was a convergence issue when hierarchical models were performed. Thus, a single-level model was performed instead. n = 1,142 in the logistic regression analysis. All effect sizes were on Hedges g. SRI also conducted an analysis based on a larger sample (n = 1,920) without controlling for student demographic information; that analysis yielded very similar results to the ones shown in Table 10. Interaction terms were added to the statistical model one by one (Condition*ACT, Condition* Age, and Condition*Gender), but none was statistically significant. Accordingly, SRI removed the interaction terms from the final model.*

Examining the Relationship Between Number of Problem Attempts and Outcomes

For this analysis, we examined the number of problems attempted in Mastering Chemistry as a measure of student Mastering Chemistry usage. The relationships between the number of problems attempted and three dependent variables (final exam scores, course grades, and binary grades) were modeled separately using multilevel models—HLM for final scores and total scores (continuous variables), and HGLM for binary grades (dichotomous variable). Note that for the spring 2017 sample, a convergence issue arose with the hierarchical model in examining binary grades; therefore, we performed a standard single-level logistic regression instead. The analytical models included ACT scores, age, and gender as covariates. Tables 11 and 12 show the results of the models for each term.

For the fall 2016 analysis of the number of problems attempted (Table 11), the estimated B parameter represents the change in the outcome score for a one-unit change in the number of problems attempted, controlling for other independent variables in the model. Specifically, each problem attempted was associated with a 0.038-point increase in the final exam score, which translates to a 1-point increase in the final exam score for every 26 problems attempted. Each problem attempted was also associated with a 0.032-point increase in the course grade, which translates to a 1-point increase in the final exam score for every 31 problems attempted. For binary grades, each problem attempted was associated with an increase of 0.0083 log odds of passing the course or 0.83 log odds for an additional 100 problems attempted.

For the spring 2017 analysis of the number of problems attempted (Table 12), the estimated B parameter represents the change in the outcome score for a one-unit change in the number of problems attempted, controlling for other independent variables in the model. Specifically, each problem attempted was associated with a 0.026-point increase in the final exam score, which translates to a 1-point increase in the final exam score for every 38 problems attempted. Each problem attempted was also associated with a 0.031-point increase in the course grade, which translates to a 1-point increase in the final exam score for every 32 problems attempted. For binary grades, each problem attempted was associated with an increase of 0.0079 log odds of passing the course or 0.79 log odds for an additional 100 problems attempted.

Table 11: Fall 2016 HLM and HGLM results for the relationship between the number of problems attempted and student achievement measures

	Solution for Fixed Effects											
	Final Exam Scores				Course Grades				Binary Grades			
Variable	***n***	**B**	***SE***	***p*-value**	***n***	**B**	***SE***	***p*-value**	***n***	**B**	***SE***	***p*-value**
Fixed effects	623				623				628			
Number of problems attempted		**0.038**	**0.003**	**< .0001**		**0.032**	**0.002**	**< .0001**		**0.008**	**0.001**	**< .0001**
Prior achievement (ACT)		1.821	0.142	< .0001		1.430	0.102	< .0001		0.322	0.057	< .0001
Age		-0.634	0.460	0.169		-0.453	0.328	0.1685				
Gender female		-3.572	0.877	< .0001		-2.434	0.622	.0001				
Gender other		-4.838	3.817	0.206		-3.291	2.704	0.2241				
Gender male		0.000	-	-		0.000	-	-				
Random effects												
Level 1 intercept		22.965	10.998	0.043		32.718	7.862	0.0002		-10.236	1.667	< .0001
Level 2 intercept		2.191	2.194	0.159		5.131	2.046	0.0061		0.108	0.296	0.3579

Note: For final exam scores, the ICC was 4.92%, and the R-squared at the student level was 42.67%. For course grades, the ICC was 8.07%, and the R-squared at the student level was 47.80%. For binary grades, SRI excluded age and gender from the analysis for model convergence.

Note: The estimated B parameter in the table represents the change in the outcome score for a one-unit change in the number of problems attempted controlling for other independent variables in the model.

Table 12: Spring 2017 HLM and Logistic Regression results for the relationship between the number of problems attempted and student achievement measures

Solution for Fixed Effects												
	Final Exam Scores				Course Grades				Binary Grades			
Variable	*n*	B	*SE*	*p*-value	*n*	B	*SE*	*p*-value	*n*	B	*SE*	*p*-value
Fixed effects	774				774				774			
Number of problems attempted		**0.026**	**0.003**	**< .0001**		**0.031**	**0.002**	**< .0001**		**0.008**	**0.001**	**< .0001**
Prior achievement (ACT)		1.844	0.140	< .0001		1.340	0.099	< .0001		0.286	0.062	< .0001
Age		-0.359	0.499	0.471		-0.450	0.356	0.2065		-0.196	0.162	.2264
Gender female		-4.781	0.825	< .0001		-2.273	0.587	< .0001		-0.070	0.195	.7191
Gender male		0.000	-	-		0.000	-	-		0.000	-	-
Random effects												
Level 1 intercept		23.359	11.571	0.049		37.799	8.228	< .0001		-4.498	3.992	.2599
Level 2 intercept		7.498	3.313	0.012		1.17	1.096	0.1423				

Note: For final exam scores, the ICC was 3.19%, and the R-squared at the student level was 26.80%. For course grades, the ICC was 2.13%, and the R-squared at the student level was 36.29%. For binary grades, SRI performed single-level logistic regression for model convergence.

Note: The estimated B parameter in the table represents the change in the outcome score for a one-unit change in the number of problems attempted controlling for other independent variables in the model.

Cost Analysis

To calculate cost, SRI collected information on the cost impacts of implementing and using Mastering Chemistry compared with other conditions.

Data Collection and Analysis Approach

SRI collected cost information through the instructor survey, site visit, and phone interviews. These data were entered into a template that captured possible cost impacts for setting up and delivering courses with Mastering Chemistry, both in the initial year and on an ongoing basis. We took an "ingredients" approach, which identifies all inputs regardless of who bears the costs. Our analyses primarily focused on costs of providing instruction, although we also considered differences in access costs for students. Costs of instructor time took into account the staffing mix (tenure-track faculty, nontenure-track faculty, and TAs) and factored in average total compensation for different types of staff. Total costs in each condition were divided by the number of student to calculate a cost per student, and any difference in cost per student was then divided by the effect size to generate a cost per unit of improvement (or decrement) relative to the comparison condition.

Findings

Because the comparison condition used online courseware in a similar use model, there were few differences in implementation costs. Specific findings from the cost analysis follow:

- No technology or infrastructure costs were incurred specifically for use of Mastering Chemistry.
- Instructors reported spending more time on setting up of Mastering Chemistry versus Sapling, including system setup, review and adjustment of product content, and professional development. We estimate these differences at roughly 20 hours per year.
- The only area Sapling instructors reported spending more time than Mastering Chemistry instructors was in communicating with developers, but that difference was very small (approximately two hours per instructor per year).
- For instruction-related activities, Sapling instructors, on average, spent an estimated seven fewer hours per year compared with Mastering Chemistry instructors.
- Subscription costs for Mastering Chemistry, in addition to the cost of textbooks, were mentioned as a possible burden for students from low-income backgrounds. A course like General Chemistry is relatively expensive, given the costs associated with textbooks, lab fees, and in some cases lab coats. In this case, because license fees for both Mastering Chemistry and Sapling are $180 per semester, access costs for students did not differ.

Overall instructors reported spending approximately 17 more hours per year using Mastering Chemistry compared with the time spent using Sapling, amounting to about 1–1.5% of time for an instructor working full time. Our calculations of the average cost per hour of instructor time indicated that the additional time spent would amount to a cost difference of $0.67 per student. This translates to $0.27 per 1-point increase in their final score for fall 2016 students; $0.49 per 1-point increase in their final score for spring 2017 students; $0.50 per 1-point increase in their course grade for fall 2016 students; $0.13 per 1-point increase in their course grade for spring 2017 students.

The differences in instructor time and cost between Mastering Chemistry and Sapling may be due to the degree of active support provided by the respective product teams. Mastering Chemistry instructors were required to make any customizations to the platform themselves, whereas Sapling instructors emailed the Sapling product team with requests for customizations. The Sapling product team would then make the respective changes to the system, thus reducing the total amount of time Sapling instructors spent on product customization.

Table 13: Cost analysis

Cost Element	Value
Time difference per instructor per year (in hours)	16.57
Average instructor compensation per year	$102,940
Cost per hour of instructor time	69.48
#of instructions	6
Total cost difference in instructor time	$3,453.96
# students (total for 6 courses)	2,590
Cost difference per student	$1.33
Instructor hours per year	1,120
% of instructor hours	1.5%
Cost difference per semester per student	$0.67

Note: cost calculations involve all six instructors who taught General Chemistry I and II courses over the academic year, including one instructor who taught a General Chemistry II course in the fall that was not included in impact analyses. Assumes major reviews of the product take place every three to four years. For the average compensation calculation we substituted an associate professor compensation for a retired full professor compensation as the latter was a significant outlier. We assume 140 working days for nine months (75 days per semester based on school calendar, minus 10 days of holiday in the year).

Limitations

This study employed a non-equivalent comparison group quasi-experimental research design. The analytic samples of this study for students met What Works Clearinghouse standards for baseline equivalence with appropriate statistical controls, including an independent academic achievement outcome (ACT scores), supporting the relevance of the findings. (Baseline equivalence was not established for instructors.) However, while the findings include many positive predictive effects of Mastering Chemistry use on student achievement, the results cannot be interpreted as evidence of a causal relationship between Mastering Chemistry use and student achievement given the study design.

There was significant attrition in the Fall 2016 semester as measured by students removed from the maximal student samples (that is, the number of students in the Fall 2016 term with course outcome data) in building the analytic samples (the number of Fall 2016 students with the required variable data for analysis). This was due to the way the student survey was administered and the ability to link the student surveys with administrative data. For Fall 2016, the overall attrition rate was 37% with a differential attrition rate of 11%. With respect to the Spring 2017 semester, students were only eligible to be included in the study if they had completed the student survey in the Fall semester and provided essential demographic information. While there was a significant gap between the number of enrolled General Chemistry students in Spring 2016 and the analytic sample, this was to be expected. Attrition among eligible students was 31%, with a differential attrition rate of 11%. The creation of the analytic samples for both terms is detailed in the Analytical Samples for Estimating Impacts section.

Given the small number of instructors involved in the study, the results presented above may be explained by differences among the instructors (such as experience or pedagogy) who used Mastering Chemistry and those that did not, rather than the use of Mastering Chemistry alone. SRI explored possible teacher confounding effects that may have influenced these findings. (Table F1 and F2 in Appendix F include a set of descriptive statistics for various background characteristics that were available about the instructors in the study.) In the fall term, instructors for comparison sections had 17 years of experience teaching college courses on average compared to 12.5 for Mastering Chemistry sections, and there are some differences in academic rank. (This information is not available for the spring term.) Another difference is that in the fall, Mastering Chemistry instructors had more experience using the product than instructors in the comparison condition had in using Sapling. The two Mastering Chemistry instructors had five to six years and nine to ten years of experience, respectively, in using Mastering Chemistry, whereas all the control instructors had between one and four years of experience in using Sapling. Thus, the instructors' degree of experience in using the respective courseware may explain at least a portion of the difference in course achievement outcomes.

In the instructor survey, however, Sapling instructors all reported feeling very prepared to use the product.

Finally, the maximal sample of this study has a total of six instructors, meaning there is insufficient statistical power to support a three-level modeling approach to control for possible instructor clustering effects. This may result in underestimated standard errors and overestimated statistical significance tests (that is, artificially small p-values).

Finally, a limitation to the cost analysis is that the majority of the data were gathered through retrospective instructor interviews, requiring instructors to estimate how much time they spent performing various tasks rather than actively tracking time throughout an instructional term. Thus, the accuracy of the calculations are dependent upon instructors’ recollections. Additionally, the cost analysis is dependent upon the specific staffing model, instructor compensation numbers, etc. for this site; it is possible that the cost analysis would arrive at different results at other institutions.

Discussion

Mastering Chemistry is primarily used to assign, support and grade homework in General Chemistry I and II at OSU. Mastering Chemistry is also used to encourage students to come to class better prepared and, in one case, to “flip” the classroom using interactive features during class time. The system also enables instructors to give students more frequent assignments and to monitor the time it takes to complete problems. Instructors report that Mastering Chemistry provides students with immediate feedback on homework assignments, whereas before Mastering Chemistry homework was not graded.

Students for their part rated Wrong Answer Feedback as by far the most frequently used feature, followed by Hints and Videos. Students reported that Mastering Chemistry helped them come to class better prepared, have access to a greater variety of learning materials, and prepare for tests/quizzes. They were less likely to report that Mastering Chemistry enabled them to receive more personalized feedback or that it increased their enjoyment of the course.

When examining impact on student outcomes for fall 2016, for all three outcome measures examined—final exam scores, course grades, and binary grades (P/F)—students using Mastering Chemistry significantly outperformed Sapling users, when controlling for student prior achievement, age, and gender, as well as accounting for the section-level effects. Similarly, for spring 2017, for two of the three outcome measures examined—course grades and binary grades—students using Mastering Chemistry significantly outperformed Sapling users, when controlling for the same factors.

The findings indicate a strong association between Mastering Chemistry use and improved student chemistry performance. When controlling for student prior achievement and demographic characteristics, the number of problems attempted and all three student course outcome measures (final scores, total scores, and binary grades) were statistically significant.

We cannot interpret these results as providing causal evidence of effectiveness given limitations in the study design. In particular, we were not able to control for instructor characteristics due to the small number of instructors involved in the study. Thus, we cannot rule out the possibility that differences in student outcomes were explained by differences among instructors and not by use of Mastering Chemistry alone. Additionally, the study is limited by high levels of attrition in student participants due to challenges matching various datasets. Future studies could increase the confidence of these findings by increasing the number of instructors involved.

We found that use of Mastering Chemistry took marginally more time for instructors than the comparison condition using Sapling. We found slightly higher costs in terms of the time it took instructors to set up Mastering Chemistry, to review and adjust problems, and to participate in professional development. There were no material differences in weekly delivery time use, and the license costs for students were the same. Overall we identified a modest difference of $.67 per student per semester for use of Mastering Chemistry, driven by instructor time.

References

Adamson, David, Gregory Dyke, Hyeju Jang, and Carolyn Penstein Rosé. 2014. "Towards an Agile Approach to Adapting Dynamic Collaboration Support to Student Needs." *International Journal of Artificial Intelligence in Education* 24 (1): 92–124. doi:10.1007/s40593-013-0012-6.

Durlach, Paula J, and Jessica M Ray. 2011. "Designing Adaptive Instructional Environments: Insights from Empirical Evidence (Technical Report 1297)." Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Kulik, James A, and J D Fletcher. 2015. "Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review." *Review of Educational Research*, 1–37. doi:10.3102/0034654315581420.

Ma, Wenting, O. O. Adesope, J. C. Nesbit, and Qing Liu. 2014. "Intelligent Tutoring Systems and Learning Outcomes: A Meta-Analysis." *Journal of Educational Psychology* 106 (4): 901–18. doi:10.1037/a0037123.

McLaren, Bruce M., and Seiji Isotani. 2011. "When Is It Best to Learn with All Worked Examples?" In Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), edited by Gautam Biswas, Susan Bull, Judy Kay, and Antonija Mitrovic, 6738:222–29. Berlin Heidelberg: Srpinger. doi:10.1007/978-3-642-21869-9_30.

McLaren, Bruce M., Krista E. Deleeuw, and Richard E. Mayer. 2011. "Polite Web-Based Intelligent Tutors: Can They Improve Learning in Classrooms?" *Computers and Education* 56 (3): 574–84. doi:10.1016/j.compedu.2010.09.019.

Nesbit, John C., Olusola O. Adesope, Qing Liu, and Wenting Ma. 2014. "How Effective Are Intelligent Tutoring Systems in Computer Science Education?" In *2014 IEEE 14th International Conference on Advanced Learning Technologies*, 99–103. doi:10.1109/ICALT.2014.38.

Steenbergen-Hu, Saiying, and Harris Cooper. 2013. "A Meta-Analysis of the Effectiveness of Intelligent Tutoring Systems on K–12 Students' Mathematical Learning." *Journal of Educational Psychology* 105 (4): 970–87. doi:10.1037/a0032447

Steenbergen-Hu, Saiying, and Harris Cooper. "A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning." *Journal of Educational Psychology* 106.2 (2014): 331. Pecjak, Regan. "Q&A with Candace Thille, 'Adaptive Learning' Expert." *The Stanford Daily*, March 9, 2016. Issue 1479. http://www.thestanforddaily.com.

VanLehn, Kurt. 2011. "The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems." *Educational Psychologist* 46 (4): 197–221. doi:10.1080/00461520.2011.611369.

Appendix A: Data management (data cleaning and linking)

Data files

To conduct the analyses, SRI analysts worked with 10 data files from three sources (see Table A1).

Table A1: source and data files used in data analyses

SRI
Pre-student survey (fall)
Post-student survey (spring)
Teacher survey (spring)
Institutions of higher education
Institutional student linking file
Course outcome file
Pearson
Use data student linking file
Courseware use file
Assignment data file
Problems data file
Dynamic module data file

Data cleaning

To prepare the data files for merging and analysis, we conducted three data cleaning functions. First, students were excluded from the sample if they were enrolled in an honors chemistry class (the focus of the study was on students enrolled in the regular General Chemistry course). Second, if students did not provide consent to participate in the study on either the student pre- or post-survey or if they reported an age under 18, they were excluded from the sample. Third, duplicate and triplicate records associated with the same student ID were dealt with on a case-by-case basis. Every attempt was made to distinguish which of the entries could be retained given the data provided, with the other entries deleted. For OSU, we also had to exclude students of instructors who did not participate in the study. The initial course outcome file provided by OSU included students associated with 12 instructors. The study included only six instructors teaching regular General Chemistry in the 2016-17 academic year.

Table A2 shows the primary cleaning steps taken for the data files and the number of observations lost per cleaning step.

Table A2: Data cleaning steps and number of records excluded from the analytical sample for each datafile and step

Data File	Initial N	Data Cleaning Step	Dropped N	Cleaned N
Pre-survey	2,844	Honors chemistry	173	2,666
		Duplicates	1	
		Reported age <18	4	
Post-survey	2,454	Duplicates	360	1,878
		Honors chemistry	148	
		Consent = No/missing or reported age<18	68	
Course Outcomes				
Fall Data	3,910	Nonparticipating instructor[1]	1,214	2,696
Spring Data	2,113	Honors chemistry	109	2,004
Mastering Chemistry Use				
Fall	1,132	No issues	0	1,132
Spring	2,281	Courses not unrelated to analysis	938	1,343

1 – OSU provided outcome data for 12 instructors' classes; the study focused on six instructors who taught regular General Chemistry courses using Mastering Chemistry or Sapling.

Merging of data files

We first describe the overall logic and strategy for merging the various data files. We then discuss and present the results of file merging.

Overall logic and strategy for merging

Impact analysis. To assemble the data file necessary for the impact analysis, we had to merge the course outcomes data file with at least one of the student survey data files (pre- or post-survey). The student demographic information used in the impact analysis was collected in those surveys.

To complete the merge, the Survey Student ID (SSID), assigned at the administration of the student pre-survey, served as the linking variable to connect the course outcome file provided by institution with the student survey data files. (Students needed to complete at least one survey to be included in the impact analysis.)

Use-Outcome Analysis. To assemble the data file necessary for the use versus outcome analysis, we followed the same process for merging the course outcome and survey files described above. We used a Student Linking File provided by the institution (student first and last name, and SSID) and a Use Linking File provided by Pearson (Product User ID and student first and last name) to merge the Pearson Use File with the outcome and survey datafiles. We used the two linking files to replace the student identifier information in the use file (Product User ID and first and last name) with the SSID; doing so allowed us to merge the use file with the outcome and survey data files.

Notes on File Merging. Merging was challenging for three reasons that affected the final sample available for both the impact and use analysis. First, OSU did not provide the SSID with their course outcome file, instead providing a student OSU ID number. As result, the project had to rely on students' accurately entering their OSU ID on the post-survey to facilitate the merging of the course outcome and student survey files. Second, a shooting incident occurred on the OSU campus on November 28, 2016, one of the days that the post-survey was administered. As a result, classes for several sections did not meet that day, and there was no opportunity for a survey makeup day. Thus the response rate was lower than expected. Of the 2,839 students who responded to the pre-survey, only 1,878 responded to the post-survey, with most of the reduction resulting because their classes were cancelled on the day of the post-survey. Consequently, the analytic sample used for both the impact and use impact analysis was inherently restricted to the number of students who (1) responded to the post-survey, and (2) provided an accurate student ID.

Other student identifying information issues also resulted in complications for merging files for the use analysis. OSU requested a change in format for the survey—from paper (pre-survey) to an online survey (post-survey). On the online post-survey, students were required to recall their SSID from the pre-survey (as well as their OSU ID) and enter it into a field in the online form. Despite the administrators' best efforts to help students recall their SSID, approximately 14% of those responding to the post-survey did not recall their SSID accurately. Thus, the project team had to use the SSID, the OSU student ID entered on the post-survey, and the Student Linking File (first and last name and SSID) provided by the institution, to merge the course outcome file, student survey file, and use file. Consequently, the analytic sample for the use analysis was restricted to the number of students who (1) responded to the post-survey, (2) provided an accurate OSU student ID, and (3) recalled the SSID.

Table A3 shows the number of student records available in the various files and the number of student records SRI analysts were able to merge.

Table A3: number of student records in each file type and number of student records merged

Students			
	Total	**Treatment**	**Control**
Fall outcome data file	2,696	1,063	1,633
Use data files	--	1,132	--
Post-survey	1,878	808	1,069
Number of student records merged			
Outcome data file and post-survey	1,734	756	977
Use data file and post-survey	--	677	--
Outcome data file, use data file, and post-survey	--	634	--

Students			
	Total	**Treatment**	**Control**
Spring outcome data file	2,004	1,326	678
Use data files	--	1,343	--
Post-survey	1,878	808	1,069
Number of student records merged			
Outcome data file and post-survey	1,187	804	383
Outcome data file, use data file, and post-survey	--	795	--

Appendix B: Pre-Survey Question Frequency Tables for Mastering Chemistry Users at OSU (Treatment)

Q1. Have you taken an online course previously?

Reported in Percentages.

		N	Yes	No	No response
OSU	Treatment	1070	47.1	51.8	1.1

Q2. Have you ever used Mastering Chemistry before?

Reported in Percentages.

		N	Yes	No	No response
OSU	Treatment	1070	10.6	88.3	1.1

Q3. Have you ever used other Pearson learning products before?

Reported in Percentages.

		N	Yes	No	Not sure	No response
OSU	Treatment	1070	37	52.2	9.7	1

Pearson

Q4a. I ask myself questions to make sure I know the material I have been studying. *Reported in Percentages.*										
		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	0.7	2.2	7.1	17.4	30.9	25	15.3	1.3

Q4b. When work is hard I either give up or study only the easy parts. *Reported in Percentages.*										
		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	25.3	45.3	17.3	6.3	3.3	1.2	0.2	1.1

Q4c. I work on practice exercises and answer end of chapter questions even when I don’t have to. *Reported in Percentages.*										
		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	4.4	14.1	18.6	26.6	20.5	10.7	3.6	1.4

Q4d. I find that when the teacher is talking I think of other things and don't really listen to what is being said.

Reported in Percentages.

		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	16.4	37.9	20.4	12.6	8	2.9	0.4	1.3

Q4e. Even when study materials are dull and uninteresting, I keep working until I finish.

Reported in Percentages.

		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	0.4	2.1	4.7	11.7	26.2	32.9	20.7	1.4

Q4f. Before I begin studying I think about the things I will need to do to learn.

Reported in Percentages.

		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	1.2	5.4	9.4	16.7	26.3	23.9	15.7	1.3

Q4g. I often find that I have been reading for class but don't know what it is all about.

Reported in Percentages.

		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	6.2	22.1	23.1	19.5	19.1	6.5	2.2	1.2

Q4h. When I'm reading I stop once in a while and go over what I have read.

Reported in Percentages.

		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	1	5.3	10	17.8	31.4	23.4	9.9	1.2

Q4i. I work hard to get a good grade even when I don't like a class.

Reported in Percentages.

		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	0.2	0.2	1.5	4.7	10.5	26.9	54.9	1.2

Q5a. I think the field of chemistry is interesting.

Reported in Percentages.

		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	2.1	4.8	6.2	13.3	22	27.5	23.1	1.2

Q5b. I can apply what we are learning in chemistry class to real life.

Reported in Percentages.

		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	1.5	4.5	11	16.6	28.3	21.6	15.2	1.2

Q5c. I expect to do well in this class.

Reported in Percentages.

		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	0.4	1.3	4.7	15.2	23.6	29.7	23.6	1.5

Q5d. To be honest, I just don't find chemistry interesting. *Reported in Percentages.*										
		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	28.7	31.4	16.5	8.5	7	3.9	2.5	1.4

Q5e. I think what we will be studying in chemistry is useful to know. *Reported in Percentages.*										
		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	0.7	1.5	5.9	12.1	23.6	33.2	21.4	1.6

Q5f. Considering the difficulty of this course and my skills, I think I will do well in this class. *Reported in Percentages.*										
		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	0.5	2.1	6.1	18.3	27.9	28.9	15	1.3

Q5g. I think what we're learning in this class will be interesting.

Reported in Percentages.

		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	0.9	2.2	5.7	12.6	25.4	29.3	22.5	1.2

Q5h. I can see how what I learn from chemistry applies to life.

Reported in Percentages.

		N	Not at all true	Mostly untrue	Somewhat untrue	Neither true or untrue	Somewhat true	Mostly true	Very true	No response
OSU	T	1070	1.1	3.7	7.5	17.5	24.4	26.1	18.4	1.3

Q6a. You have a certain amount of intelligence and you really can't do much to change it.

Reported in Percentages.

		N	Strongly agree	Agree	Mostly agree	Mostly disagree	Disagree	Strongly disagree	No response
OSU	T	1070	0.7	5.4	16.4	22.2	34.2	19.6	1.3

Q6b. You can learn new things, but you can't really change your basic intelligence.

Reported in Percentages.

		N	Strongly agree	Agree	Mostly agree	Mostly disagree	Disagree	Strongly disagree	No response
OSU	T	1070	2.6	10.4	22.8	22.7	28.1	12	1.4

Q6c. Your intelligence is something about you that you can't change very much.

Reported in Percentages.

		N	Strongly agree	Agree	Mostly agree	Mostly disagree	Disagree	Strongly disagree	No response
OSU	T	1070	1.3	6.4	13.6	25	34	18.4	1.4

Appendix C: Post-Survey Question Frequency Tables for Mastering Chemistry Users (Treatment)

Q1. How often on average did you use Mastering Chemistry to help you learn chemistry?

Reported in Percentages.

		N	Daily or almost every day	2-3 times each week	One time per week	One time every 2-3 weeks or less	Never	[missing]
OSU	T	808	54.5	39.9	3.5	1.5	0.6	0.1

Q3. When do you primarily use Mastering Chemistry for this course?

Reported in Percentages.

		N	In class, during regular class time	In computer lab, as part of scheduled lab time	Outside of regularly scheduled class or lab time	no response
OSU	T	808	9.8	0.1	90	0.1

Q4a. How easy is it to get your account set up?

Reported in Percentages.

		N	Easy	Not easy	no response
OSU	T	808	93.8	5.9	0.2

Q4b. How easy is it to login?

Reported in Percentages.

		N	Easy	Not easy	no response
OSU	T	808	96.7	2.8	0.5

Q4c. How easy is it to find your grade on a homework assignment?

Reported in Percentages.

		N	Easy	Not easy	no response
OSU	T	808	90.8	8.4	0.7

Q4d. How easy is it to find the relevant information in the digital textbook? / How easy is it to find information you are looking for?

Reported in Percentages.

		N	Easy	Not easy	no response
OSU	T	808	62	37.5	0.5

Q4e. How easy is it to get help when you don't understand something?

Reported in Percentages.

		N	Easy	Not easy	no response
OSU	T	808	65.1	34.4	0.5

Q4f. How easy is it to find out how much work you have finished?

Reported in Percentages.

		N	Easy	Not easy	no response
OSU	T	808	96.3	3.1	0.6

Q5a. When using Mastering Chemistry, I believe I came to class better prepared.

Reported in Percentages.

		N	Very true	Mostly true	Sometimes true	Not at all true	Not sure / uncertain	no response
OSU	T	808	31.8	44.1	20.9	2.2	0.6	0.4

Q5b. When using Mastering Chemistry, I believe I enjoyed this class more.

Reported in Percentages.

		N	Very true	Mostly true	Sometimes true	Not at all true	Not sure / uncertain	no response
OSU	T	808	15.6	28.2	35.1	17.8	2.7	0.5

Q5c. When using Mastering Chemistry, I believe I was more engaged in the learning experience.

Reported in Percentages.

		N	Very true	Mostly true	Sometimes true	Not at all true	Not sure / uncertain	no response
OSU	T	808	27.7	40.1	24.1	6.8	0.7	0.5

Q5d. When using Mastering Chemistry, I believe I received more personalized feedback on my work.

Reported in Percentages.

		N	Very true	Mostly true	Sometimes true	Not at all true	Not sure / uncertain	no response
OSU	T	808	20.8	25.6	30	20.9	2.2	0.5

Q5e. When using Mastering Chemistry, I believe I was able to use different approaches to help me learn.

Reported in Percentages.

		N	Very true	Mostly true	Sometimes true	Not at all true	Not sure / uncertain	no response
OSU	T	808	24.4	32.5	33	8.8	0.9	0.4

Q5f. When using Mastering Chemistry, I believe I had access to a greater variety of learning materials.

Reported in Percentages.

		N	Very true	Mostly true	Sometimes true	Not at all true	Not sure / uncertain	no response
OSU	T	808	32.5	36	24.9	4.7	1.5	0.4

Q5g. When using Mastering Chemistry, I believe it helped me learn new problem-solving skills.

Reported in Percentages.

		N	Very true	Mostly true	Sometimes true	Not at all true	Not sure / uncertain	no response
OSU	T	808	26.4	32.4	29.5	10.1	1	0.6

Q5h. When using Mastering Chemistry, I believe it increased my confidence that I can learn new things on my own without an instructor.

Reported in Percentages.

		N	Very true	Mostly true	Sometimes true	Not at all true	Not sure / uncertain	no response
OSU	T	808	29.7	32.1	25.7	10.6	1.1	0.7

Pearson

Q5i. When using Mastering Chemistry, I believe it helped me prepare for tests and quizzes.

Reported in Percentages.

		N	Very true	Mostly true	Sometimes true	Not at all true	Not sure / uncertain	no response
OSU	T	808	37.4	35.8	19.8	6.1	0.5	0.5

Q6a. I think the field of chemistry is interesting.

Reported in Percentages.

		N	Not at all true of me	Mostly untrue	Somewhat untrue	Neutral	Somewhat true	Mostly true	Very true of me	no response
OSU	T	808	5.2	5.9	8.2	10.6	23.8	26	19.8	0.5

Q6b. I can apply what we are learning in chemistry class to real life.

Reported in Percentages.

		N	Not at all true of me	Mostly untrue	Somewhat untrue	Neutral	Somewhat true	Mostly true	Very true of me	no response
OSU	T	808	4.5	7.1	9.4	17.5	27.2	21.2	12.7	0.5

Q6c. To be honest, I just don't find chemistry interesting.

Reported in Percentages.

		N	Not at all true of me	Mostly untrue	Somewhat untrue	Neutral	Somewhat true	Mostly true	Very true of me	no response
OSU	T	808	24.5	30.6	15.5	7.9	8.7	5.1	7.3	0.5

Q6d. I think what we are studying in chemistry is useful to know.

Reported in Percentages.

		N	Not at all true of me	Mostly untrue	Somewhat untrue	Neutral	Somewhat true	Mostly true	Very true of me	no response
OSU	T	808	3.8	6.6	10.9	15.5	24.1	22.8	15.6	0.7

Q6e. I think what we're learning in this class is interesting.

Reported in Percentages.

		N	Not at all true of me	Mostly untrue	Somewhat untrue	Neutral	Somewhat true	Mostly true	Very true of me	no response
OSU	T	808	3.3	5.2	9.3	14.6	25.2	24.5	17.3	0.5

Q6f. I can see how what I learn from chemistry applies to life.

Reported in Percentages.

		N	Not at all true of me	Mostly untrue	Somewhat untrue	Neutral	Somewhat true	Mostly true	Very true of me	no response
OSU	T	808	2.7	4.6	10	15.8	27.8	22.9	15.6	0.5

Q7a. You have a certain amount of intelligence and you really can't do much to change it.

Reported in Percentages.

		N	Strongly Agree	Agree	Mostly agree	Mostly disagree	Disagree	Strongly Disagree	no response
OSU	T	808	2.5	6.9	19.7	24.9	30.6	15	0.5

Q7b. You can learn new things, but you can't really change your basic intelligence.

Reported in Percentages.

		N	Strongly Agree	Agree	Mostly agree	Mostly disagree	Disagree	Strongly Disagree	no response
OSU	T	808	3.5	16.7	22.8	24.4	22.8	9.4	0.5

Q7c. Your intelligence is something about you that you can't change very much.

Reported in Percentages.

		N	Strongly Agree	Agree	Mostly agree	Mostly disagree	Disagree	Strongly Disagree	no response
OSU	T	808	2.4	7.5	18.9	26.7	29.6	14.4	0.5

Q8a. Insufficient group tutorial sessions/recitations impacted my use of Mastering Chemistry. / Insufficient tutorial sessions impacted my use of Mastering Chemistry.

Reported in Percentages.

		N	Significant impact	Moderate Impact	Small Impact	No impact	no response
OSU	T	808	6.7	24.4	29	39	1

Q8b. Insufficient individual tutoring sessions with TAs or instructors impacted my use of Mastering Chemistry. / Insufficient extra tutoring availability impacted my use of Mastering Chemistry.

Reported in Percentages.

		N	Significant impact	Moderate Impact	Small Impact	No impact	no response
OSU	T	808	8.4	23.4	26.5	40.7	1

Q8c. Insufficient peer study groups impacted my use of Mastering Chemistry.

Reported in Percentages.

		N	Significant impact	Moderate Impact	Small Impact	No impact	no response
OSU	T	808	5.3	20.2	27.2	46.3	1

Q8d. Other reasons impacted my use of Mastering Chemistry.

Reported in Percentages.

		N	Significant impact	Moderate Impact	Small Impact	No impact	no response
OSU	T	808	3	4.6	4.5	82.9	5.1

Pearson

Q9a. Lack of access to a computer impacted my use of Mastering Chemistry. / Insufficient computers impacted my use of Mastering Chemistry.

Reported in Percentages.

		N	Significant impact	Moderate Impact	Small Impact	No impact	no response
OSU	T	808	9.2	4.8	7.8	76.6	1.6

Q9b. Too slow computer impacted my use of Mastering Chemistry.

Reported in Percentages.

		N	Significant impact	Moderate Impact	Small Impact	No impact	no response
OSU	T	808	5.4	9.7	17.9	65.2	1.7

Q9c. Insufficient bandwidth/too slow internet impacted my use of Mastering Chemistry.

Reported in Percentages.

		N	Significant impact	Moderate Impact	Small Impact	No impact	no response
OSU	T	808	6.2	10.8	21	60.6	1.4

Q9d. Slow load time for homework questions impacted my use of Mastering Chemistry.

Reported in Percentages.

		N	Significant impact	Moderate Impact	Small Impact	No impact	no response
OSU	T	808	5.8	9.9	23.1	59.5	1.6

Q9e. Getting a question marked incorrect because incorrectly formatted impacted my use of Mastering Chemistry.

Reported in Percentages.

		N	Significant impact	Moderate Impact	Small Impact	No impact	no response
OSU	T	808	23.9	30.7	26.5	17.7	1.2

Q9f. Lack of technical or help desk support from my campus impacted my use of Mastering Chemistry.

Reported in Percentages.

		N	Significant impact	Moderate Impact	Small Impact	No impact	no response
OSU	T	808	3.8	6.2	15	73.4	1.6

Pearson

Q9g. Lack of technical or help desk support from Pearson impacted my use of Mastering Chemistry.

Reported in Percentages.

		N	Significant impact	Moderate Impact	Small Impact	No impact	no response
OSU	T	808	4.7	6.2	12.1	75.9	1.1

Q9h. Other reasons impacted my use of Mastering Chemistry.

Reported in Percentages.

		N	Significant impact	Moderate Impact	Small Impact	No impact	no response
OSU	T	808	0.5	0.9	0.4	88.2	10

Q10a. How often did you use Tutorial Problems?

Reported in Percentages.

		N	Often	Sometimes	Almost Never	Did Not Use / Not Available	no response
OSU	T	808	19.6	32.4	26.7	20.2	1.1

Pearson

Q10b. How often did you use Hints?

Reported in Percentages.

		N	Often	Sometimes	Almost Never	Did Not Use / Not Available	no response
OSU	T	808	41.1	34.9	17.2	5.8	1

Q10c. How often did you use Videos?

Reported in Percentages.

		N	Often	Sometimes	Almost Never	Did Not Use / Not Available	no response
OSU	T	808	28.6	41.1	22	7.3	1

Q10d. How often did you use Dynamic Study Modules?

Reported in Percentages.

		N	Often	Sometimes	Almost Never	Did Not Use / Not Available	no response
OSU	T	808	17.8	28.1	31.3	21.7	1.1

Q10e. How often did you use Wrong Answer Feedback?

Reported in Percentages.

		N	Often	Sometimes	Almost Never	Did Not Use / Not Available	no response
OSU	T	808	63.2	25.6	7.8	2.4	1

Q10f. How often did you use Adaptive follow-up Homework?

Reported in Percentages.

		N	Often	Sometimes	Almost Never	Did Not Use / Not Available	no response
OSU	T	808	22.3	27	25.6	24.1	1

Q10g. How often did you use Learning Catalytics?

Reported in Percentages.

		N	Often	Sometimes	Almost Never	Did Not Use / Not Available	no response
OSU	T	808	52	15.1	13.4	18.4	1.1

Pearson

Q11a. How helpful was the Tutorial Problems aspect?

Reported in Percentages.

		N	Very Helpful	Somewhat Helpful	Not Helpful	Did Not Use	no response
OSU	T	808	38	36	4	21	1

Q11b. How helpful was the Hints aspect?

Reported in Percentages.

		N	Very Helpful	Somewhat Helpful	Not Helpful	Did Not Use	no response
OSU	T	808	66.6	23.9	2.8	5.6	1.1

Q11c. How helpful was the Videos aspect?

Reported in Percentages.

		N	Very Helpful	Somewhat Helpful	Not Helpful	Did Not Use	no response
OSU	T	808	36.8	41.7	10.3	10.1	1.1

Q11d. How helpful was the Dynamic Study Modules aspect?

Reported in Percentages.

		N	Very Helpful	Somewhat Helpful	Not Helpful	Did Not Use	no response
OSU	T	808	24.1	32.7	9.7	32.1	1.5

Q11e. How helpful was the Wrong Answer Feedback aspect?

Reported in Percentages.

		N	Very Helpful	Somewhat Helpful	Not Helpful	Did Not Use	no response
OSU	T	808	67.7	24.3	3.2	3.6	1.2

Q11f. How helpful was the Adaptive Follow-up Homework aspect?

Reported in Percentages.

		N	Very Helpful	Somewhat Helpful	Not Helpful	Did Not Use	no response
OSU	T	808	28.7	26.7	8.7	34.9	1

Q11g. How helpful was the Learning Catalytics aspect?

Reported in Percentages.

		N	Very Helpful	Somewhat Helpful	Not Helpful	Did Not Use	no response
OSU	T	808	41.8	25.6	7.9	23.4	1.2

Q13. Would you prefer your instructor made more or less use of Mastering Chemistry in this class?

Reported in Percentages.

		N	More Use	Less Use	Don't change; it's about right	no response
OSU	T	808	14.9	20.5	63.6	1

Q15. Do you intend to continue taking chemistry courses in the future?

Reported in Percentages.

		N	Yes	No	Not Sure	no response
OSU	T	808	79.5	10.9	8.8	0.9

Appendix D. Use Descriptives

Descriptive statistics on use

In analyzing the fall and spring use data, SRI focused on data related to assignments and problem attempts. Pearson distinguishes activity by assignment type—homework, extra practice, test, and adaptive. However, instructors can also create their own assignment types. Therefore to ensure consistency across institutions, SRI regrouped assignment types with guidance from Pearson staff. Table E1 shows how assignment types were grouped for this analysis. Tables E2 and E3 provide descriptive statistics for Mastering Chemistry assignment types and use variables for the fall and spring semesters, respectively. Figures E1 through E4 provide Mastering Chemistry use statistics for the fall and spring semesters.

Table D1: Reassignment of assignment types in Mastering Chemistry use data

Pearson Group Suggestion	Assignment Types	SRI Label
Group 1	Homework Pre-class Assignments Pre-lecture Pre-lecture Assignments Assignments	Homework
Group 2	Extra Practice Assessment of Learning End of Chapter Exam Prep Exam Prep Follow-Up Practice Additional Practice	Extra Practice
Group 3	Test	Test
Group 4	The Grind Final Exam Grind	Adaptive

Pearson

Table D2: Descriptive statistics for Mastering Chemistry use variables based on student-level system data for the fall semester

Variable	*n*	Mean	SD	Min	Median	Max
Overall						
Unique Days	677	50.81	13.67	1	51	84
Percent of Class Days Used	677	52%	14%	1%	52%	86%
Time Spent on Mastering Chemistry (hours)	677	31.16	11.38	0.32	30.12	65.57
Unique Assignments	677	71.84	26.26	1	70	121
Unique Problems	677	533.71	155.83	4	556	919
Homework						
Time Spent on Mastering Chemistry (hours)	677	12.63	8.24	0	13.63	40.97
Unique Assignments	677	23.34	12.46	0	32	36
Unique Problems	677	96.82	41.84	0	116	155
Extra Practice						
Time Spent on Mastering Chemistry (hours)	677	16.44	8.61	0	15.30	51.43
Unique Assignments	677	360.31	130.26	0	393	681
Unique Problems	677	34.30	10.05	0	37	58
Test						
Time Spent on Mastering Chemistry (hours)	677	0.43	0.85	0	0	3.62
Unique Assignments	677	26.16	50.37	0	0	149
Unique Problems	677	0.59	1.09	0	0	3
Adaptive						

Time Spent on Mastering Chemistry (hours)	677	1.66	1.84	0	1.12	7.25
Unique Assignments	677	50.44	50.34	0	34	114
Unique Problems	677	13.62	13.29	0	11	30
Wrong Attempts per Problem	677	1.28	0.71	0.11	1.15	4.27
Hints per Problem	677	0.25	0.26	0.00	0.19	3
Points per Problem	677	0.89	0.08	0.58	0.89	1
Raw Score per Problem	677	0.83	0.08	0.51	0.84	1
Score per Problem	677	0.80	0.10	0.30	0.81	1
% of Assigned Credits Finished	677	97%	6%	27%	99%	1

Table D3: Descriptive statistics for Mastering Chemistry use variables based on student-level system data for the spring semester

Variable	*n*	Mean	SD	Min	Median	Max
Overall						
Unique Days	1311	41.86	12.41	1	42	84
Percent of Class Days Used	1311	0.43	0.13	0.01	0.43	0.86
Time Spent on Mastering Chemistry (hours)	1311	28.90	11.87	0.06	28.40	79.82
Unique Assignments	1311	60.22	29.37	1	47	158
Unique Problems	1311	452.07	151.04	1	513	939
Homework						
Time Spent on Mastering Chemistry (hours)	1311	19.4	10.4	0	16.62	54.58
Unique Assignments	1311	29.3	10.4	0	23	44
Unique Problems	1311	256.9	169.4	0	152	471
Extra Practice						

Time Spent on Mastering Chemistry (hours)	1311	6.86	9.20	0	0.45	46.95
Unique Assignments	1311	131.81	156.16	0	17	528
Unique Problems	1311	25.26	31.16	0	4	107
Test						
Time Spent on Mastering Chemistry (hours)	1311	0.75	0.88	0	0	3.95
Unique Assignments	1311	28.04	29.46	0	0	95
Unique Problems	1311	0.94	0.98	0	0	2
Adaptive						
Time Spent on Mastering Chemistry (hours)	1311	0.78	1.45	0	0	8.59
Unique Assignments	1311	20.19	40.22	0	0	242
Unique Problems	1311	3.28	5.89	0	0	30
Wrong Attempts per Problem	1311	1.20	0.70	0.07	1.05	6.00
Hints per Problem	1311	0.26	0.27	0.00	0.19	3.00
Points per Problem	1311	0.96	0.04	0.00	0.96	1.09
Raw Score per Problem	1311	0.83	0.09	0.50	0.84	1.00
Score per Problem	1311	0.78	0.12	0.45	0.79	1.00
% of Assigned Credits Finished	1310	0.89	0.15	0.09	0.97	1.00

Figure D1: Distribution of time spent on Mastering Chemistry by student for the fall semester

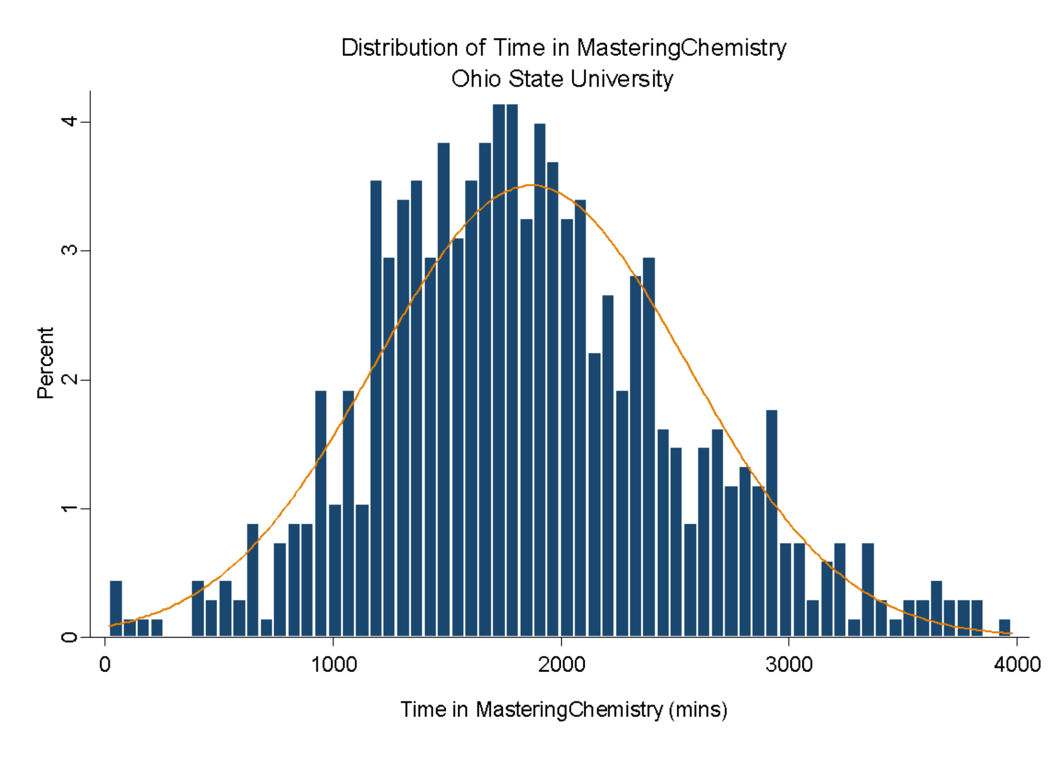


Figure D2: Distribution of number of problems attempted in Mastering Chemistry for the fall semester

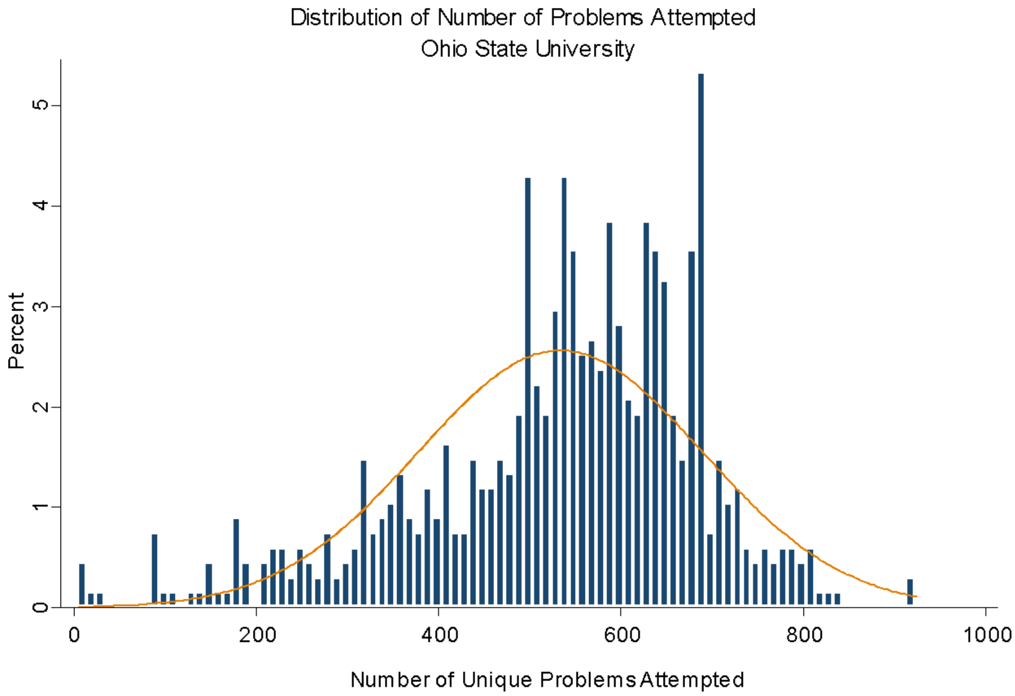


Figure D3: Distribution of time spent on Mastering Chemistry by student for the spring semester

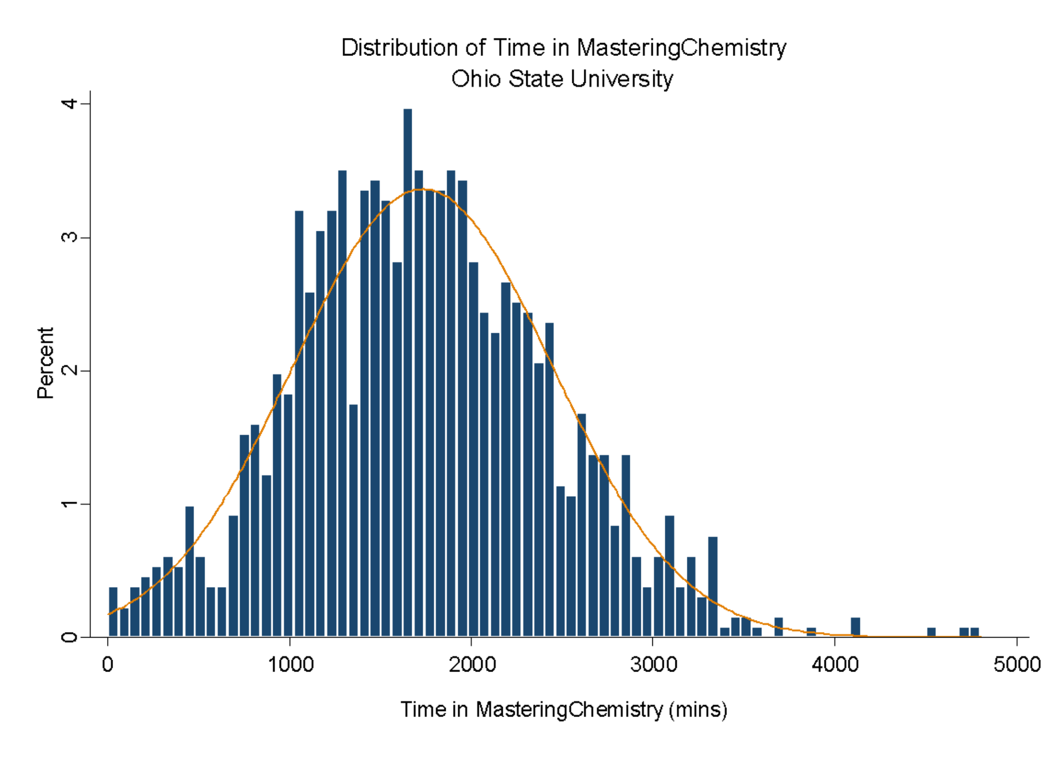
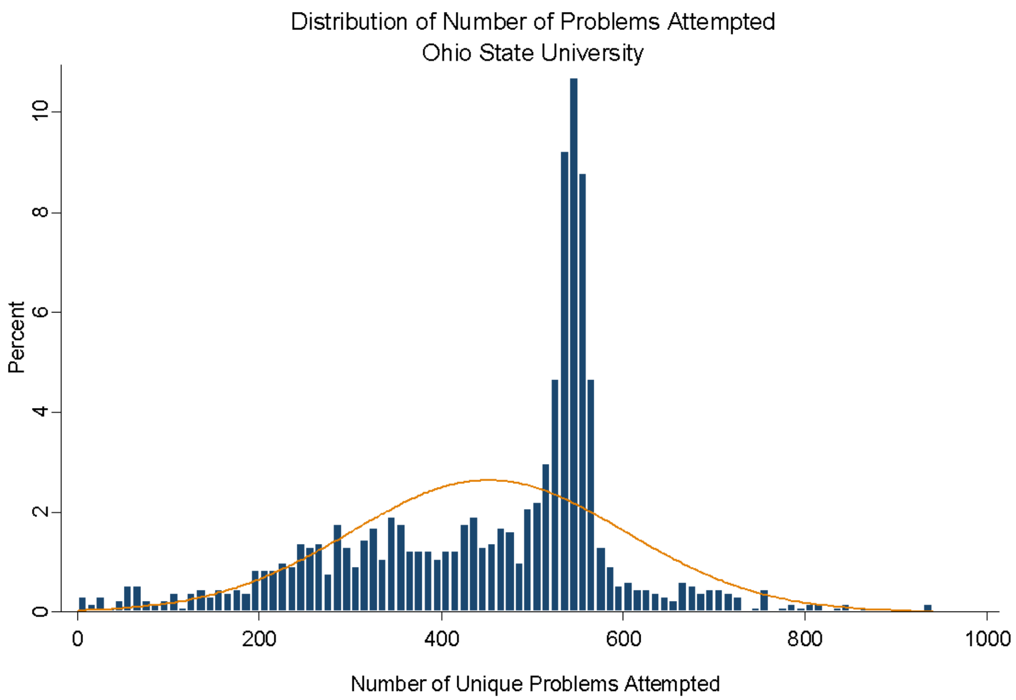


Figure D4: Distribution of number of problems attempted in Mastering Chemistry for the spring semester



Appendix E. Teacher Characteristics Descriptives

Table E1: Fall General Chemistry I instructor characteristics

	Teacher A	Teacher B	Teacher C	Teacher D	Teacher E	Teacher F
Condition	Treatment	Treatment	Control	Control	Control	Control
Product	Mastering Chemistry	Mastering Chemistry	Sapling	Sapling	Sapling	Sapling
Number of Sections Taught	17	27	14	19	14	14
Years' experience with product	5-6	9-10	3-4	1-2	3-4	1-2
Years teaching college courses	15	10	9	20	13	26
Years teaching at OSU	14	10	9	20	13	26
Position at OSU	Associate professor, non-tenured	Full-time lecturer	Lab assistant	Full-time lecturer	Part-time/ adjunct	Full-time lecturer

Table E2: Spring General Chemistry II instructor characteristics

	Teacher G / H (sections co-taught by two instructors)	Teacher I	Teacher K
Condition	Treatment	Treatment	Control
Product	Mastering Chemistry	Mastering Chemistry	Sapling
Number of Sections Taught	27	27	28
Years' experience with product	NA / NA	9-10	1-2
Years teaching college courses	NA / NA	10	20
Years teaching at OSU	7 / 5	10	20
Position at OSU	Associate professor / Professor	Full-time lecturer	Full-time lecturer