Efficacy Report

MasteringChemistry

March 23, 2016

Product Summary

MasteringChemistry is an online tutorial and classroom engagement program designed to improve results by helping learners gain proficiency more quickly, master concepts, and develop problem-solving skills. Learners benefit from self-paced tutorials featuring specific, wrong-answer feedback and hints that emulate a human tutor. Additional key features include *Learning Catalytics* (a classroom engagement tool), *dynamic study modules* (an adaptive, conceptual mastery tool), and an eText.

Intended Outcomes

Overview of Intended Outcomes

MasteringChemistry is designed to engage learners and provide a positive learning experience so that they can improve their results in the course, master concepts and develop problem-solving skills. It does this by providing learners access to a range of learning and assessment features, such as dynamic study modules that personalize the learning experience. These features help learners study effectively on their own by continuously assessing performance in real time and providing feedback.

Intended Outcome 1: Learners have access to learning resources
MasteringChemistry is designed to be accessible anytime via phones and tablets, as a means of increasing engagement in the classroom and beyond.

Intended Outcome 2: Learners are engaged and have a positive learning experience
MasteringChemistry has a number of features designed to fully engage learners, from interactive pre-lecture videos that provide an introduction on key topics, to embedded assessment to help students prepare before a lecture. An additional key feature is a classroom engagement tool, *Learning Catalytics*, which augments lectures by providing guiding questions for in-class discussion that encourage student participation and foster a positive learning experience.

Intended Outcome 3: Learners complete tasks/activities/course on time
MasteringChemistry learners benefit from self-paced tutorials featuring specific, wrong-answer feedback and hints that emulate a human tutor with the aim of making them more likely to complete activities and, ultimately, the course on time.

Intended Outcome 4: Learners improve in skills, competency and expertise in chemistry
MasteringChemistry can be used to develop understanding of the material through pre-reading, quizzes and dynamic study modules assigned prior to class. After class, it can contribute to the mastery of concepts through homework, quizzes and adaptive learning follow-ups.

Foundational Research

Overview of Foundational Research

MasteringChemistry is designed to support students as they learn content, such that their knowledge can flexibly transfer to new topics later in the course and into the future. The tutorial problems align closely with educational research on designing effective 'worked examples' which have been shown to lead to effective and efficient learning for novice students, as they help focus learners' attention and manage cognitive load (Sweller & Cooper, 1985; Atkinson, Derry, Renkl, & Wortham, 2000). In addition, chemistry education research has highlighted the importance of the kind of high-quality, self-directed exercises that MasteringChemistry provides in promoting achievement outcomes, such as higher final course grades (Cuadros, Yaron, & Leinhardt, 2007).

MasteringChemistry leverages the best in instructional research and design. The areas of particular strength are:

Scaffolding and Fading
In MasteringChemistry there are a variety of learner support tools to help students struggling with assessment items (e.g. hints, reference tables, e-text). Some questions provide multiple levels of hints to provide more gradual levels of support. The tutorial problems help break down complex content into manageable ideas, an approach that is aligned to research on how 'worked examples' can support learning. Research has found that novices learn and process information in fundamentally different ways to those with more background knowledge (Chi, Feltovich, & Glaser, 1981). Specifically, novices require more support, as they do not have a body of relevant knowledge and strategies to draw upon to help them solve new problems or learn new information. As such, it is critically important to *scaffold*, or support, novice learners in a variety of ways and one of the most well-researched approaches is through the use of 'worked examples' (Sweller & Cooper, 1985). At its core, a 'worked example' shows students how to complete a problem-solving procedure, usually followed by opportunities for the student to practice this themselves.

Memory Strategies
MasteringChemistry assessment items are designed and offered in a way that allow for increased opportunities for practice, which research shows leads to a deeper understanding of procedures and rules. A very robust finding from research about memory strategies is that increasing the number of opportunities to recall information over time can help students to learn (Karpicke & Roediger III, 2010), compared to isolated presentations of information.There is a growing body of evidence that this kind of 'retrieval practice' contributes to both long-term retention and to learning in a way that can flexibly transfer to help solve new problems (Roediger III & Butler, 2011; Carpenter, 2012).

Adaptivity
Research has identified two types of adaptivity in learning technologies. One type relates to adaptive responses to students (i.e., adaptive feedback). Adaptive systems which provide timely feedback to students as they engage with the learning technology have been seen to be as effective as human tutors (Van Lehn, 2011). The other mode of adaptivity relates to adapting a

learning sequence based on an understanding of a student's current proficiency. One way in which this can be done is by estimating mastery based on student performance, and ensuring that students receive enough practice to achieve fluency with the content. This "Knowledge Tracing' has been used to great effect (Corbett & Anderson, 1995).

In MasteringChemistry items are selected based on performance on previous items and adaptive study guides are provided to help target areas of learning need. Dynamic study modules provide adaptive experiences to support improvement in competency in chemistry concepts and skills, providing extra practice problems when students are unsure or incorrect. Adaptive follow-up assignments provide coaching and targeted opportunities for remediation, personalized to each student's strengths and weaknesses.

Intended Product Implementation

Overview of Intended Product Implementation

MasteringChemistry is designed to be flexible to support different implementation models based on the goal the faculty member is trying to achieve. MasteringChemistry can be used to gauge understanding of the material through pre-reading, quizzes and dynamic study modules assigned prior to class and after class. In addition, mastery of concepts is supported through homework, quizzes and adaptive follow-ups. Currently the most popular implementation approach is as an auto-graded online homework tool.

Product Research

Overview of Product Research

Over the past 10 years there have been 21 educator studies completed by Pearson MasteringChemistry users. These reports describe how instructors have implemented MasteringChemistry and provide users with examples of how MasteringChemistry can be used within their own courses. In many of these studies instructors also provide data describing students' experiences and course engagement while using MasteringChemistry, and report on levels of academic achievement and passing rates attained in the course following the implementation of MasteringChemistry. However, because these studies were not designed using experimental or quasi-experimental designs that could robustly attribute the changes in learner outcomes to MasteringChemistry, for this report Pearson researchers completed a systematic search and review of published, peer-reviewed studies that assessed the impact of MasteringChemistry on learner outcomes. Four studies related to MasteringChemistry have been published in peer-reviewed academic journals and one as a dissertation. From among these, only one study provides robust information about the relationship of the use of MasteringChemistry on students' scores on the American Chemical Society (ACS) test in general chemistry.

About the claims included in this report

In reporting causal claims about the impact of our products on learner outcomes, Pearson relies on guidance by the US Department of Education What Works Clearinghouse (WWC) to determine which available studies have been conducted and reported using the following standards of research that would yield valid and reliable information.

- **Research design.** In accordance with WWC standards, we focused on randomized controlled trials, quasi-experimental designs, regression discontinuity designs and single case designs. If conducted properly, these are the research designs that are of sufficient rigour to allow us to make causal claims.
- **Attrition.** Using WWC guidance, we considered both overall and differential attrition - the extent to which both users and non-users of the Pearson product remain in the study - in relation to each other - to determine if we could use the results from the studies.
- **Baseline equivalence.** Equivalence of the intervention and comparison groups on observable characteristics at baseline must be presented for the analytic sample in order for us to draw valid conclusions about the extent to which changes in student outcomes could be reliably and validly attributed to use of the Pearson product.
- **Outcomes eligibility.** Outcomes examined in the studies must have face validity and be of sufficient reliability for us to use the results from the outcomes. These outcomes must also be aligned to Pearson's defined learner outcomes. There must also be no over-alignment of the outcomes to the Pearson product and data collection for the outcomes must be done in the same manner for both the treatment and control groups.

Our Search and Review Process for Evidence of Impact Provided by Independent and Published Studies

The search for external studies on a particular Pearson courseware product was done primarily through Google Scholar and the rest were found through snowballing from relevant articles on the Pearson product. In Google Scholar, the keyword of the Pearson product was used before a scan of the title and abstract was done to determine relevance. No studies earlier than 2008 were reviewed since they were probably based on an outdated version of the Pearson product. Articles that discussed implementation but did not examine any Pearson defined learner outcomes on a population in higher education were also excluded.

During our review process, it was found that some of the quasi-experimental studies did not examine baseline differences at all. Others were of slightly higher rigour in that prior differences were examined or acknowledged but no attempts were made to take into account these prior differences in the analysis. Thus, we would consider these studies correlational studies and not studies that can lead to causal claims.

The next level of better designed studies were those that tried to control for prior differences of students in some way, such as through regression analysis or ANCOVA. Other studies used random assignment. However, many of these quasi-experimental or randomized studies assigned the treatment or control condition at the class level but failed to consider clustering at the class/instructor level when there were multiple instructors involved in the study. Thus, the potential confounding effects of instructors on student achievement were not taken into account in these studies. As a result, in most cases, the estimates of the statistical reliability of the impact estimates from these studies are incorrect and misleading, leading readers to believe they should have more confidence in the results than warranted. Thus, in summary, the studies that could be used to make causal claims were those that used a rigorous research design, had acceptable overall and differential attrition rates, addressed baseline equivalence, and used the right method of analysis.

Research Studies

Study 1: Impact of an online homework program on the achievement of college students enrolled in the first semester general chemistry.	
Study Citation	Huesgen, B. D. (2012). *Impact of an online homework program on the achievement of college students enrolled in the first semester general chemistry* (Unpublished doctoral dissertation). University of Missouri-St. Louis, St. Louis, MO.
Research Study Contributors	Author and the institution indicated above
Type of Study	Simple regression analysis
Sample Size	Sample of 95 students, with 58 classified as low-performing and

	37 classified as high-performing
Description of Sample	The students varied in age from traditional freshmen to older, non-traditional students. The course under study was designed for science majors and engineers though it was not exclusive to these majors. Before enrolling in general chemistry I, the students were required to have completed both algebra and trigonometry but concurrent enrollment in trigonometry was allowed. The course was the first part on a two-course series in general chemistry and students were required to take both.
Outcomes Measured	Achievement : Learners improve in competency or expertise in chemistry.

Introduction

The study, conducted as part of a doctoral dissertation, analysed students taking the general chemistry I course at the University of Missouri-Saint Louis in fall 2007. In the semester the instructor assigned online homework using MasteringChemistry. This study examined the relationship between performance in MasteringChemistry, used as an online homework platform, and students' understanding of material in university level general chemistry, as measured by their performance on the standardized American Chemical Society (ACS) First Semester General Chemistry final examination, a widely accepted benchmark in North America for assessing understanding of general chemistry. Before analyzing the relationship, students were split into two groups—high-performing and low-performing - based on whether they scored above or below the mean on the ACS final exam. The purpose of this study was to investigate whether use of MasteringChemistry was associated with different levels of learning gains among lower-performing and higher-performing students.

Method

Students enrolled in the fall semester of 2007 general chemistry I used MasteringChemistry for online homework. Students who scored above and below the mean on the ACS final exam were split into high-performing and low-performing groups of students. The total of 95 students were grouped into two with 37 classified as high-performing and 58 classified as low-performing students. Then within each of the two groups, the relationship of the MasteringChemistry score with the ACS final exam score was determined through a simple regression analysis where MasteringChemistry score(s) was used as the only predictor.

Results

For students who scored below the mean on the ACS final exam (low-performing students whose scores were 32 or below), there was no relationship between performance on MasteringChemistry and scores on the ACS final, $F(1, 56) = 0.136$, $p > .05$, $R^2 = .002$. For students who scored above the mean on the ACS final exam (high-performing students), there was a significant positive relationship between performance on MasteringChemistry and scores on the ACS final, $F(1, 35) = 6.274$, $p < .05$, $R^2 = .152$.

The author also compared the students from 2007 fall semester who used MasteringChemistry with a previous cohort of students in fall 2005 who did not use MasteringChemistry and instead, as indicated by the author, used a less sophisticated online program. Demographics in terms of ethnicity, age and gender distributions were noted by the author to be similar between the two groups. It was found that the two semesters of students did not perform (significantly) differently from each other in terms of the ACS final. However, though the author noted that the two groups were similar in terms of demographics, the author still did not address baseline equivalence of the two cohorts, such as prior achievement and socio-economic status, which are important covariates in higher education. Hence, we cannot make a claim based on this comparison.

Discussion

This study provided evidence of a modest association between two independent variables, MasteringChemistry scores and ACS final exam scores. However, sample sizes were relatively small, and the association between performance on MasteringChemistry and performance on a standardized final exam did not rule out confounding factors that might explain the association other than MasteringChemistry. For example, one of the important possible confounders is that high ability students would tend to perform well on both MasteringChemistry and the final exam. The author tried to address this in an inadequate way by splitting students into a high-performing group and a low-performing group based on their final exam score before analysis. However, splitting students into high-performing and low-performing groups resulted in smaller sample sizes for analysis and also inadvertently restricted the range of the ACS scores in calculating the correlation, which might then attenuate the correlations.

Lastly, this study was conducted in 2007 and since then, MasteringChemistry might have changed which could make the results from this study less relevant today.

Future Research Plans

Overview of Future Research Plans

There are four planned studies, including two quasi-experimental studies and two efficacy trials. One of the quasi-experimental studies will be conducted by SRI International, a non-profit independent research center, and the other case control quasi-experimental study currently in process under the direction of researchers in Pearson's Global Product Efficacy and Analytics Team, in collaboration with Pearson's North America Customer Experience and Engagement Team and university faculty at the study sites. The two efficacy trials were designed to test new learning functionalities that Pearson has recently developed: Adaptive Module and Early Alert.

Future Research Plans

Study 1: Quasi-experimental Propensity-Score Matching Study on the Efficacy of MasteringChemistry	
Intended Start Date	January, 2016
Anticipated Length of Study	Until end of 2017
Type of Study	Quasi-experimental
Research Leads	SRI International
Intended Sample Size	4 institutions with a total sample of over 6,000 users and 6,000 matched non-users
Description of Sample	MasteringChemistry users and non-users from the 4 institutions
Outcomes to be Measured	Experience, Completion, Academic Achievement, and Progression

Proposed Study Description
Pearson has commissioned SRI International, an independent research firm recognized for its expertise in evaluations of education technology products, to conduct a quasi-experimental study that uses propensity score matching techniques to compare users and non-users of MasteringChemistry. SRI will implement a rigorous, quasi-experimental design (QED) to collect evidence on the potential impact of students' interactions with the Pearson product on key student outcomes, including completion, academic achievement and progression outcomes. SRI has recruited four higher education institutions and 16 instructors who have used MasteringChemistry for at least two years. This one-year (two semesters) study will employ the following data collection approaches: (a) surveys of students at the beginning and end of each semester to gather data related to students' background characteristics, academic intentions, motivations and attitudes that may in turn influence the use of MasteringChemistry, and that also may impact on course completion, achievement and progression; (b) observations of classrooms that are using MasteringChemistry; (c) interviews of instructors to gather information

about their course goals and how MasteringChemistry is implemented; (d) end of course surveys of instructors to gather their feedback about the utility, advantages and disadvantages associated with the use of MasteringChemistry; and (e) student records provided by the institutions in order to abstract information related to students achievement and to perform the propensity score matching procedures.

Study 2: A Quasi-Experimental Study of the Efficacy of MasteringChemistry	
Intended Start Date	December, 2015
Anticipated Length of Study	Until end of 2017
Type of Study	QED - Case Control
Research Leads	Christine Leow, PhD; Carmen Arroyo, PhD; Betsy Nixon - Pearson in collaboration with the instructor at the institution.
Intended Sample Size	100 students from one institution
Description of Sample	MasteringChemistry users from the institution
Outcomes to be Measured	Experience, Completion, Achievement, and Progression

Proposed Study Description

Researchers in the Pearson Efficacy, Research and Impact Evaluation team are conducting a case-control quasi-experimental study in collaboration with the Pearson Customer Experience and Engagement Team and instructors at two institutions where MasteringChemistry is used. The goal of this study is to isolate the contribution of MasteringChemistry to student success when all other factors that research has indicated are important predictors of student success are taken into account and are controlled for, statistically. Specifically, this two-year longitudinal study will take into account factors that impact students' academic achievement and progression and can differentiate students who succeed from those who fail to achieve and progress. Such factors include: students' prior academic experiences, their overall approach to academic work, out-of-school responsibilities, the type and intensity of academic advising students receive when they enroll in college, and the support they receive within their developmental courses. The first results from this study will be available in April 2016.

Study 3: Lab-based Randomized Control Trial (RCT) on Efficacy of Adaptive/Dynamic Modules	
Intended Start Date	March, 2016
Anticipated Length of Study	1 week to collect data
Type of Study	Lab-based RCT; students will be randomly assigned to complete a learning activity in MasteringChemistry either with or without a dynamic study module component. Both groups will complete pre- and post-tests, and their performance will be compared.
Research Leads	Dan Belenky, PhD - Pearson in collaboration with faculty at the sponsoring university
Intended Sample Size	150
Description of Sample	Sample will be drawn from beginning chemistry students at a small state University
Outcomes to be Measured	Focus will be on achievement, with measures of experience and completion also collected. As it is a short lab-study, the goal is to extrapolate the findings regarding learning to predict the kinds of outcomes we may expect over a longer period.

Proposed Study Description

One of the newer adaptive features in MasteringChemistry are the dynamic study modules. These quiz students on content in a way that is designed to help improve learning and retention. Specifically, these questions are multiple-choice, and students are given two ‘votes’ to use. If they use both on one response, they are indicating a high-level of confidence in that option. If they are somewhat unsure, they may choose two of the options. This kind of approach allows for better scaffolding of a student’s metacognition, or their own ability to monitor their own understanding and learning. Feedback is provided after a block of these multiple-choice questions are completed, and this slight delay in providing feedback helps reactivate memory. Finally, based on student performance, an adaptive engine underlying the system can choose which information students still need additional practice with, and serve up content related to that topic, specifically, helping students to achieve mastery.

During Spring 2016, researchers in the Global Product Efficacy and Research Team will partner with faculty at a small northeastern state university to conduct a randomized controlled trial designed to assess the efficacy of the modules. This study will be conducted as a laboratory experiment. Students will be randomly assigned to one of two groups - a group that completes a learning activity in MChem that includes a dynamic study module activity, and a group that will complete the same learning activity without the dynamic study modules.

Study 4: Quasi-experimental on Efficacy of New Assessment Tool to Predict Students At Risk From Course Drop-out	
Intended Start Date	January, 2016
Anticipated Length of Study	Until end of Spring 2016
Type of Study	Quasi-experimental
Research Leads	Efficacy and Research Team in collaboration with Mastering product management team
Intended Sample Size	200 from two institutions
Description of Sample	Mastering product users and non-users from the 2 institutions
Outcomes to be Measured	Experience, Completion, and Achievement

Proposed Study Description

Improving the academic achievement of students struggling in class is one of the greatest challenges faced by instructors. Pearson has developed a pilot version of 'early alerts' that could indicate to instructors which students might be struggling so that in-time help might be provided. To investigate the usefulness of the early alerts and its potential influence on student outcomes, Pearson has partnered with two instructors from two institutions for this pilot. This pilot study will explore benefits of the early alerts, how early alerts affect the way instructors teach, and the relationship between the use of the early alerts and student achievement. The study will use a mixed-method design comprised of: (a) an instructor survey including questions about the background characteristics and teaching experiences of instructors and about the course(s) in which the early alerts capability is used; (b) weekly follow-up surveys that record instructors' experiences with the early alerts; (c) analysis of platform data to assess functionality of the early alerts algorithm, and; (d) collection and analysis of student end-of-course grade data provided by the institution.

References

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, *70*(2), 181-214.

Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science, 21*(5), 279-283.

Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121-152.

Corbett, A. T., Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4*(4), 253-278.

Cuadros, J., Yaron, D., & Leinhardt, G. (2007). "One firm spot": The role of homework as a lever in acquiring conceptual and performance competence in college chemistry. *Journal of Chemical Education, 84*(6), 1047-1052.

Huesgen, B. D. (2012). *Impact of an online homework program on the achievement of college students enrolled in the first semester General Chemistry* (Unpublished doctoral dissertation). University of Missouri-St. Louis, St. Louis, MO.

Karpicke, J., & Roediger III, H. L., (2010). The critical importance of retrieval for learning. *Science, 319,* 966-968.

Roediger III, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Science, 15*(1), 20-27.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*(1), 59-89.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197-221.