JOURNAL OF CHEMICAL EDUCATION
Article
pubs.acs.org/jchemeduc

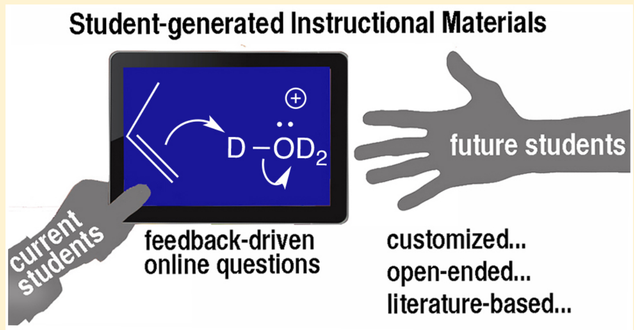
Using Student-Generated Instructional Materials in an e-Homework Platform

Danielle M. Zurcher, Sameer Phadke, Brian P. Coppola,* and Anne J. McNeil*

Department of Chemistry and Macromolecular Science and Engineering Program, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109-1055, United States

Supporting Information

ABSTRACT: Feedback-driven online homework systems provide students with a comprehensive set of practice questions that can accompany and enhance other instructional resources. However, the available e-homework systems do not contain content that aligns well with our course objectives, provide too few questions in key areas, and use assessment format(s) that do not match the ones used on our exams. Motivated to create our own questions, we used this gap as an opportunity to engage students in constructing and reviewing course-aligned content within a commercial e-homework platform. The students successfully generated approximately 1,000 largely open-ended organic chemistry questions, some with mechanistic and structural drawing capabilities, by modifying old exam questions. The students' questions spanned a variety of cognitive levels that skewed, as intended, toward skill-building. According to our assessment scheme, 75% of the questions were evaluated to be of the highest quality. As a consequence, we advocate that collaborating with undergraduate students in a "teaching team" can be a broadly useful way for instructors to generate high-quality instructional materials aligned with their course content.



KEYWORDS: *First-Year Undergraduate/General, Second-Year Undergraduate, Organic Chemistry, Computer-Based Learning, Internet/Web-Based Learning, Problem Solving/Decision Making, Student-Centered Learning, Mechanisms of Reactions*

INTRODUCTION

Instructors provide homework assignments to help guide engagement with course material and promote learning and retention. Online homework has begun to replace traditional paper-based homework in many large introductory courses because it encourages active learning, provides students with immediate feedback, and reduces faculty time spent on grading.[1−3] Researchers have studied online homework systems for organic chemistry and found a correlation between online homework performance and final grade,[4] as well as students' perception of its usefulness.[5] In addition, online homework use has led to improved class averages.[6−8] In contrast, Smithrud and Pinhas found that high online homework completion rates did not always lead to exam success and suggested that combining aspects of traditional homework and online homework is better.[9] Qualitative findings from these studies report positive student perceptions of online homework and high levels of perceived helpfulness in learning course content.[1,4−6] The learning advantages are hypothesized to be due to the students' ability to rework problems based on the immediate feedback they receive. Feedback reduces "discrepancies between [students'] current understandings/performance and a [desired] goal."[10] However, the type of feedback given affects learning outcomes.[10] For example, giving students information/hints that guide them toward an answer was more helpful than telling students their response was correct/incorrect.[11] For feedback to be effective it must be timely, specific, and clear.[10] Online homework can provide these types of effective feedback by suggesting ways students could fix an incorrect response.

Historically, we have not used online homework systems in the introductory organic chemistry program at the University of Michigan (U-M). We opted instead to emphasize peer-based and peer-led options, which take advantage of our highly residential, campus-based environment.[12] Moreover, most online e-homework systems are not aligned with the course assessments in our organic sequence: our exams require the students to answer open-ended questions about new and unfamiliar literature-based examples (Figure 1).[13,14] Our rationale for literature-based questions is 2-fold: through this testing strategy, we are transmitting as clearly as we can that students need to transfer knowledge gained during study to these unfamiliar examples, encouraging them to rely less on rote memorization and more on developing their reasoning skills.[15,16] In addition, by connecting what they learn in class to a virtually endless supply of real-world examples we aim to increase the students' appreciation for the course relevance.[17,18]

Received: May 25, 2016
Revised: September 18, 2016
Published: October 7, 2016

© 2016 American Chemical Society and Division of Chemical Education, Inc.

DOI: 10.1021/acs.jchemed.6b00384
J. Chem. Educ. 2016, 93, 1871−1878

Complete the following step, which was reported for the synthesis of Aspidospermine (a biologically interesting compound; *J. Org. Chem.* **2000**, *65*, 2642). In the first answer space, place the single structure **Y** implied by mechanistic arrows associated with the structure **X** (**Y** has all closed shell atoms). By examining the structure of the ultimate product **Z**, you should also be able to draw the mechanistic arrows needed to go from structure **Y**, with Bronsted base "B:" to give **Z**.

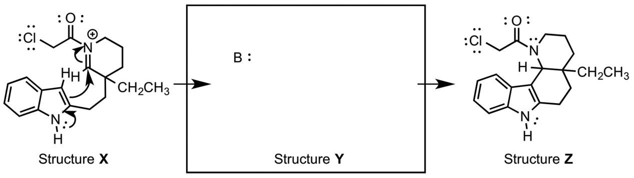


Figure 1. A representative exam question.

Our students understand that just becoming familiar with textbook questions or past exams is not a useful study strategy.

One obvious challenge with using high-level questions on exams is that deficiencies in lower-level skills become a significant barrier. For example, we want to aim at 100% of students achieving literacy in the curved arrow formalism, given the instructional reliance on it; however, it is realistic to assume some students are burdened by their deficient understanding of this topic and so their struggle to learn increases throughout the course.[19] Ideally, we want every student to master the curved arrow formalism before we ask them to follow and construct arguments using it. We believe students that carry a deficiency in their skill-based knowledge can benefit from a large number of questions in certain topical areas.

On the basis of the need described above, we reasoned that the custom use of a feedback-driven online homework platform might be an ideal vehicle for such questions. Such customization would enable us to create skill-based questions that are aligned with our course assessments (i.e., open-ended, literature-based). Several organic chemistry online homework platforms have been developed in recent years; each with a similar core functionality and differing specialized functionality.[20] For a number of reasons, specific to our needs, we decided to use the Sapling Learning system.[21] This platform is user-friendly and able to handle open-ended questions with generative answers. Moreover, its authoring system was robust and particularly amenable to novice users. This latter point was nontrivial for us because we were interested in teaching undergraduate students to be the content-generators. Although the method described herein is optimized for the Sapling Learning system, it could be implemented using a variety of online homework platforms with some modifications to suit the needs of the instructor and institution.

We chose students (rather than instructors) as content-generators because they can generate a significant quantity of material in a short time and the feedback they provide is from a student's perspective. Involving students in generating instructional materials encourages them to engage with both the course content and their peers, and the authentic audience (i.e., future students) lends purpose to their work.[22,23] Although the long-term goal was to release these student-generated questions to future students, we first determined whether students could be trained to generate high quality questions, with structural drawing and mechanistic functionality, in an online homework system. The answer is a resounding "yes." Herein, we describe a successful model for engaging students in generating instructional material to get course-aligned content. Furthermore, we demonstrate with various metrics that the question quality was high.

■ APPROACH

e-Homework Platform

PeerWise is perhaps the most identifiable leader in the area of student-generated questions, but it is limited to multiple-choice questions.[24,25] As alluded to above, our choice of Sapling Learning was purely based on functionality for our intended goals (content and generation method). This system provided access to traditional formats (i.e., multiple choice, fill-in-the-blank) along with handling student-generated responses to open-ended questions with an organic-friendly interface (i.e., structural drawings and mechanistic functionality).[21] Moreover, the authoring interface is user-friendly, which facilitated student training, and the company was open and receptive to working with us on this experiment. The Sapling Learning platform offers multiple forms of feedback (e.g., specific, default, and solution) that can be integrated into each question. For example, specific feedback is triggered when the incorrect answer was anticipated and programmed in with hints associated with the specific misconception. In contrast, the default feedback is triggered if an incorrect answer does not match one of the anticipated answers. An example of specific and default feedback from a student-generated question is given in Figure 2. When students view the correct answer they are also provided with an explanation (Figure 2). All questions created by our students include these three types of feedback (for additional student-generated questions see Supporting Information (SI)).

Project Structure

For our pilot project, we engaged the students in the Structured Study Group (SSG) program, a supplemental instruction option for students enrolled in first-semester organic chemistry.[26] Thus, students self-selected into SSG and were not selected based on performance. SSG classes meet 2 h per week for 15 weeks and are led by junior- and senior-level undergraduate students. Most students who participated in SSG during our pilot were first- and second-year students ($N = 142$, Fall 2013). During the pilot project, students generated questions as just one assignment in the SSG curriculum (see SI for pilot project syllabus). This format, unsurprisingly, did not provide sufficient time for multiple rounds of review and refinement, but it did provide proof of concept and allowed us to rapidly understand some good lessons about working with student authors. Because our practical goal was to generate high quality questions in a short, controlled period with the intent of releasing these questions to the entire population of students taking the organic courses ($N \sim 1400$ in the Fall and ~600 in the Winter, for the first-term course offering), we moved to the next phase of our project.

To continue with using student authors, we asked interested students to enroll in a one-credit independent study course (*Teaching Experiences for Undergraduates*) where the work would be focused solely on authoring questions. We invited students who generated high quality questions in the pilot project to enroll in this course. We imagined that involving fewer students with more intensive effort would naturally limit the number, and the 4 junior- and senior-level group leaders recommended that groups of 6−8 would be ideal, from their experience in the pilot. The number of participating students was indeed significantly less, due to a combination of intrinsic interest, time demand, and the need to enroll in a separate course (Winter 2014, $N = 31$; Fall 2014, $N = 12$; Winter 2015, $N = 16$). Thus, these content-generators met once a week for 1

DOI: 10.1021/acs.jchemed.6b00384
J. Chem. Educ. 2016, 93, 1871−1878

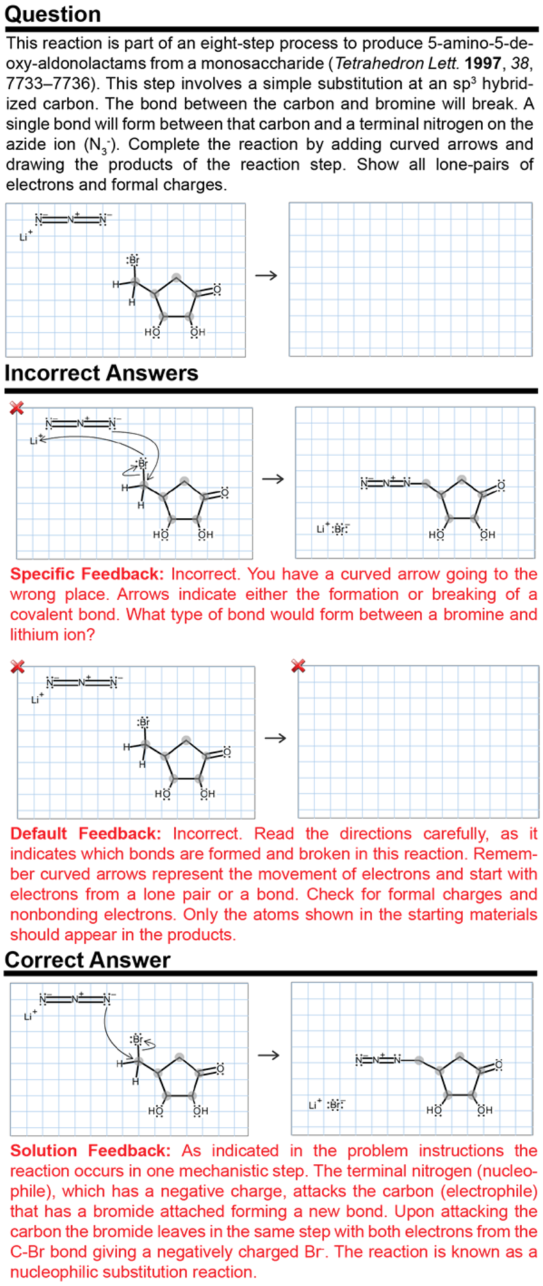


Figure 2. Student-generated question with incorrect and correct answers as well as the feedback.

h, and each class was led by one of the junior- and senior-level undergraduate SSG leaders. Participating students received credit for constructing questions, incorrect answers, and feedback (see SI for full project syllabi). In addition, each student earned \$250 if they programmed their questions, answers, and feedback into the Sapling Learning platform. Although this approach creates certain limitations (a small set of students generates questions, funding was used, etc.), we do not think these are critical criteria for creating student-generated instructional materials, but merely represent how we implemented the idea. Although the funding came from internal initiatives aimed at improving undergraduate education at the University of Michigan, one could implement this project by rewarding the students with additional credit hours (or points) rather than pay. We have no basis for prescribing how this project might work in different contexts, and we think this decision is an inherently local one.

Experienced faculty instructors defined a set of organic chemistry skills that were imagined to benefit students through having access to a large set of targeted practice problems. Something we found attractive about a software-based environment was its intrinsic mercilessness—in the case of basic skills, getting things "right" or "wrong" with no option to forgive oneself for small errors is intuitively appealing. The skill-based topics identified for the first semester course were: curved-arrow notation, resonance, acid—base chemistry, individual stereochemistry relationships, comparative stereochemistry relationships, electrophilic addition, elimination, substitution, transition states, electrophilic aromatic substitution, reaction mechanisms, and aromaticity. For the second semester, during which many more reactions are presented, skill-based questions included topics from epoxide chemistry, aldehyde/ketone chemistry, acyl transfer reactions, enolate chemistry, Diels—Alder reactions, peptide chemistry, and carbohydrate chemistry.

Content Creation

For this report, however, we return to the question of using students to generate this instructional content. The training period for the pilot project in the SSG course took 5 weeks and is outlined in Figure 3. Students were first trained to create usable questions, answers, and feedback and then learned how to program the questions into Sapling Learning. Training started with introducing the students to the Sapling Learning interface, followed by dividing them into small groups and giving them the same example question to solve (SI). They were instructed to create an array of plausible incorrect answers with feedback and a general feedback response that hinted at key concepts. After sharing the incorrect answers and feedback with the class a set number were selected. Finally, each group generated the example question in Sapling Learning, peer-reviewed the question, and addressed suggested edits.

After the training period, students were tasked with generating new questions following four different curved arrow formats and using literature sources for inspiration (SI). Overall, 172 questions were generated and subjected to an internal review process (SI). First, their classmates reviewed the questions guided by a rubric and suggested edits. Once the edits were complete, the questions were reviewed by the class leader and finally by a graduate student overseeing the project, who evaluated them on a pass/fail scale post hoc. Each question was evaluated for chemistry, grammar, clarity, programming, accuracy of written feedback, and functionality in the interface. In the large-scale first round, only 64/172 (37%) questions passed this internal review (Table 1). The majority of questions failed the graduate student's review due to incorrect chemistry, which we speculated might stem from the students' inexperience in creating questions from the literature (e.g., unable to distinguish whether a reaction occurred under acidic or basic conditions, or if the mechanism was concerted or stepwise). We concluded that having the students use the primary literature as an inspiration for questions was not a viable approach.

We made several changes to address the shortfalls of the pilot project. First, we replaced primary literature with old exams as the source. Because the vast majority of our exam problems are literature-based, faculty instructors, by default, have already done the most difficult part of the screening process. Also,

DOI: 10.1021/acs.jchemed.6b00384
J. Chem. Educ. 2016, 93, 1871–1878

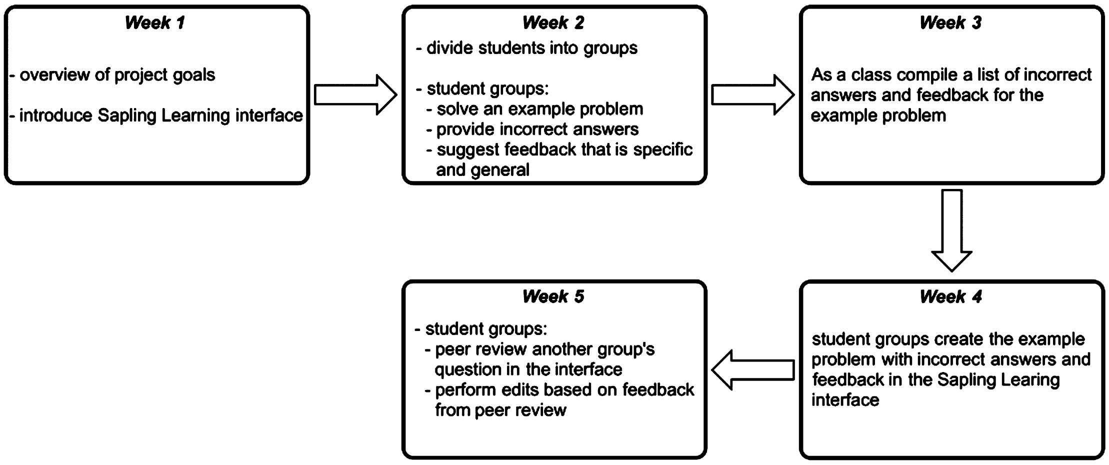


Figure 3. Time line for training students to generate questions in Sapling Learning.

Table 1. Number of Questions Generated and Publishable after the Internal Review Process

questions	pilot (F 2013)	1-credit course (W 2014)	1-credit course (F 2014)	1-credit course (W 2015)	total
generated	172	639	192	290	1,293
publishable (% passed)	64 (37)	627 (98)	167 (87)	256 (88)	1,114 (86)

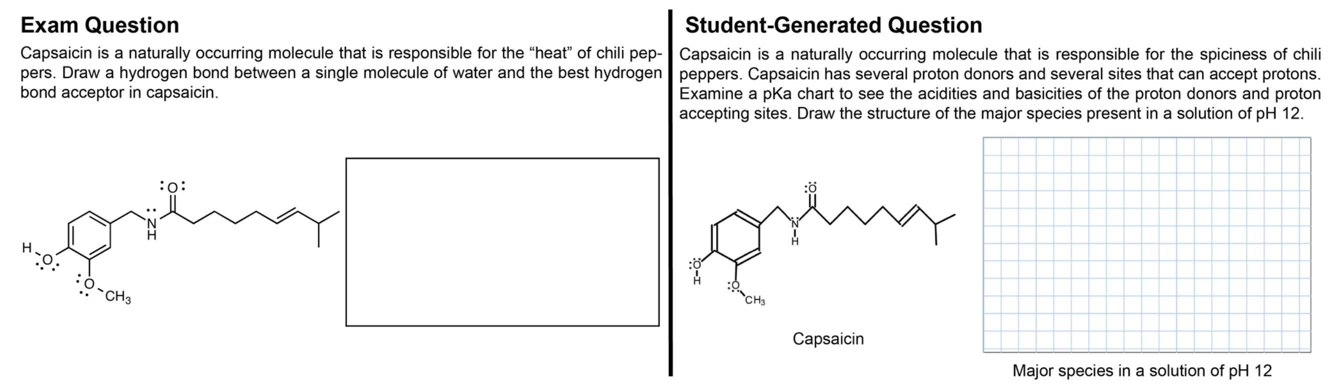


Figure 4. Original exam question (left) and student-generated question (right). Inspiration from literature source.[28]

Table 2. Differences between Each Iteration of the Project

parameters	pilot (F 2013)	1-credit course (W 2014)	1-credit course (F 2014)	1-credit course (W 2015)
source material	literature journal	old test questions	old test questions	old test questions
course	organic chemistry I	organic chemistry I	organic chemistry II	organic chemistry II
# of participants	142	31	12	16
# of questions generated per student or group	3	20	16	18

because we never repeat questions, which are generated freshly by teams of instructors for every exam, we have a deep repository of retired exam problems. Instead of creating problems *de novo*, students were instead tasked with modifying an old question to create a new question, and to focus on highlighting one of the given skills we were after. For example, an exam question focused on hydrogen-bonding might be transformed into an acid–base question (Figure 4). We reasoned that this approach enables students to generate an original question, but with chemistry that was previously vetted for the course. The second key change was moving from the SSG format to a stand-alone 1-credit course, meeting 1 h each week (Table 2). Students who participated in the first stand-alone course began generating content immediately because they were previously trained during the pilot project. Each student was responsible for generating two questions per week, where a different skill-based topic was selected for each week.[27] In the pilot, generating questions was just a small part of the SSG curriculum; it was never meant to be the dedicated mechanism for generating content. In the dedicated course, students also spent 3 weeks reviewing and editing questions. With this accelerated timeline, students generated 639 questions over a single semester, with 98% of the questions created deemed usable after internal review (Table 1). The course-based strategy was implemented two more times, in which 87% (Fall 2014) and 88% (Winter 2015) of questions passed the internal review. (These later two courses focused on creating content for the second semester of organic chemistry.) Alternative strategies that could work in other contexts include reducing the number of questions generated over a semester or using a longer time scale.

DOI: 10.1021/acs.jchemed.6b00384
J. Chem. Educ. 2016, 93, 1871−1878

QUESTION QUALITY

External Review

To better assess the quality of questions generated over the four semesters, several additional rounds of review were carried out (Figure 5). The questions that passed the above-mentioned internal review were evaluated using two types of external review by a Sapling Learning technician. First, 703 questions were evaluated based on functionality in the interface, of which 677 (96%) passed the review (Table 3). Then, a more in-depth review was carried out on a randomly selected set of 113 questions. Although 31 (27%) needed no further edits, 62 (55%) contained technical and formatting problems and 20 (18%) contained content errors. Technical and format issues included insufficient programming with respect to benzene rings (i.e., programming in just one resonance form for each answer), improper bond angles, unclear feedback, and inefficient question layout. Looking more closely at the content errors, only 4 questions actually had incorrect chemistry (e.g., wrong regioselectivity in a Diels–Alder reaction), whereas the other 16 questions contained missing counterions or lone pairs. In response, the peer review criteria and authoring manual were updated to explicitly draw attention to these common errors (SI). These updated criteria should negate the need for an external review in future iterations, rendering this method independent of the choice of platform.

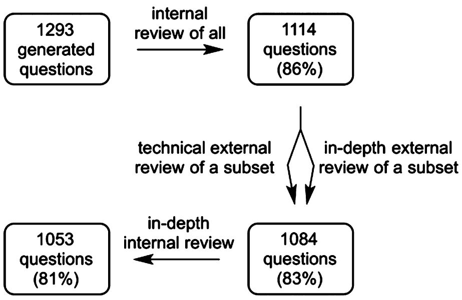


Figure 5. Overview of the different stages of review. (The percentages refer to the original set of 1293 questions.)

Table 3. Quality of Questions Determined by the External Review

type of review	# of questions submitted	# of questions with formatting/technical issues	# of questions with content errors	# of questions passed
external technical (% passed)	703			677 (96)
external in-depth (% of questions)	113	62 (55)	20 (18)	31 (27)

In-Depth Internal Review

To further probe question quality, we performed an in-depth internal analysis of the 1084 questions that passed all prior reviews. Undergraduate students were hired and trained to classify each question on a set of criteria (described below) using a method reported by Bates and co-workers, who used undergraduates to reliably evaluate questions in Peerwise.[25c] Our undergraduate raters were given access to the question stem, solution, and incorrect responses. Their analytical calibration was directed by two postdoctoral scholars over 3 weeks, wherein 40 questions were independently reviewed on a set of six criteria and the results discussed at weekly meetings. After calibration, the absolute percent agreement for each criterion was calculated on a set of 20 randomly selected questions comparing each undergraduate rater to the postdoctoral scholars, and was found to be >80% (SI).[29] For the criterion where each question was categorized into the revised Bloom's taxonomy, Cohen's kappa was also calculated and the inter-rater agreement was found to be >0.71 for each undergraduate (SI). The questions with rating discrepancies between undergraduate students were re-evaluated as a group that included the postdoctoral scholars to make a final determination.

Each question was first assessed on whether the solution was correct and programmed accurately. Of the 1084 questions evaluated, 17 questions (1.5%) were found to have an inaccurate solution and 14 questions (1.3%) had solutions that were insufficiently programmed. These 31 questions were removed from the database and not evaluated further. The remaining 1053 questions were evaluated for question clarity and categorized into different cognitive levels using Bloom's taxonomy.[30,31] Additional criteria were used to evaluate the solution explanation, specific feedback, and default feedback (SI).

Clarity of each question was evaluated on a binary scale (yes/no). We found 16 questions were not clearly worded, with 10 of those questions on the topic of resonance. These questions were immediately fixed through adding and/or deleting a phrase. Each question was then classified into a cognitive level based on the revised Bloom's taxonomy (Table 4 for descriptions of each level), which is a frequently used method to classify the objectives of student assignments.[31] The revised Bloom's taxonomy hierarchy starts with lower-order thinking skill categories (e.g., remember) and goes to higher-order thinking skill categories (e.g., create). Questions within our database were in four categorization levels of Bloom's taxonomy: remember, understand, apply, and analyze. No questions were found at either the evaluate or create levels. A breakdown of each topic with respect to Bloom's taxonomy can be found in the SI. Some topics had questions across the four

Table 4. Bloom's cognitive categorization levels[a]

level	description	number of questions at each categorization level (%), $N = 1053$
remember	factual recall, remembering facts and information	272 (26)
understand	understand information, interpret facts, construct meaning from text or graphs	132 (13)
apply	use information, methods, concepts, and theories in new situations to solve problems	427 (40)
analyze	see patterns, break down material into parts, identify components	222 (21)
evaluate	put together elements or parts to form a whole and forming judgements based on criteria and standards through checking and critiquing	0
create	compare and discriminate between ideas, assess value of theories, judge and check the value of material for a given purpose	0

[a]Descriptions adapted from ref 31.

DOI: 10.1021/acs.jchemed.6b00384
J. Chem. Educ. 2016, 93, 1871–1878

levels (e.g., resonance) whereas others had questions in only one level (e.g., transition states) (Table 4).

The analytical results were quite satisfying, as we intended our student authors to create more fundamental, skill-based questions rather than to mirror the demands on the exam questions themselves. To test this, we repeated the categorical analysis on the last eight organic chemistry final exams given at U-M, which showed that questions were skewed toward the higher-order cognitive levels compared with the student-generated practice questions (Figure 6 and SI).

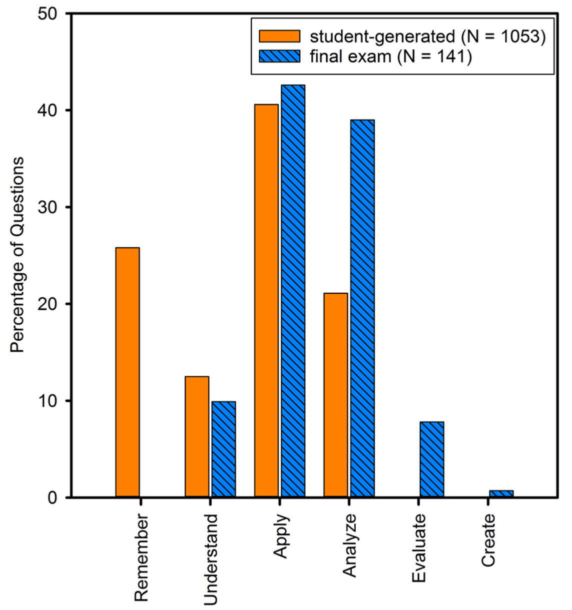


Figure 6. Percentage of student-generated questions (orange) and final exam questions (blue) in each category of Bloom's taxonomy.

The quality and type of feedback was also evaluated because providing useful feedback was an important objective. Solution explanations were classified as either good, minimal, insufficient, or missing (Table 5 for descriptions). The vast majority of solution explanations (95%) were classified as minimal or good (Table 5). A remarkable 71% (34/48) of the questions that were missing a solution explanation were on the topic of curved arrows, likely because these questions were generated during the pilot project, which had a less rigorous review process (SI).

Table 5. Solution Explanation Categorization Levels[a]

identifier	description	number of questions at each identifier (%), N = 1053
missing	no explanation provided	48 (4.6)
insufficient	wrong reasoning	1 (0.1)
minimal	insufficient explanation or justification of the correct answer such that some parts may be unclear or incorrect	40 (3.8)
good	clear and adequately detailed description the correct answer	964 (91.5)

[a]Descriptions adapted from ref 25c.

The specific feedback was categorized by whether the response addressed a misconception, hinted at a minor error (e.g., missing lone pairs), or had an explanation that was incorrect or missing. For 1053 questions, 5403 specific feedback responses were generated, with a range of 1−16 per question. Overall, 88% of specific feedback responses addressed a misconception, 11% hinted at a minor error and 1% were missing an explanation or had an incorrect explanation. The default feedback was similarly categorized and we found 97% of questions provided a general hint with the other 3% had an explanation related to minor errors (e.g., missing counterions) or the explanation was missing/incorrect.

Although all 1053 questions contained correct chemistry, the overall question quality was evaluated using stricter criteria. Thus, the question quality was rated to be high, medium or low on a set of criteria described in Table 6. Overall, 964 of the 1053 questions evaluated (i.e., 92%; or 75% of all questions generated) were considered high quality. To improve the question quality, students currently enrolled in the SSG program are continuing to edit/revise questions. For example, they are asked to generate additional specific feedback responses and, at the same time, address formatting issues. Future efforts will focus on whether feedback quality can be improved by incorporating animations, mechanistic drawings, or additional information beyond text.

■ STUDENT PERCEPTIONS

We surveyed students who participated in the 1-credit course to assess several aspects of the project.[32] Students were asked about their experience in using the Sapling Learning interface to create questions. Students reported that programming in questions was mostly straightforward, but reported lower favorability with respect to specific interface qualities such as ease of use, utility (i.e., ability to perform several functions), and room for creativity (SI). Altogether, most students reported spending less than 1 h programming each question and associated feedback into the interface. (SI).

■ SUMMARY

Herein, we generated a successful model for a process wherein students generate instructional material for their peers. An extensive review process revealed that the 75% of the questions generated were rated as "high" quality. As a consequence, this approach may be used as-is or modified by other instructors who want to better align instructional materials with their specific course, including open-ended questions. Advantages of

Table 6. Overall Question Quality

measure	description	number of questions (%), N = 1053
high	correct solution, at least minimal solution explanation, clearly worded, no specific feedback responses were incorrect/missing, default feedback offered a general hint	964 (92)
medium	correct solution but one other aspect was not up to par; that is, not clearly worded, missing a solution explanation, one specific feedback response missing, or default feedback offered a programming hint	65 (6)
low	correct solution but more than one other aspect was not up to par; that is, not clearly worded, missing a solution explanation, one specific feedback response missing, or default feedback offered a programming hint	24 (2)

DOI: 10.1021/acs.jchemed.6b00384
J. Chem. Educ. 2016, 93, 1871−1878

using students as content-generators is that approximately 1000 usable questions were generated over just four semesters. These questions are now being utilized by our currently enrolled students. Efforts to assess the effectiveness of this new resource on student learning are underway. Aspects of what we are reporting are contextual to our setting (no traditional publisher material, idiosyncratic course, infrastructure to support group-based instructional development). However, we have provided a framework for engaging students in content-generation, including suggestions for instructors to imagine what, in their own settings, might be accomplished by partnering with students to solve whatever problem might be vexing their teaching program. Finally, we note that this project aligns well with the history of technology development, where content generation is successfully moved from a centralized authority to more decentralized and localized settings, resulting in a broader impact.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available on the ACS Publications website at DOI: 10.1021/acs.jchemed.6b00384.

> Specific assignment and course descriptions, topic handouts, examples of student-generated questions, and survey results. (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: ajmcneil@umich.edu.
*E-mail: bcoppola@umich.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge the students and SSG leaders who participated. (SSG Leaders: Nicholas Carducci, Anitha Menon, Kevin Zhang, Alex Kokaly, Alex Blaty, Fahad Sarvari, Xiao Wang, Matt Thimm, Amber Young, Lawrence Chen, Paul Parker, and Lynn Daboul.) We thank the Sapling Learning technicians and staff, as well as our colleagues at Michigan for their helpful discussions. This work was partially supported by the Howard Hughes Medical Institute (HHMI) through an HHMI Professors Program grant to A.J.M. We also thank the University of Michigan for partial support of this work via a Transforming Learning for the Third Century Grant, College of Literature Science and Arts Instructional Technology New Initiatives/New Infrastructure Grant and Level II Grant.

■ REFERENCES

(1) Parker, L. L.; Loudon, G. M. Case Study Using Online Homework in Undergraduate Organic Chemistry: Results and Student Attitudes. *J. Chem. Educ.* **2013**, *90*, 37−44.

(2) Richards-Babb, M.; Drelick, J.; Henry, Z.; Robertson-Honecker, J. Online Homework, Help or Hindrance? What Students Think and How They Perform. *J. Coll. Teach.* **2011**, *40*, 81−93.

(3) Butler, M. B.; Zerr, R. J. The Use of Online Homework Systems to Enhance Out-of-Class Student Engagement. *Int. J. Technol. Math. Educ.* **2005**, *12*, 51−58.

(4) Richards-Babb, M.; Curtis, R.; Georgieva, Z.; Penn, J. H. Student Perceptions of Online Homework Use for Formative Assessment of Learning in Organic Chemistry. *J. Chem. Educ.* **2015**, *92*, 1813−1819.

(5) Chamala, R. R.; Ciochina, R.; Grossman, R. B.; Finkel, R. A.; Kannan, S.; Ramachandran, P. EPOCH: An Organic Chemistry Homework Program That Offers Response-Specific Feedback to Students. *J. Chem. Educ.* **2006**, *83*, 164−169.

(6) Malik, K.; Martinez, N.; Romero, J.; Schubel, S.; Janowicz, P. A. Mixed-Methods Study of Online and Written Organic Chemistry Homework. *J. Chem. Educ.* **2014**, *91*, 1804−1809.

(7) Penn, J. H.; Nedeff, V. M.; Gozdzik, G. Organic Chemistry and the Internet: A Web-Based Approach to Homework and Testing Using the WE_LEARN System. *J. Chem. Educ.* **2000**, *77*, 227−231.

(8) Chen, J. H.; Baldi, P. Synthesis Explorer: A Chemical Reaction Tutorial System for Organic Synthesis Design and Mechanism Prediction. *J. Chem. Educ.* **2008**, *85*, 1699−1703.

(9) Smithrud, D. B.; Pinhas, A. R. Pencil−Paper Learning Should Be Combined with Online Homework Software. *J. Chem. Educ.* **2015**, *92*, 1965−1970.

(10) Hattie, J.; Timperley, H. The Power of Feedback. *Review of Educational Research* **2007**, *77*, 81−112.

(11) Bangert-Drowns, R. L.; Kulik, C.-L. C.; Kulik, J. A.; Morgan, M. The Instructional Effect of Feedback in Test-Like Events. *Review of Educational Research* **1991**, *61*, 213−238.

(12) (a) Varma-Nelson, P.; Coppola, B. P. Team learning. In *Chemists' Guide to Effective Teaching*; Pienta, N. J., Cooper, M. M., Greenbowe, T. J., Eds.; Prentice Hall: Saddle River, NJ, 2005; pp 155−166. (b) Varma-Nelson, P.; Banks, J. PLTL: Tracking the trajectory from face-to-face to online environments. In *Trajectories of Chemistry Education Innovation and Reform*; Holme, T., Cooper, M. M., Varma-Nelson, P., Eds.; American Chemical Society: Washington, DC, 2013; pp 95−110. (c) Smith, J.; Wilson, S. B.; Banks, J.; Zhu, L.; Varma-Nelson, P. Replicating Peer-Led Team Learning in Cyberspace: Research, Opportunities, and Challenges. *J. Res. Sci. Teach.* **2014**, *51*, 714−740.

(13) Coppola, B. P. Literature-Based Examinations and Grading Them: Well Worth the Effort. In *College Pathways to the Science Education Standards*; Siebert, E. D., McIntosh, W. J., Eds.; National Science Teachers Association Press: Arlington, Virginia, **2001**; pp 84−86.

(14) Coppola, B. P.; Ege, S. N.; Lawton, R. G. The University of Michigan Undergraduate Chemistry Curriculum 2. Instructional Strategies and Assessment. *J. Chem. Educ.* **1997**, *74*, 84−94.

(15) Anderson, T. L.; Bodner, G. M. What can we do about 'Parker'? A case study of a good student who didn't 'get' organic chemistry. *Chem. Educ. Res. Pract.* **2008**, *9*, 93−101.

(16) Bhattacharyya, G.; Bodner, G. M. "It Gets Me to the Product": How Students Propose Organic Mechanisms. *J. Chem. Educ.* **2005**, *82*, 1402−1407.

(17) Schaller, C. P.; Graham, K. J.; Jones, T. N. Synthesis Road Map Problems in Organic Chemistry. *J. Chem. Educ.* **2014**, *91*, 2142−2145.

(18) Shea, K. M.; Gorin, D. J.; Buck, M. E. Literature-Based Problems for Introductory Organic Chemistry Quizzes and Exams. *J. Chem. Educ.* **2016**, *93*, 886−890.

(19) Grove, N. P.; Cooper, M. M.; Rush, K. M. Decorating with Arrows: Toward the Development of Representational Competence in Organic Chemistry. *J. Chem. Educ.* **2012**, *89*, 844−849.

(20) For other online homework platforms, see: (a) WileyPlus. https://www.wileyplus.com/WileyCDA/ (accessed Sept 2016). (b) Connect. http://connect.mheducation.com (accessed Sept 2016). (c) MasteringChemistry. http://www.pearsonmylabandmastering.com/northamerica/ (accessed Sept 2016).

(21) Sapling Learning. http://www2.saplinglearning.com (accessed Sept 2016). (Note that their website lists three modes of question authoring (1) content requests [authored by a Sapling Learning technician], (2) editing existing questions, and (3) writing questions from scratch.).

(22) For recent summaries, see: (a) Coppola, B. P. Do Real Work, Not Homework. In *Chemistry Education: Best Practices, Opportunities and Trends*; García-Martínez, J., Serrano-Torregrosa, E., Eds.; Wiley-VCH: Weineim, Germany, 2015; pp 203−258. (b) *Student-generated digital Media in Science Education: Learning, Explaining and*

DOI: 10.1021/acs.jchemed.6b00384
J. Chem. Educ. 2016, 93, 1871−1878

Communicating Content; Hoban, G., Nielsen, W., Shepherd, A., Eds.; Routledge: New York, 2015.

(23) For examples, see: (a) Shultz, G. V.; Winschel, G. A.; Inglehart, R. C.; Coppola, B. P. Eliciting Student Explanations of Experimental Results Using an Online Discussion Board. *J. Chem. Educ.* **2014**, *91*, 684−686. (b) Lawrie, G.; Bartle, E. Chemistry Vlogs: A Vehicle for Student-Generated Representations and Explanations to Scaffold their Understanding of Structure-Property Relationships. *Int. J. Innov. Sci. Math. Educ.* **2013**, *21*, 27−45. (c) Bottomley, S.; Denny, P. A Participatory Learning Approach to Biochemistry Using Student Authored and Evaluated Multiple-choice Questions. *Biochem. Mol. Biol. Educ.* **2011**, *39*, 352−361. (d) Evans, M. J.; Moore, J. S. A Collaborative, Wiki-Based Organic Chemistry Project Incorporating Free Chemistry Software on the Web. *J. Chem. Educ.* **2011**, *88*, 764−768. (e) Moy, C.; Locke, J. R.; Coppola, B. P.; McNeil, A. J. Improving Science Education and Understanding through Editing Wikipedia. *J. Chem. Educ.* **2010**, *87*, 1159−1162.

(24) PeerWise. http://peerwise.cs.auckland.ac.nz (accessed Sept 2016).

(25) For examples, see: (a) Ryan, B.; Raighne, A. M.; Casey, M.; Howard, R. Student Attitudes to an Online, Peer-instruction, Revision Aid in Science Education. *JPAAP* **2015**, *3*, 49−60. (b) Hardy, J.; Bates, S. P.; Casey, M. M.; Galloway, K. W.; Galloway, R. K.; Kay, A. E.; Kirsop, P.; McQueen, H. A. Student-Generated Content: Enhancing Learning Through Sharing Multiple-Choice Questions. *Int. J. Sci. Educ.* **2014**, *36*, 2180−2194. (c) Bates, S. P.; Galloway, R. K.; Riise, J.; Homer, D. Assessing the quality of a student-generated question repository. *Phys. Rev. St. Phys. Educ. Res.* **2014**, *10*, 020105−1−020105−11. (d) Bates, S.; Galloway, R. Student-Generated Assessment. *Educ. Chem.* **2013**, *50*, 18−21. (e) Denny, P.; Luxton-Reilly, A.; Simon, B. Quality of Student Contributed Questions Using PeerWise. In *Proceeding ACE '09, Proceedings of the eleventh Australasian Conference on Computing Education*; Australian Computer Society, Inc.: Darlinghurst, Australia, 2009; pp 55−63. (f) Denny, P.; Luxton-Reilly, A.; Hamer, J. Peerwise. In *Proceeding ACE '08, Proceedings of the tenth Australasian Conference on Computing Education*; Australian Computer Society, Inc.: Darlinghurst, Australia, 2008; pp 69−74.

(26) Coppola, B. P.; Daniels, D. S.; Pontrello, J. K. Using Structured Study Groups To Create Chemistry Honors Sections. In *Student-Assisted Teaching: A guide to Faculty-Student Teamwork*; Miller, J. E., Groccia, J. E., Miller, M. S., Eds.; Anker: Bolton, MA, 2001; pp 116−122.

(27) To create questions on each topic, students were given a set of "formats" that were developed to help student choose appropriate modules (i.e., structural drawing) in Sapling Learning. The specific formats can be found in the SI (i.e., Winter 2014 handouts and Fall 2014 handouts).

(28) Reyes-Escogido, M. L.; Gonzalez-Mondragon, E. G.; Vazquez-Tzompantzi, E. Chemical and Pharmacological Aspects of Capsaicin. *Molecules* **2011**, *16*, 1253−1270.

(29) Graham, M.; Milanowski, A.; Miller, J. *Measuring and promoting inter-rater agreement of teacher and principal performance ratings*; Technical Report for Center for Educator Compensation reform, Department of Education: Washington, DC, February 2012. Retrieved from http://files.eric.ed.gov/fulltext/ED532068.pdf (accessed Sept 2016).

(30) Bloom, B. S. *Taxonomy of Educational Objectives: The Classification of Educational Goals; Handbook I: Cognitive Domain*; David Mckay Co. Inc.: New York, 1956.

(31) *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*; Anderson, L. W., Krathwohl, D. R., Eds.; Longman: New York, NY, 2001.

(32) The research survey was classified as Exemption #1 by the University of Michigan Institutional Review Board for human subjects research (exempt ID: HUM00099765, 3/18/2015). Only students who participated in the 1-credit course were surveyed. Participants in the Winter 2014 and Fall 2014 classes were sent an online survey three months to one year after participating in the course (survey respondents = 12, 35% response rate). The students from in the Winter 2015 course were surveyed at the end of the semester (survey respondents = 16, 100% response rate).

DOI: 10.1021/acs.jchemed.6b00384
J. Chem. Educ. 2016, 93, 1871−1878