JOURNAL OF MICROBIOLOGY & BIOLOGY EDUCATION
DOI: https://doi.org/10.1128/jmbe.v19i1.1324

Research

Curricular Activities that Promote Metacognitive Skills Impact Lower-Performing Students in an Introductory Biology Course †

Nathan V. Dang, Jacob C. Chiang, Heather M. Brown, and Kelly K. McDonald
Department of Biological Sciences, California State University, Sacramento, CA 95695-6077

This study explores the impacts of repeated curricular activities designed to promote metacognitive skills development and academic achievement on students in an introductory biology course. Prior to this study, the course curriculum was enhanced with pre-assignments containing comprehension monitoring and self-evaluation questions, exam review assignments with reflective questions related to study habits, and an optional opportunity for students to explore metacognition and deep versus surface learning. We used a mixed-methods study design and collected data over two semesters. Self-evaluation, a component of metacognition, was measured via exam score postdictions, in which students estimated their exam scores after completing their exam. Metacognitive awareness was assessed using the Metacognitive Awareness Inventory (MAI) and a reflective essay designed to gauge students' perceptions of their metacognitive skills and study habits. In both semesters, more students over-predicted their Exam 1 scores than under-predicted, and statistical tests revealed significantly lower mean exam scores for the over-predictors. By Exam 3, under-predictors still scored significantly higher on the exam, but they outnumbered the over-predictors. Lower-performing students also displayed a significant increase in exam postdiction accuracy by Exam 3. While there was no significant difference in students' MAI scores from the beginning to the end of the semester, qualitative analysis of reflective essays indicated that students benefitted from the assignments and could articulate clear action plans to improve their learning and performance. Our findings suggest that assignments designed to promote metacognition can have an impact on students over the course of one semester and may provide the greatest benefits to lower-performing students.

INTRODUCTION

Metacognition is the process of being cognizant of, reflecting on and comprehending one's knowledge (1, 2). Simply put, metacognition is "thinking about thinking," and it can be critical to the process of learning (3). Two components of metacognition have emerged from research on the subject: *knowledge* and *regulation* (4). Metacognitive *knowledge* refers to one's awareness of cognitive processes, while metacognitive *regulation* refers to one's ability to act on that awareness (5).

Early research on metacognition was largely focused on students in the K–12 system (1), while more recent studies have concentrated on post-secondary students in the humanities and social sciences (1, 6–9). Although the use of metacognitive skills and strategies is assumed to be important across disciplines, few study designs allow for such comparisons. One report involving college undergraduates from multiple disciplines did find that motivation and certain cognitive strategies such as rehearsal, which facilitates learning and information recall, could be used to predict the academic performance of students in the natural sciences and social sciences, but not in the humanities (10). However, the researchers suggest their study may be confounded by the fact that science students tend to be taught through lectures and tested with multiple-choice questions, whereas humanities students more often engage in discussions and are frequently assessed through essays or free-response formats.

Regardless of discipline, strong metacognitive skills have been shown to correlate with student learning and academic success. For instance, students who achieve higher exam scores are often better at predicting their exam performance than their lower-scoring peers (8, 10, 11). Several studies have further shown that higher-performing students are more likely to underestimate their performance, while lower-performing students overestimate their academic abilities (1, 6, 7, 8). This phenomenon is an extension of

*Corresponding author. Mailing address: California State University, 6000 J Street, Sacramento, CA 95819-6077 Phone: 916-278-5836. Fax: 916-278-6993. E-mail: mcdonald@csus.edu.
Received: 28 March 2017, Accepted: 22 November 2017, Published: 16 February 2018
†Supplemental materials available at http://asmscience.org/jmbe

the Dunning-Kruger Effect, which states that individuals with minimal expertise in a given area or proficiency for a particular skill are unaware of their deficiencies (12). It presents a "double burden" because "[t]he act of evaluating the correctness of one's…response draws upon the exact same expertise that is necessary in choosing the correct response in the first place." Therefore, in situations that require self-evaluation, individuals exhibiting the Dunning-Kruger Effect will be less likely to recognize their mistakes and make appropriate corrections. It should also be noted that metacognitive *awareness* may be independent of academic achievement (4). In other words, a student may be aware of poor study skills (*knowledge*), but lack the ability to make the appropriate adjustments (*regulation*). In support of this, Miller and Geraci (7) found that lower-performing students, while still overestimating their exam performance, were less confident in their predictions than higher-performing students, suggesting they were aware of their metacognitive deficiencies.

Thus, the question becomes, how can instructors promote the development of metacognitive awareness and skills that lead to student success? Some researchers posit that metacognition must be taught directly, and a variety of strategies have been documented. Tanner (13) places emphasis on three components of metacognitive regulation, suggesting students approach class, assignments, and exams by taking the time to think about their cognitive processes before (Planning), during (Monitoring), and after (Evaluating) each activity. A student beginning a homework assignment, for example, could be prompted with questions such as "What is the goal of this assignment?" (Planning), "Is the approach I am taking working or not?" (Monitoring), and "Did I achieve the assignment's goals, and did my approach work?" (Evaluating). Ehrlinger and Shain (2) suggest involving students more deeply in the learning process by having them self-test, evaluate their learning while reading, and summarize texts. Furthermore, Nilson's *Creating Self-Regulated Learners* (14) provides a thorough guide with supporting evidence from the literature for incorporating activities into virtually every part of the curriculum.

The resources described above provide instructors with a variety of ideas and strategies for enhancing their curriculum with metacognitive activities; however, it is also important to evaluate the impact of these activities (15). Historically, metacognitive ability has been assessed by means of online testing modules or lengthy verbal interviews. Free-response survey questions and reflections have also been employed to determine the effectiveness of interventions that promote metacognition and gaining information about student thinking (5, 16–18). As an alternative, Schraw and Dennison (4) developed and validated the Metacognitive Awareness Inventory (MAI) as a means of measuring metacognitive awareness in adolescents and adults. The MAI is a 52-item survey, with each item corresponding to one of eight subcategories of Knowledge or Regulation (Fig. 1). Knowledge is classified as Declarative (knowledge of facts or of one's own skills and abilities), Procedural (knowledge related to the process of learning), or Conditional (knowledge related to the timing and rationale for applying certain learning strategies). Regulation includes Planning, Information Management, Comprehension Monitoring, Debugging, and Evaluation, which collectively relate to an individual's ability to set goals, process information, evaluate learning, and modify strategies when necessary. Since its development, the MAI has been used in a variety of settings as a self-report measure of metacognition. Doyle (19), for example, studied metacognition in a pre-nursing course and used the MAI before and after an intervention to measure metacognitive skills development. While a statistically significant increase in knowledge of cognition was observed in this study, no difference was detected for regulation of cognition or total MAI score. As a self-report tool, the MAI is limited to measuring a student's perceived level of metacognition, rather than their actual metacognitive ability. More direct measures of metacognition involve focusing on a specific component and prompting students to engage in the actual activity. For example, self-evaluation, an MAI sub-category related to regulation of cognition, has been measured using exam score postdictions, in which students estimate their exam scores after completing the exam but before receiving their scores (1, 6, 8, 20).

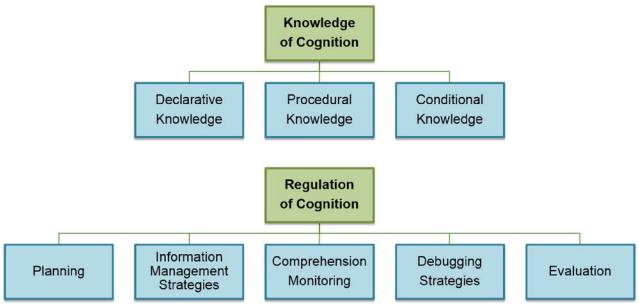

FIGURE 1. Metacognitive constructs and sub-categories measured by the Metacognition Awareness Inventory (MAI).

In an effort to strengthen students' metacognitive skills, particularly those related to self-evaluation, the authors of this study designed and embedded metacognitive activities throughout the curriculum of an introductory biology course. While students reacted positively and instructors felt the activities added value, there was no evidence supporting the effectiveness of the activities. Therefore, our research question sought to determine whether the specific metacognitive activities ("the intervention") could influence students' self-evaluation skills, learning strategies, and exam performance over the course of the semester. We hypothesized that 1) students' self-evaluation skills, as measured by exam score postdictions, would improve over the course of the semester and 2) that students who more accurately estimated their exam scores would score higher on the exams. Lastly, we hypothesized that higher scores on the MAI would correlate with higher exam scores and greater

postdiction accuracy. The MAI reflective essays were further used to gain insight into students' study habits/behaviors, their ability to monitor and modify their habits, and the challenges they faced related to studying.

METHODS

Course context and participants

BIO2: Introduction to Cells, Molecules, and Genes, at California State University, Sacramento, is the second course in a two-semester general biology sequence required for biology and biochemistry majors. The course includes two 75-minute lectures, one 3-hour laboratory, and one 2-hour activity (i.e., discussion) session per week. This study focused on the lecture component of the course. During the data collection semesters, the BIO2 lecture provided an active and collaborative learning environment intended to promote conceptual understanding of course material and the development of interpersonal, critical thinking, metacognitive, and study skills.

We examined the spring 2015 (n=119) and spring 2016 (n=52) cohorts of BIO2, hereafter referred to as SP15 and SP16, respectively. The SP15 cohort contained two sections taught by two different instructors who worked closely to align their course curriculum and teaching practices. SP16 was a single section taught by one of the instructors from SP15, using a nearly identical curriculum.

This study was approved by the CSUS Institutional Review Board (Protocol #14-15-127).

Curricular interventions

Pre-lecture assignments. Students were given guided homework assignments to be completed before each class meeting. Pre-lecture assignments were designed to prepare students for in-class activities by familiarizing them with important terms and concepts and often requiring them to organize or establish relationships between terms (e.g., using concept maps or comparing/contrasting). Most pre-lecture assignments included a metacognitive prompt/question asking students to describe their prior knowledge of a topic, evaluate their work, or report their confidence level related to concepts/topics covered by the pre-lecture assignment or describe how the topics on the assignment may be relevant to their lives (see Appendix 1). Pre-lecture assignments were collected randomly and scored for completion (5 points each in a course, with approximately 1,000 total points).

Collaborative group work. During class, students were expected to engage in higher-order thinking activities in groups of three to five individuals. Peer-discussion and individual and group quizzes provided students with opportunities to evaluate their own content mastery, as well as that of their peers.

Exam review assignments. Following each of the first three semester exams (out of four total), students completed an in-depth reflective assignment requiring them to 1) correct the answers to missed questions on their exam, 2) cite the source(s) they used to make their corrections, 3) describe why the answer they provided could not be correct, 4) attempt to diagnose the reasons for their mistakes, and 5) identify the study skills and behaviors that they employed to prepare for the exam. In addition, each Exam Review activity included several open-ended reflective questions asking students 1) to reflect on the material that gave them the most difficulty, 2) to evaluate their performance (e.g., "Do you believe your exam grade reflects your knowledge and the time you spent in preparation? Why or Why not?"), and 3) to further describe the study strategies and behaviors that worked for them or that needed to change for improved performance on the next exam (Appendix 2). This assignment is similar to cognitive wrappers or exam wrappers that have been employed across disciplines to help students improve their learning and self-regulatory skills (21, 22). See Andaya (15) for a full description of the Exam Review activity used in this study and its impact on students.

Data collection and analysis

We used two modes of assessment to evaluate student metacognitive development. In SP15 and SP16, students' self-evaluation skills were measured via a question at the end of the first and third exams asking them to postdict their exam score; this question was worth one point (on a 100-point exam). Second, at the beginning and end of the SP16 semester, students were offered an extra credit assignment (Pre-MAI and Post-MAI Survey and Reflective Essay; Appendices 3 and 4). The pre-MAI assignment asked students to watch two videos about metacognition and deep versus surface learning, complete the MAI, and write a personal reflection based on several prompts. The prompts asked students what they learned from the video and MAI, what they perceived to be their strengths and weaknesses regarding learning and studying, and what strategies they planned to use to be successful in BIO2. The pre-MAI exercise was assigned during the first and collected during the second week of the semester. The post-MAI required students re-take the MAI and respond to several prompts related to the effectiveness of the study strategies they employed and their thoughts regarding their metacognitive development over the course of the semester. The post-MAI exercise was offered during week 14 and collected during week 15 of the semester (Fig. 2).

Data analysis

For Exams 1 and 3, students were categorized as over- or under-postdictors based on a comparison of their actual and postdicted exam scores. Students whose actual exam scores were higher than their estimates (even by one point)

were classified as under-postdictors, and those whose actual exam scores were lower than their postdictions were classified as over-postdictors. Students who failed to make a postdiction or who postdicted correctly (actual score = postdiction score) were excluded from the analyses. The percentage of over- and under-postdictors was determined, and mean exam scores of these two groups were compared using independent sample *t*-test analyses. Postdiction accuracy was calculated using calibration scores, which measure how close a student's assessment of their performance is to their actual performance (6). Calibration scores were computed using the difference between students' actual and self-assessed (postdicted) exam scores, with the following equation from Miller and Geraci (8):

Students were further categorized as lower- and higher-performing by averaging all four exam scores for each student and selecting the median of this distribution as the cut-off point for the two groups. Independent sample *t*-test analyses were used to compare calibration scores of higher- and lower-performing students from Exam 1 to Exam 3.

The MAI essays were analyzed qualitatively by content analysis. Essay prompts often served as *a priori* themes and informed the coding strategy. For instance, on the pre-MAI essay assignment, "What might you be doing right?" and "What might you be doing wrong?" (while studying) were two of the prompts, and student responses to these questions were grouped into the categories "Student Strengths" and "Student Weaknesses." *In vivo* coding led to a list of activities that were assigned to these categories and then further sub-categorized by the MAI construct (Fig. 1). For the prompt, "What new strategies will you try this semester?", researchers developed a coding scheme that classified students into groups according to those who 1) described an action plan for studying, 2) discussed non-specific learning strategies, or 3) described no action plan for their learning. Inter-rater reliability (IRR), using percent agreement, averaged 91% and ranged from 85% to 100% for the individual codes. After approximately 30% of the pre-MAI reflections and all of the post-MAI essays had been coded independently, researchers convened to discuss and reconcile differences. The remainder of the pre-MAI essays were coded by one researcher, but checked for agreement with a second researcher.

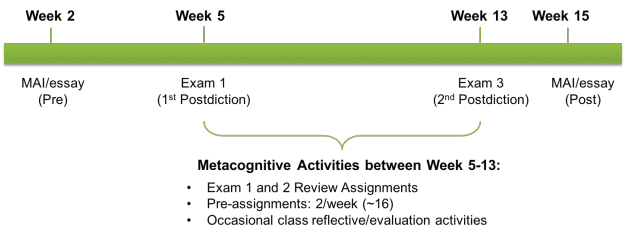

FIGURE 2. Timeline of metacognitive activities and assessments embedded in the curriculum over the course of semester.

Descriptive statistics were calculated to summarize pre- and post-MAI scores, and Pearson's correlations were used to determine whether there was a relationship between students' MAI scores, exam score postdictions and actual exam scores.

RESULTS

Postdiction direction and exam performance

To measure students' self-evaluation skills, we examined the relationship between actual and estimated (postdicted) exam scores following Exam 1, near the beginning of the semester, and Exam 3, after students had routinely practiced evaluating their work through the course curriculum. In both semesters, the percentage of students who over-postdicted their scores on Exam 1 (68% in SP15, 62% in SP16) exceeded the number who under-postdicted their scores (32% in SP15, 38% in SP16) (Fig. 3). However, after Exam 3, the number of under-postdictors increased (54% in SP15 and 77% in SP16), exceeding the number of over-postdictors (46% in SP15 and 23% in SP16). To determine whether over- or under-postdiction was related to exam performance, we compared mean exam scores for both groups. In both semesters and on both exams, we found that over-postdictors performed significantly lower on exams than under-postdictors (Table 1). Significant differences existed for SP15 over- and under-postdictors on Exam 1 ($t = 1.668$, df = 67, $p < 0.001$, 2-tailed) and Exam 3 ($t = 1.994$, df = 71, $p < 0.001$, 2-tailed), as well as SP16 Exam 1 ($t = 2.011$, df = 48, $p < 0.001$, 2-tailed) and Exam 3 ($t = 2.131$, df = 15, $p = 0.003$, 2-tailed).

Postdiction accuracy and student performance

Categorizing a student as an over- or under-postdictor provided information related to the direction, but not the accuracy, of their evaluation error. Thus, we used calibration scores to measure students' postdiction accuracies, with a

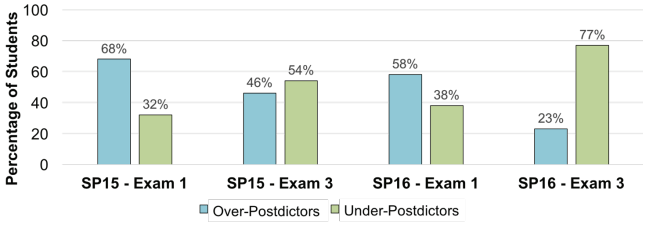

FIGURE 3. Frequency of over- and under-postdiction on Exams 1 and 3 for spring 2015 (SP15) and spring 2016 (SP16) cohorts. Blue bars represent the percentage of students who over-postdicted (estimated exam score was higher than actual exam score). Green bars represent the percentage of students who under-postdicted (estimated exam score was lower than actual exam score). Students who failed to make a prediction or who postdicted correctly (actual score = postdiction score) were excluded from the analysis. HP = high performing; LP = low performing.

TABLE 1.
Exam performance by postdiction group.

	Over-Postdictor Mean Score ± SD	Under-Postdictor Mean Score ± SD	*P* value
Exam 1 – SP15	69±11	79±10	<0.001
Exam 3 – SP15	70±13	87±7	<0.001
Exam 1 – SP16	64±15	85±10	<0.001
Exam 3 – SP16	72±11	85±9	<0.001

Mean scores plus standard deviations on Exams 1 and 3 for the Spring 2015 (SP15) and Spring 2016 (SP16) cohorts are shown for students who over- and under-postdicted their scores. Under-postdictors had significantly higher exam scores across all comparisons (p<0.001). Students who failed to make postdictions or who postdicted exam scores correctly were excluded from the analysis.

higher calibration score indicating greater accuracy, or less difference between estimated and actual exam score. In both semesters, there was no significant change in calibration scores from Exam 1 to Exam 3 for higher-performing students (SP15: t = 1.982, df = 109, p = 0.647, 2-tailed; SP16: t = 2.011, df = 48, p = 0.490, 2-tailed); lower-performing students, on the other hand, exhibited a significant increase in calibration scores from Exam 1 to Exam 3 (SP15: t = 1.987, df = 90, p = 0.019, 2-tailed; SP16: t = 2.018, df = 42, p = 0.007, 2-tailed) (Fig. 4). This resulted in an elimination of the significant difference in calibration scores observed between higher- and lower-performing students in both semesters from Exam 1 (SP15: t = 1.989, df = 82, p = 0.005, 2-tailed; SP16: t = 2.020, df = 41, p = 0.000, 2-tailed) to Exam 3 (SP15: t = 1.984, df = 99, p = 0.969, 2-tailed; SP16: t = 2.010, df = 49, p = 0.645, 2-tailed). Irrespective of performance group, 83.6% of SP15 students who improved their postdiction accuracy from Exam 1 to Exam 3 also demonstrated improvement in their exam scores, and 81.8% of students demonstrated this improvement in SP16.

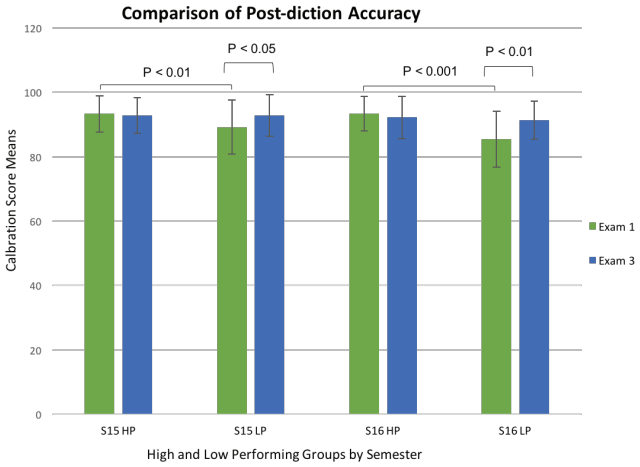

FIGURE 4. Postdiction accuracy by performance group on Exams 1 and 3 for spring 2015 (SP15) and spring 2016 (SP16) cohorts. Higher calibration score = more accurate. Students were categorized as LP and HP by averaging all foaur exam scores for each student and selecting the median of this distribution as the cut-off point for the two groups. Error bars represent standard deviation of the mean. Comparison bars with p values are shown for groups that differ significantly from one another in their calibration scores. LP = lower-performing; HP = higher-performing.

MAI survey and reflective essay

Analysis of the pre-MAI reflective essay indicated that 100% of the SP16 students who completed the pre-MAI assignment (n=37) found the activity to be beneficial. For example, in response to the prompt "Did you discover anything new?", one student replied, "The whole process was an eye opener and I can't wait to use the new learned skills and improve on my older ones for better understanding." Students reported on various strengths, with "Information Management Strategies" identified most often (39% of students); however, more students discussed their weaknesses related to studying (Table 2). The two most commonly reported weaknesses were "Focusing on memorizing, rather than understanding" (81%) and "Struggling with distractions, multi-tasking and procrastination" (55%). One student indicated several weaknesses included "…thinking about my goals prior to studying, which I do not do regularly. I also don't question my methods of learning much." Out of the 37 students, 68% described a clear, thoughtful action plan for studying that they intended to implement during the semester (Table 3). Another 30% of students discussed learning strategies in a general manner but failed to outline specific strategies or behaviors they would apply when studying. Only one student (3%) revealed no intention of implementing any new learning or study strategies during the semester.

Twenty students completed the post-MAI essay at the end of SP16, and 85% of this group indicated that they discovered and applied new study strategies. Of the 70% that included specific examples in their essays, the most commonly adopted methods/behaviors included studying and reading in advance of class, studying with other students or groups, and practicing problems, self-assessing, and checking comprehension. Eighteen students (90%) believed that their metacognitive skills improved to some degree over the course of the semester; however, over half of these students recognized that they still struggled with certain aspects of studying or metacognitive regulation. They attributed their

TABLE 2.
Summary of student responses to the pre-MAI reflection essay prompts, "What do you thing you are doing right when studying?" and "What might you be doing wrong when studying?"

	# of Students (% of Students)	MAI Construct
Student Strengths		
Take appropriate time to learn/pacing Paraphrase materials Make connections Use visual aids Break down learning into manageable parts	12 students (39%)	Information management strategies
Recognize weaknesses/misunderstandings Recognize understandings Use strategies for remembering information	9 students (29%)	Declarative knowledge
Focus on understanding Review material	9 students (29%)	Comprehension monitoring
Plan for studying Reduce distractions/control study environment	6 students (19%)	Planning
Select appropriate learning strategies	3 students (10%)	Procedural knowledge
Summarize material Critique work	2 students (6%)	Evaluation
Ask for help Address confusions early	2 students (6%)	Debugging
Student weaknesses		
Focus on memorizing, surface learning	17 students (55%)	Comprehension monitoring
Struggle with distractions, multi-tasking, procrastination	14 students (45%)	Planning
Don't ask for help, don't self-test, have difficulty changing habits	9 students (29%)	Debugging
Fail to make connections	7 students (23%)	Information management
Overpredict performance/overconfident	4 students (13%)	Evaluation
Fail to set goals	2 students (6%)	Planning

Specific strengths and weaknesses were identified through content analysis of student pre-MAI essays and categorized according to MAI construct.

challenges to lack of time and time management skills, stress in or outside of class, poor motivation, and difficulty with techniques related to studying or test taking. Although the prompt did not specifically ask students about changes in performance, 45% reported an improvement in grades over the course of the semester. See Appendices 5–7 in Supplemental Materials for additional post-MAI reflective essay data.

A subset of students (n=14) completed both the pre- and post-MAI surveys in SP16, allowing for the comparison of individual MAI constructs and total scores from the beginning and end of the semester. While we observed no significant difference between pre- and post-MAI means (Table 4), 64% of these students exhibited an increase in their total MAI score. We also found no evidence for correlations between pre- or post-MAI survey scores and students' actual exam grades or calibration scores; however, there was a significant positive relationship between students' exam grades and calibration scores for Exam 1, which supports the data showing that higher-performing students are also more accurate self-assessors (see Pearson correlation data in Appendices 8 and 9). Our results indicate that, while the MAI survey may not have been predictive of exam or metacognitive performance, students were using it in their essays to reflect on their strengths and weaknesses and devise plans for studying.

TABLE 3.
Summary of student responses to the pre-MAI reflection essay prompt, "What new strategies will you try this semester?"

Action Plan Code	# of Students (% of Students)	Sample Student Response
Described clear action plan	25 students (68%)	"Some new strategies I'm going to do this year is write notes while I read the text and also watch videos on the topic I'm learning to make it more interesting and see it in a different light than the textbook..."
Discussed non-specific learning strategies	11 students (30%)	"This semester I am going to try to manage my time more wisely and develop better learning habits."
No indication of implementing action plan	1 student (3%)	N/A

Responses were coded based on the extent to which students described an action plan for their studying, and example responses are provided for two of the codes. N/A = not applicable.

TABLE 4.
Comparison of pre-MAI and post-MAI survey scores.

MAI Constructs (Max Points)	Pre-MAI Mean Score ± SD	Post-MAI Mean Score ± SD
Declarative knowledge (8)	5.9±0.4	6.5±0.4
Procedural knowledge (4)	3.1±0.2	3.4±0.2
Conditional knowledge (5)	3.9±0.3	4.4±0.3
Comprehension monitoring (7)	4.4±0.3	4.9±0.5
Evaluation (6)	3.7±0.4	3.9±0.5
Planning (7)	4.6±0.4	4.6±0.5
Debugging strategies (5)	4.5±0.2	4.4±0.2
Information management strategies (10)	8.0±0.4	7.4±0.4
MAI total survey (52)	38.1±5.2	39.4±8.2

Mean scores and standard deviations are shown for each of the 8 MAI constructs and for the total survey. Numbers in parentheses indicate maximum possible score. There was no significant difference in pre- and post-MAI total means (t = 2.074, df = 22, p = 0.627, 2-tailed).

DISCUSSION

Strong metacognitive skills have been linked to greater learning and higher performance in students of various developmental stages engaged in different academic disciplines; yet few studies have focused on metacognitive awareness or development in college students pursuing science majors. Our study sought to determine whether measurable gains in metacognition could be detected over the course of a single semester in an introductory biology course designed with the goal of promoting metacognitive awareness and skills development. We used exam score postdictions to directly measure students' self-evaluation skills, the MAI survey to gauge metacognitive awareness, and reflective essays to gain a deeper understanding of students' perceptions related to their metacognitive activities and study habits. The pre-MAI assignment served as a measurement tool for collecting baseline data but also provided students with an opportunity to learn about metacognition and engage in several related activities, such as self-evaluation and planning.

Our study results support our hypothesis that students who more accurately estimated their exam scores, a measure of self-evaluation skills, would also score higher on the exams. We observed this pattern in both semesters, as higher-performing students had significantly higher calibration scores, and thus higher postdiction accuracy, than their lower-performing counterparts. We also found that students who over-estimated their exam performance in both semesters had significantly lower exam scores than students who under-estimated, which is consistent with the findings of previous studies (6, 8, 22). Aligned with the goals of the curriculum, we observed a significant decrease in over-postdictors from Exam 1 to Exam 3 in both semesters, and also found that lower-performing students exhibited a significant increase in postdiction accuracy. This suggests that lower-performing students are improving in self-evaluation skills over the course of the semester, and this may be a factor in their improved performance. While student confidence was not measured in this study, Ehrlinger and

Shain (2) found that overconfidence could be a problem for students if it leads them to terminate studying early based on their belief that they have mastered the content, when in fact they have not. The overconfidence may persist in students with less developed metacognitive skills, as "students who do not monitor [their learning] are likely to maintain confidence in their skill level even when it is not merited" (2). Furthermore, Hacker, Bol, and Bahbahani (6) found that students who performed higher academically often attributed their shortcomings to under-confidence, perhaps as a self-defense mechanism against disappointment. Our study results do not provide an explanation for the quality of students' self-evaluation or how they approached their postdictions; however, the themes above were expressed by students during both office hours with the course instructors and on responses to Exam Review reflective questions. Specifically, lower-performing students often expressed surprise regarding their grades and indicated that they felt prepared after minimal studying. Under-confidence, or an intentional effort to under-postdict to avoid disappointment, was also voiced by students who performed well on exams and had high postdiction accuracy. These are anecdotal observations, and additional research is needed to confirm a relationship between confidence, metacognition and performance in our student population.

Our second hypothesis that higher MAI scores would correspond with higher exam scores and postdiction accuracy was not supported by our data, as Pearson correlations revealed no significant association between MAI scores and actual exam scores or postdicted exam scores. This suggests a disconnect between students' perceived and actual metacognitive skills; it could also indicate that the MAI, which is a Likert-like self-report survey, is limited in its ability to predict metacognitive aptitude. As an alternative method for examining changes that may have occurred during the semester that could not be discerned by the MAI, we used student reflections. Qualitative analyses of pre-MAI reflections revealed that most students found the MAI assignment beneficial, could identify strengths and weaknesses in their study strategies, and developed action plans to improve their learning. Post-MAI analyses revealed that most students employed new study strategies and about half believed that their metacognitive and study skills improved along with their exam performance. The majority of the students who completed the MAI assignments also communicated an awareness of factors that were likely influencing their academic performance.

Our observations are promising and suggest that curricular activities designed to promote metacognition do indeed help students improve their self-evaluation skills and may preferentially help lower-performing students. This study lacks a true control group, as instructors and researchers felt it unethical to withdraw the curricular components serving as the "interventions," the benefits of metacognitive and active learning activities being well-documented. However, the pre-post design provides evidence that individual students and grouped cohorts experienced gains, potentially due to the regular opportunities they had to reflect on their learning and study habits as a means of fostering metacognition. Instruction could benefit from further research concerning the types and frequency of curricular activities that best promote the development of self-evaluation and other metacognitive skills in science students.

SUPPLEMENTAL MATERIALS

Appendix 1: Pre-assignment metacognitive questions
Appendix 2: Exam Review assignment metacognitive questions
Appendix 3: Pre-MAI assignment
Appendix 4: Post-MAI assignment
Appendix 5: Summary of student responses to the post-MAI reflection essay prompts
Appendix 6: List of new study strategies/behaviors
Appendix 7: Reasons students provided in response to the post-MAI reflection essay prompt
Appendix 8: Correlation between pre-MAI scores, exam grades and calibration scores
Appendix 9: Correlation between post-MAI Scores, exam grades and calibration scores

ACKNOWLEDGEMENTS

The authors declare that there are no conflicts of interest.

REFERENCES

1. Vadhan V, Stander P. 1994. Metacognitive ability and test performance among college students. J Psychol 128:307–309.
2. Ehrlinger J, Shain EA. 2014. How accuracy in students' self perceptions relates to success in learning. *In* Benassi VA, Overson CE, Hakala CM (ed), Applying science of learning in education: infusing psychological science into the curriculum. Society for the Teaching of Psychology. [Online.] http://teachpsych.org/ebooks/asle2014/index.php.
3. Pintrich PR. 2010. The role of metacognitive knowledge in learning, teaching and assessing. Theory Pract 41:219–225.
4. Schraw G, Dennison RS, 1994. Assessing metacognitive awareness. Contemp Educ Psychol 19:460–475.
5. Stanton JD, Neider XN, Gallegos IJ, Clark NC. 2015. Differences in metacognitive regulation in introductory biology students: when prompts are not enough. CBE Life Sci Educ 14:1–12.
6. Hacker DJ, Bol L, Bahbahani K. 2008. Explaining calibration accuracy in classroom contexts: the effects of incentives, reflection, and explanatory style. Metacog Learn 3:101–121.
7. Miller TM, Geraci L. 2011. Unskilled but aware: reinterpreting overconfidence in low-performing students. J Exp Psychol Learn Mem Cog 37:502–506.

8. Miller TM, Geraci L. 2011. Training metacognition in the classroom: the influence of incentives and feedback on exam predictions. Metacog Learn 6:303–314.
9. Dunlosky J, Rawson K. 2012. Overconfidence produces underachievement: inaccurate self evaluations undermine students' learning and retention. Learning Instr 22:271–280.
10. VanderStoep SW, Pintrich PR, Fagerlin A. 1996. Disciplinary differences in self-regulated learning in college students. Contemp Educ Psychol 21:345–362.
11. Miller D. 2015. Learning how students learn: an exploration of self-regulation strategies in a two-year college general chemistry course. J Coll Sci Teach 44:11–16.
12. Dunning D. 2011. The Dunning-Kruger effect: on being ignorant of one's own ignorance. Adv Exp Soc Psychol 44:247–296.
13. Tanner KD. 2012. Promoting student metacognition. CBE Life Sci Educ 11:113–120.
14. Nilson LB. 2013. Creating self-regulated learners: strategies to strengthen students' self-awareness and learning skills. Stylus Publishing, LLC, Sterling, VA.
15. Andaya G, Hrabak V, Parks S, Diaz RE, McDonald KK. 2017. Examining the effectiveness of a postexam review activity to promote self-regulation in introductory biology students. J Coll Sci Teach 46:84–92.
16. Deslauriers L, Harris SE, Lane E, Wieman CE. 2012. Transforming the lowest-performing students: an intervention that worked. J Coll Sci Teach 41:80–88.
17. Welsh AJ. 2012. Exploring undergraduates' perceptions of the use of active learning techniques in science lectures. J Coll Sci Teach 42:80–87.
18. Hill KM, Bronzel VS, Heiberger GA. 2014. Examining the delivery modes of metacognitive awareness and active reading lessons in a college nonmajors introductory biology course. J Microbiol Biol Educ 15:5–12.
19. Doyle BP. 2013. Metacognitive awareness: impact of a metacognitive intervention in a pre-nursing course. PhD thesis. Louisiana State University, Baton Rouge, LA.
20. Siegesmund A. 2016. Increasing student metacognition and learning through classroom-based learning communities and self-assessment. J Microbiol Biol Educ 17:204–214.
21. Lovett MC. 2013. Make exams worth more than grades: using exam wrappers to promote metacognition in using reflection and metacognition to improve student learning. Stylus Publishing, LLC, Sterling, VA
22. Gezer-Templeton PG, Mayhew EJ, Korte DS, Schmidt SJ. 2017. Use of exam wrappers to enhance students' metacognitive skills in a large introductory food science and human nutrition course. J Food Sci Educ 16:28–36.