Teaching Students How to Check Their Work While Solving Problems in Genetics

By Lisa McDonnell and Martha Mullally

Although problem solving is a highly valued skill, the processes involved are rarely taught in our undergraduate biology classes. An essential component skill of monitoring and reflection during problem solving is work checking, a process used by experts while solving problems to determine if their solution is achieving the goal. The results of work checking may reveal errors or inconsistencies, indicating a need for iteration. Using think-aloud interviews, we identified that most students in our undergraduate genetics course did not engage in checking while solving problems. In response, we made changes to the course curriculum to include explicit teaching of work checking, practice and feedback on work checking, and allotting a small number of course points for demonstration of checking on quizzes and exams. Analysis of students' written answers and think-aloud interviews revealed that significantly more students engaged in checking after our curriculum change, even when not prompted. Results of this study highlight the value of explicitly teaching context-specific, problem-solving processes in our science courses.

Problem solving is a valued skill, and improving this skill is a desired outcome for undergraduates pursuing STEM (science, technology, engineering, and mathematics) degrees (Brewer & Smith, 2011; National Academy of Science, 2011). To transfer knowledge beyond a single postsecondary STEM course, students need to become strong critical thinkers and problem solvers. Further, with a need to move toward more active classrooms in STEM undergraduate courses (Freeman et al., 2014), students will use critical thinking and problem-solving skills more often, and hence it is important to incorporate support for the development of these skills in our undergraduate STEM teaching. Successful problem solving requires conceptual and factual knowledge (declarative knowledge); problem-solving skills, also referred to as strategic knowledge (de Jong & Ferguson-Hessler, 1996); and metacognitive skills (Anderson & Nashon, 2007; Zohar & Dori, 2012). Although course curricula typically focus on declarative knowledge, of equal importance is how to help students develop the necessary strategic and metacognitive knowledge and skills required to solve problems.

The problem-solving processes used by experts may vary depending on the context and type of problem. There are also many different processes and skills needed to successfully solve problems (Adams & Wieman, 2015). However, the problem-solving framework described by PISA (Organisation for Economic Co-operation and Development [OECD], 2012) and others (e.g., Carlson & Bloom, 2005; Larkin, McDermott, Simon, & Simon, 1980) identifies four broad components of problem solving common to most disciplines: (a) exploring and understanding the problem, (b) representing and formulating, (c) planning and executing, and (d) monitoring and reflecting.

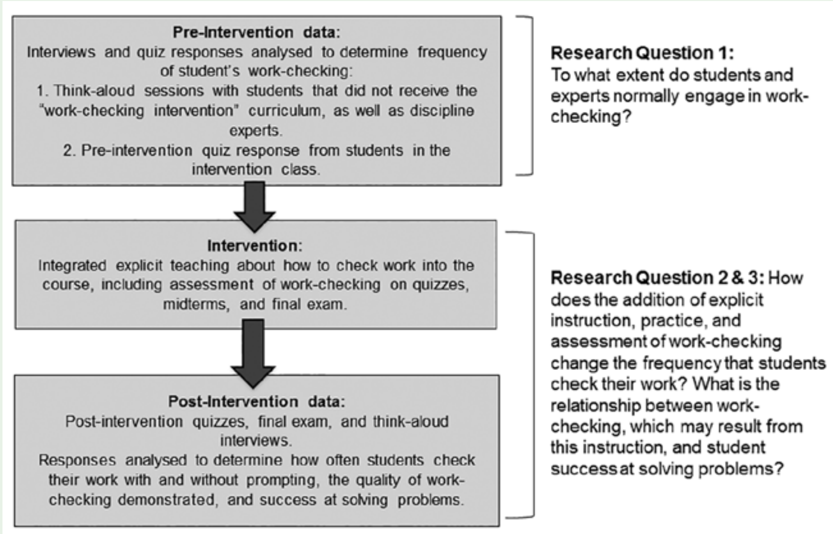
An aspect of monitoring and reflecting can involve the solver using various methods to check or verify their solution against stated or intuited criteria to determine if the solution is reasonable. The criteria or goals may be outlined in the question, as well as determined by the constraints of the concepts involved. For example, if asked about determining the speed of a car given the distance travelled over a period of time, solvers may check that their answer falls within a certain order of magnitude that makes sense based on their past experience, and the information provided in the question. If the results of checking indicate errors or inconsistencies, then one should cycle back to earlier processes (e.g., a–c in previous paragraph). Checking work has been described as an important reflective component of successful problem solving in various contexts, such as mathematics (Carlson & Bloom, 2005; Schoenfeld, 1985, 1987), physics (Holmes, 2015), writing (Mayer, 1998), and genetics (Smith & Good, 1984).

Despite the fact that checking is necessary for successful problem solving, it is likely that many undergraduate STEM students do not often engage in these processes of monitoring and reflecting. Others have documented that students normally do not monitor their work (Holmes, 2015; Schoenfeld, 1985, 1987; Smith & Good, 1984). It may be that many students engage in low-level or naïve checking whereby the criteria against which they monitor their progress is superficial, such as providing an answer to a question or listing a genotype as requested but not checking if the genotype follows the rules of inheritance relevant to the question.

Pedagogical interventions to improve students' problem-solving skills have yielded mixed results. For example, explicit instruction of problem-solving strategies has been shown to improve at least some aspects of students' use of various strategies, such as organizing and representing information in a problem (DiLisi et al., 2006; Huffman, 1997). However, DiLisi et al. (2006) found that students had a low adoption rate of a monitoring process they called "dimensional analysis" and postulated that students did not appreciate the value in engaging in such monitoring. In contrast, educational supports in an undergraduate physics lab that encouraged work checking and decisions based on that checking increased the proportion of students who monitored and used the results of the monitoring to improve their work (Holmes, 2015).

FIGURE 1

Overview of the study design.



Our research focused on improving students' ability to engage in productive monitoring, specifically through checking their work. For the remainder of this article, "work checking" is how we describe the form of monitoring and reflection targeted in this research. We examined the following research questions:

1. To what extent do students normally engage in work checking and how does this compare to experts?
2. How does the addition of explicit instruction, practice, and assessment of work checking affect the frequency that students check their work?
3. What is the relationship between the work checking, which may result from this instruction, and student success at solving problems?

Methods

Course and study design

This study took place in a second-year genetics course at the University of British Columbia, a large, selective, research university in Canada. The course covers multiple topics, from gene and chromosome structure through to introductory cancer genetics. Solving genetic problems (genetic analysis) composes 50% of the course curriculum. The course is required for all biology majors. There are between 60 and 200 students in each class.

Figure 1 provides a visual overview of the study design. To answer the first research question, we performed think-aloud interviews with students who had not received explicit instruction on problem-solving strategies (control group, $n = 15$ from a class of 160 students in 2012), and with discipline experts (n = 4, three course instructors and a graduate student). Experts were interviewed so we could compare problem-solving processes of students to those of experts (see next section for methods on interview protocols and analysis). Guided by the results of these interviews, we implemented curricular changes in the summer 2013 term to address research questions 2 and 3. The curricular change, "work-checking intervention," involved adding explicit instruction, practice, feedback, and assessment of work checking. Students who received this intervention were considered the treatment group (n = 60). Analysis of students' work checking on quizzes, exams, and postintervention think-aloud interviews was used to assess the impact of the intervention.

Think-aloud interviews

Think-aloud interviews were conducted following recommendations of Ericsson and Simon (1998) using a method that aims to capture subject thinking while attempting not to alter their normal behaviors. Interviews were audio recorded. At the interview, the subject (expert or student) was presented with a set of exam-style genetics problems and asked to solve the problems as they would normally and verbalize their thought process as much as possible while working (think aloud). We were careful to not interrupt their thinking with any probing questions, only occasionally reminding them to continue to think aloud. Interview notes and audio recordings were transcribed. Both authors independently analyzed a subset of transcripts with the aim of identifying common steps or approaches used by the interviewees while they were solving the problems. The problem-solving processes identified by each author were compared and discrepancies were discussed and resolved by reviewing the original transcripts. From the first round of analysis we developed a rubric allowing us to classify observed work and behaviors into five problem solving steps, similar to those described by others (e.g., OECD, 2013): (a) organize information in the problem, (b) hypothesis formation, (c) solving (applying hypothesis), (d) work checking, and (e) considering alternative solutions. For the purposes of this research, each interview was then scored for whether work checking was observed (1 for some work checking, 0 for no checking). An example of student processes captured during the think-aloud interview can be found in the Supplemental Material S1 (available at http://www.nsta.org/college/connections.aspx).

Intervention and data analysis

Previous iterations of the genetics course spent a significant fraction of in-class time (10%–35% per week), and approximately 90% of tutorial/recitation time, solving genetics problems, but students were not taught how to use work checking while solving problems, and there was no structured practice, feedback, or assessment of work checking. In the summer 2013 term, all

TABLE 1

Scoring of students' written work on the problem-solving pretest and posttest.

Item scored	0 points	1 point	2 points
Evidence of work checking	No checking obvious	Some form of work checking obvious in written work, e.g., an arrow linking their solution to a piece of information given in the question, or a comment made indicating how checking revealed an inconsistency	2 points not used
Uses checking to revise original hypothesis/solution	No revision detected	A second hypothesis/solution proposed	2 points not used
Quality of work checking (final exam only)	No checking evident	*Superficial:* compared work with a subset of the criteria given	*Thorough*: compared work with all of the criteria given
Correctness	No correct elements in final solution	Mostly correct, but missing a piece of information, or one element of answer is incorrect	Completely correct, satisfying all the criteria given in the question

of the course content was the same as the control course, with the exception of the addition of the work-checking intervention, which was captured with the addition of the following course learning outcome: *Show how you check your hypotheses/explanations to verify they support data you are given.*

The intervention spanned the entire course and began with a 45-minute in-class lesson on Day 4 of the course. To start the lesson, students were given a challenging genetics problem to solve and asked to try work checking with the following prompt: "Checking your work: What work have you done to check that your hypothesis explains the data provided? Highlight that work above, or do the work here." After students had a sufficient amount of time to work on the problem, the instructor provided feedback to the class by modeling problem-solving processes. The instructor solved the problem in real time, verbalizing her solving process (hypothesis formation, stating any assumptions she's making, describing how and why she was checking her work, describing and showing how to deal with the results of work checking) and showing her work using real-time writing.

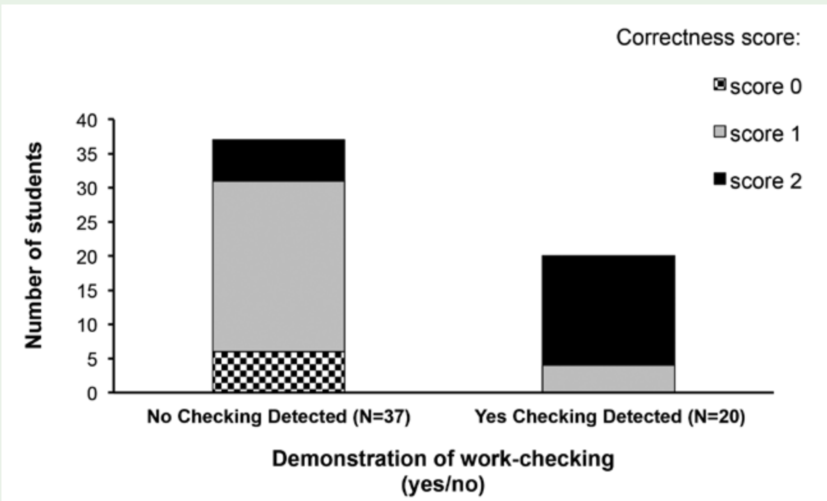
Throughout the term students were provided with opportunities, both in and outside of class, to practice work checking (3x per week), and work checking was explicitly assessed on quizzes (1x per week), on each of two midterms, and the final exam. Questions to assess these behaviors included explicit prompts, where students were asked to write their hypothesis and show how they checked their work. See Supplemental Material S2 (available at http://www.nsta.org/college/connections.aspx) for an example of a test question containing an assessment of checking. Marks for demonstrating work checking were worth a total of 5 course points out of 100 (1.5 on quizzes, 1.7 on a midterm, 1.8 on the final exam).

A preintervention quiz was administered before the work-checking lesson to assess how frequently students demonstrate forms of work checking in their problem-solving work. On the quiz, students were asked to solve a single problem with the explicit instructions "to show all of your work, including any wrong attempts you made, and how you went about deciding if your answer was correct." Two days after the work-checking lesson, students were given a postquiz (a new question), which included explicit instructions to show how they checked their work (Supplemental Material S3, available at http://www.nsta.org/college/connections.aspx). Student answers to the pre- and postquiz problems were scored on three dimensions, described in Table 1. At the end of the term, final exam work checking was also scored, and an additional quality dimension was scored (Table 1).

Because of the categorical nature of the data, Fisher's exact tests were used for the following comparisons: (a) the number of students who demonstrated work checking and achieved a high correctness score on

FIGURE 2

Frequency of work checking demonstrated on preintervention quiz and distribution of correctness scores. Before the work-checking intervention began, the majority of students did not show work checking in their answers on a prequiz question (compared 37 students with no checking to 20 students with checking). Students who did show work checking were more likely to achieve a higher correctness score. The number of students with a score of 2 was significantly higher in the work-checking group (Fisher's exact, $p <$ .001). These results answer research question 1.



the pretest compared with those that did not demonstrate work checking and achieved a high correctness score on the pretest; (b) the number of students demonstrating work checking between the pre- and posttest; (c) the number of students demonstrating work checking on the pretest compared with the final exam; and (d) the number of students from the control and intervention groups demonstrating work checking during think-aloud interviews. A Kruskal-Wallis test, followed by a Dunn's post hoc test, was used to determine if there was a significant difference between a score on the exam question and the type of work checking demonstrated on that question (none, superficial, and thorough).

Results

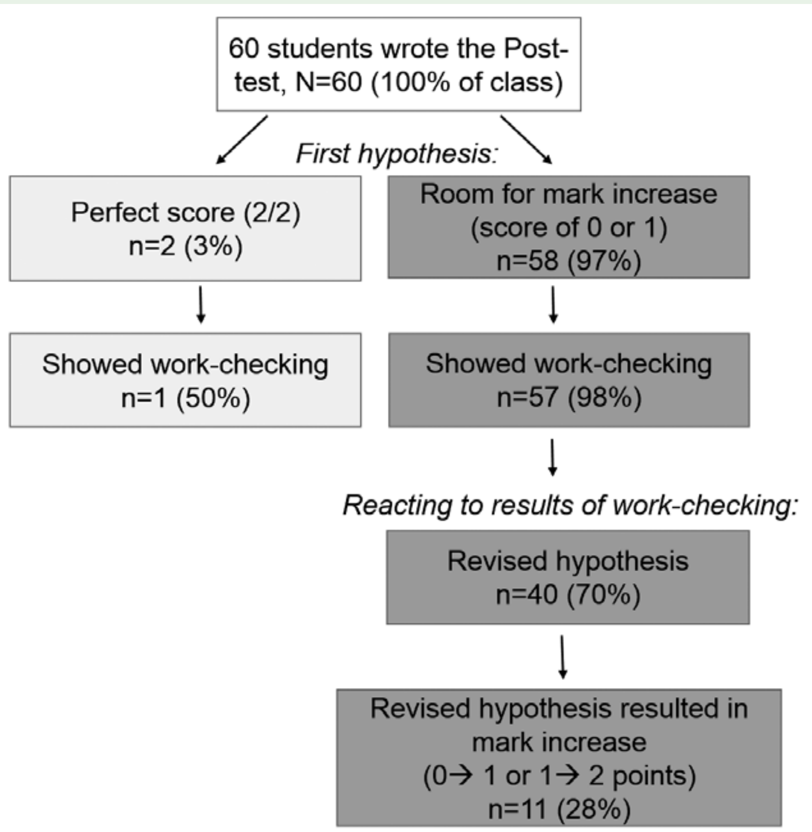
Low frequency of work checking while solving problems

Think-aloud interviews of students ($n = 15$) and experts (instructors, $n = 3$) from the control course (no work-checking intervention) revealed the problem-solving processes students commonly use and how those differed from experts. Students often omitted the work-checking step (only 40% of interviewed students demonstrated work checking on a given problem). All of the interviewed experts consistently used work checking as a part of their problem-solving process.

Analysis of students' written work on the preintervention quiz revealed a similar trend that was measured by the think-aloud interviews: Most students did not demonstrate work checking. Thirty-five percent of students' prequiz responses included some form of work checking (Figure 2). A correctness score was also assigned (0, 1, or 2 for correctness); students who demonstrated work checking were more likely to achieve a higher score (Figure 2), suggesting to us that supporting students to learn how to check their work may improve their overall success on course problems.

FIGURE 3

Frequency of work checking and using checking to revise work on the postintervention quiz. Thirty-five percent of students demonstrated work checking on the prequiz, before the work-checking in-class lesson. A significantly larger proportion of the class (98%) checked their work and realized errors (70%) on the postquiz. This answers research question 2.



Teaching students to engage in work checking

On the postintervention quiz, 98% of students demonstrated some form of work checking (Figure 3). The number of students checking their work is significantly greater on the post-test, after the problem-solving lesson, compared with the pretest ($p < .001$, Fisher's exact test). This indicated that the lesson, brief amount of practice, and points awarded for demonstrating work checking were sufficient to increase the number of students who checked their work in response to the prompt. Of the students who had a first hypothesis worth 0 points out of a possible 2 points ($n = 2$) or 1 point ($n = 56$), 57 of them checked their work, and 40 of the 57 revised their initial hypothesis to generate a second, alternative hypothesis. This indicates that many of the students reflected and used the checking to monitor their work. An example of student work—revealing the checking, reflecting on the checking, and revising of their hypothesis—can be found in Supplemental Material S3 (available at http://www.nsta.org/college/connections.aspx).

However, this checking resulted in limited improvement on performance. Only 11 of these 40 students generated a second hypothesis that resulted in a marked increase, compared with the correctness score achieved with their original hypothesis (0→1 point, or 1→2 points). This indicates the challenging nature of this particular problem and suggests that many students may lack the conceptual knowledge required to generate an alternative hypothesis, a known difficulty associated with problem solving in a context where one is not an expert (Adams & Wieman, 2015).

Persistence and quality of work checking

Throughout the course, grade points were offered for demonstrating work checking. To assess whether students would continue to use work checking when there were no explicit grade incentives, we analyzed responses to an exam question where no work-checking prompt was present and no points were offered for showing work checking. Over 60% of the class showed some form of work checking (Table 2), significantly higher than the number of students who showed work checking on the preintervention quiz (Fisher's exact, $p < .01$). A closer look at students' work checking on this exam question revealed variation in the quality of checking. We coded quality of work checking by looking at all the criteria in the problem that could be checked against. If a student only checked against a subset of those criteria, we classified it as superficial work checking, but comparing to all criteria we classified as thorough work checking. Approximately 65% of the students who checked their work demonstrated superficial checking (Table 2). As expected, the mean correctness score was significantly higher for the group of students who showed thorough checking compared with superficial checking or no checking (Kruskal-Wallis followed by Dunn's post hoc test, $p < .001$, Table 2). Only 3 of the 14 students who showed thorough checking on the exam question also showed work checking on the prequiz, suggesting that many of the students who showed thorough work checking may have attained these skills from the work-checking intervention and practice in the course.

TABLE 2

Number of students demonstrating thorough and superficial work checking on a postintervention exam question in the absence of a prompt or point incentive.

Quality of work checking on the exam question	Number of students	Mean number of points on exam question, max = 6 (*SD*)
None	24	3.6 (1.3)
Superficial	26	3.8 (1.1)
Thorough	14	5.7 (0.6)*

*Mean score of the thorough checking group is significantly higher than the no-checking (none) and superficial group mean scores (Kruskal-Wallis followed by Dunn's post hoc test, $p < .001$).

We also conducted a small number ($n = 10$) of think-aloud interviews with students from the intervention class on the last day of class, before the final exam. Compared with the control group, more students demonstrated work checking spontaneously in the absence of prompts (Figure 4). These interview results reinforce the trend observed on the exam from the intervention class—that many students adopted work checking as part of their problem-solving routines, even under no-rewards and low-stakes situations.

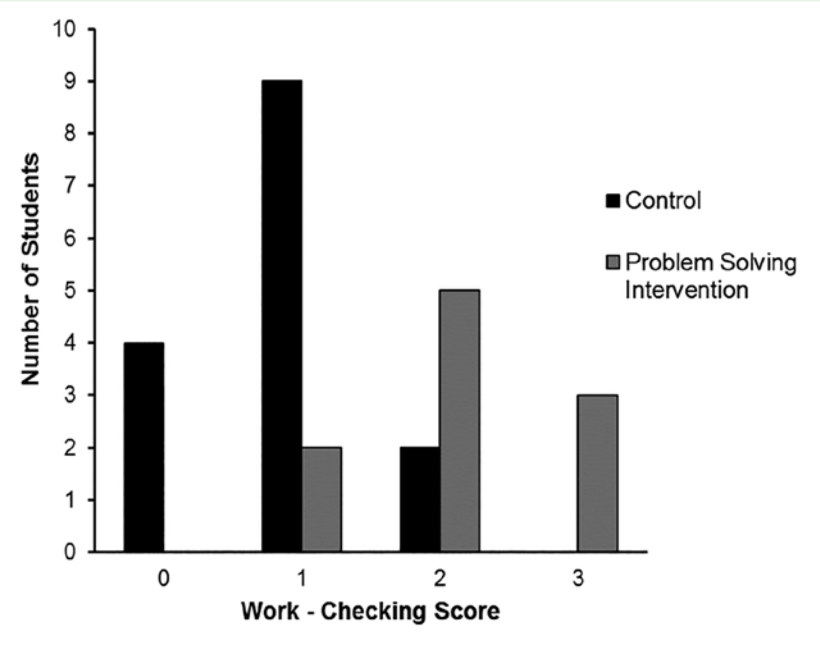
Conclusions

To improve problem-solving skills, we must find ways to effectively teach and assess these skills in our courses (DeHaan, 2009; Hoskinson, Caballero, & Knight, 2013; Maskiewicz, Griscom, & Welch, 2012). This research demonstrates that the majority of our students lack the skills to engage in context-specific problem solving and expert-like metacognitive processes such as monitoring, in the form of work checking against stated or conceptually constrained criteria and goals. For many students, it may be that they are unaware of how to engage in monitoring and reflection, beyond very novice checking such as "did I answer the question?"

Our results show that the context-specific, course-based intervention approach resulted in a large proportion

FIGURE 4

Distribution of work-checking scores of think-aloud interviews from control and intervention groups. Students could receive a maximum score of 3 for work checking (demonstrating checking on all three questions coded from the think-aloud interview). Students from the class receiving the work-checking intervention demonstrated work checking more often during the interview compared with students from the control group (no intervention). Number of students demonstrating more checking (2 or 3 points) is significantly higher in the treatment group than control group (Fisher's exact test, $p < .001$).



of the students in our undergraduate genetics class both learning and spontaneously engaging in reflection of their work and iteration (checking and revising their solution). This indicates that the intervention successfully provided metacognitive tools for students to use while solving problems. However, many students engaged only in superficial work checking. Superficial checking is superior to the very low-level checking ("did I answer the question?"), because it still involves checking against some of the criteria stated in the question. However, the quality of superficial checking is such that errors and inconsistencies may not be revealed. This result suggests that the intervention was missing repeated exposure to modeling of thorough work checking and using it to improve answers, which may explain why many students engaged in superficial checking on the exam question. It could also be that time pressure during an exam situation limits the investment students will voluntarily make into demonstrating work checking. Holmes (2015) observed that having students repeatedly reflect on their work and make decisions that are based on that reflection resulted in more thorough monitoring. This continued after scaffolding was removed, including in a subsequent course. This suggests a worthwhile addition to our intervention. We predict that such an addition will increase the frequency of students who engage in these key problem-solving behaviors and may also increase transfer of these skills to novel situations (Mayer & Wittrock, 1996; Ogilvie, 2009; Salomon & Perkins, 1989). ■

Acknowledgments

We enthusiastically acknowledge the following people who made this research possible: Jennifer Klenz for delivering the problem-solving lesson to her class and assisting with the development of the work-checking intervention; Jennifer Klenz, Craig Berezowsky, Shannon May-McNally, and Pam Kalas for participating in think-aloud interviews; student volunteers for contributing their time to participate in interviews; Trish Schulte, Sarah Gilbert, and Megan Barker for helpful conversations about the research project; and Carl Wieman, Laura Weir, and Jane Maxwell for critical feedback on the manuscript.

References

Adams, W. K., & Wieman, C. E. (2015). Analyzing the many skills involved in solving complex physics problems. *American Journal of Physics*, *83*, 459–467.

Anderson, D., & Nashon, S. (2007). Predators of knowledge construction: Interpreting students' metacognition in an amusement park physics program. *Science Education*, *91*, 298–320.

Brewer, C. A., & Smith, D. (Eds.).

(2011). *Vision and change in undergraduate biology education*: *A call to action.* Washington, DC: American Association for the Advancement of Science. Retrieved from http://visionandchange.org/files/2013/11/aaas-VISchange-web1113.pdf

Carlson, M. P., & Bloom, I. (2005). The cyclic nature of problem solving: An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics*, *58*, 45–75.

DeHaan, R. L. (2009). Teaching creativity and inventive problem solving in science. *CBE—Life Sciences Education, 8,* 172–181.

de Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist*, *31*, 105–113.

DiLisi, G. A., Eulberg, J. E., Lanese, J. F., & Padovan, P. (2006). Establishing problem-solving habits in introductory science courses. *Journal of College Science Teaching*, *35*(5), 42–47.

Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, & Activity, 5,* 178–186.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, USA*, *111*, 8410–8415.

Holmes, N. G. (2015). *Structured quantitative inquiry labs: Developing critical thinking in the introductory physics laboratory*. Unpublished doctoral dissertation, University of British Columbia. Retrieved from http://circle.ubc.ca/handle/2429/51363

Hoskinson, A.-M., Caballero, M. D., & Knight, J. K. (2013). How can we improve problem solving in undergraduate biology? Applying lessons from 30 years of physics education research. *CBE—Life Sciences Education*, *12*, 153–161.

Huffman, D. (1997). Effect of explicit problem solving instruction on high school students' problem-solving performance and conceptual understanding of physics. *Journal of Research in Science Teaching*, *34*, 551–570.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*(4450), 1335–1342.

Maskiewicz, A. C., Griscom, H. P., & Welch, N. T. (2012). Using targeted active-learning exercises and diagnostic question clusters to improve students' understanding of carbon cycling in ecosystems. *CBE—Life Sciences Education*, *11*, 58–67.

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science*, *26*(1-2), 49–63.

Mayer, R. E., & Wittrock, M. C. (1996). Problem-solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 47–62). New York, NY: Macmillan.

National Academy of Science. (2011). *The current status and future direction of biology education research*. Washington, DC: Author.

Ogilvie, C. A. (2009). Changes in students' problem-solving strategies in a course that includes context-rich, multifaceted problems. *Physical Review Special Topics—Physics Education Research*, *5*(2), 020102.

Organisation for Economic Co-operation and Development (OECD). (2013). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy.* Paris, France: OECD Publishing. Available at http://www.oecd.org/pisa/pisaproducts/PISA%202012%20framework%20e-book_final.pdf

Salomon, G., & Perkins, D. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational Psychologist*, *24*, 113–142.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, FL: Academic Press.

Schoenfeld, A. H. (1987). What's all the fuss about metacognition? In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 189–215). Hillsdale, NJ: Erlbaum.

Smith, M. U., & Good, R. (1984). Problem solving and classical genetics: Successful versus unsuccessful performance. *Journal of Research in Science Teaching*, *21*, 895–912.

Zohar, A., & Dori, Y. J. (Eds.). (2012). *Metacognition in science education: Trends in current research*. New York, NY: Springer Science.

Lisa McDonnell (lmcdonnell@ucsd.edu) *was an assistant teaching professor and* ***Martha Mullally*** *was a biology instructor, both in the Zoology Department at the University of British Columbia in Vancouver, BC, Canada, at the time this paper was written.* ***Lisa McDonnell*** *is now in the Division of Biological Science at the University of California San Diego in La Jolla and* ***Martha Mullally*** *is now the director (acting) of the Dalhousie Integrated Science Program at Dalhousie University in Halifax, Nova Scotia, Canada.*