Tech., Inst., Cognition and Learning, Vol. 8, pp. 255–271
Reprints available directly from the publisher
Photocopying permitted by license only

Improving Students' Problem Solving in a Web-based Chemistry Simulation Through Embedded Metacognitive Messages

CAROLE R. BEAL* AND RONALD H. STEVENS

University of Arizona Cognitive Science Program,
P.O. Box 210025, Tucson AZ 85721 USA
University of California Los Angeles IMMEX Laboratory,
5601 West Slauson Ave. #272, Los Angeles CA 90230 USA

The goal of the study was to improve high school students' problem solving performance as they worked with IMMEX, an online chemistry problem solving simulation. The intervention included brief text messages added to the simulation to encourage students to reflect on their problem solving strategies and to adopt more effective problem solving behaviors. Students who worked with the message-enhanced version were more likely to solve the problems correctly, and to use more effective problem solving strategies than students who worked with the original version. Benefits of the messages were observed for students with relatively poor problem solving skills, and for students who used exhaustive strategies.

Keywords: science education, metacognition, technology-based learning, problem solving, web-based chemistry simulation

INTRODUCTION

Problem solving is now recognized as a central component of science proficiency. There is increasing consideration in the science education community of the distinction between learning *about* science, and learning to *do* science (Chinn & Malhotra, 2002; Duschl, 2008; National Research Council [NRC], 2006). When students are learning about science, they are focused primarily on learning

---

*Corresponding author: crbeal@email.arizona.edu

scientific content, including factual information, knowledge of theories, and understanding of processes. When they are learning to do science, they use their knowledge of scientific content and processes to define and solve scientific problems. This distinction is explicitly represented in the science education content standards for many states, and at the national level (National Academies Press, 1996). For example, California includes "Investigation and Experimentation Skills" as a distinct area of science proficiency, in addition to the mastery of science facts and content knowledge (California Department of Education, 1998).

Although problem solving is recognized as an important skill, recent assessments indicate that students do not always perform well in this area (Augustine, 2005; Kuhn, 2002; OECD, 2006; Zimmerman, 2000). For example, Krajcik, Blumenfeld, Marx, Bass and Fredricks (1998) found that middle school students struggled to generate research questions, and needed considerable support to develop a plan for answering their research questions. Schauble and colleagues found that students tended to view experiments as opportunities to explore how various combinations of features produced different outcomes, rather than to discover the causal relations between features (Schauble, 1990, 1996; Schauble, Klopfer & Raghavan, 2006). Klahr and Nigam (2004) followed students who learned about the "control of variables" (CVS) strategy for conducting science experiments either through self-directed experimentation or direct instruction. Their results indicated that most students exhibited greater mastery of the CVS concept after direct instruction, suggesting that self-directed experimentation was not always productive. Overall, prior research suggests that students' undirected problem solving in science domains tends to be relatively unsystematic and that, without assistance, students are often unselective with regard to the evidence that is collected and considered (Kirschner, Sweller & Clark, 2006; Mayer, 2004; Moreno, 2004).

Students' difficulties with problem solving can be especially evident in technology-based learning environments, which often require careful planning and progress monitoring to use effectively (Schauble, 1990; Stark et al., 1999). When students can readily explore multiple sources of information and experiment with different combinations of factors in a technology-based environment, they can easily become distracted from the primary objective of using the information to solve the problem. Schwarz and White (2005) found that when students worked with software for designing models to describe various scientific phenomena, they did not always understand that the purpose of the models was to compare predictions and test hypotheses. Rather, the students appeared to view the task as one of exploring the ways in which different settings in the software led to different outcomes. Similar results were reported by Schauble (1990) who studied students' strategies for identifying factors that affected the speed of racecars in a computer-based microworld. In many cases, students combined various factors unsystematically and thus were not able to identify the relations between the causal factors with the outcomes.

One approach to improving students' problem solving is to link the technology-based activity with classroom activities designed to help students adopt good problem solving strategies. Such activities would remind students to make sure that the goal of the problem is clearly understood, identify the information that will be most helpful in solving the problem, and monitor their progress towards the solution. However, prior research indicates that teachers do not necessarily provide such guidance. In the IMMEX project, students solve realistic case-type science problems in a web-based simulation, searching through an array of menu-based resources to find content information relevant to the problem. The design of IMMEX is informed by the work of Newell and Simon (1972) who viewed problem solving as a process of reducing the difference between the current state and the goal through effective search through available information. Other research supports the view of good problem solving as involving the effective and efficient use of information (Heider & Frensch, 1996; O'Connor, Mulley & Wennberg, 2003).

Extensive prior research has shown that students vary widely in how systematically and effectively they approach IMMEX problems (Stevens et al., 2004; Stevens & Soller, 2005). Some students carefully and systematically look for information sources that are appropriate for the current case, keep track of the information that they are accessing, and answer when the information they have reviewed is sufficient to support the answer, whereas other students are less systematic, often reinspecting information they have already viewed (Stevens & Thadani, 2007; Soller & Stevens, 2007).

Thadani, Stevens and Tao (2009) found that students' strategies varied with the way that science teachers implemented IMMEX activities in the classroom. Some teachers framed the activity for students, reminding them to check that they clearly understood the problem before starting to work, and then encouraging students to organize their resources, take notes, and reflect on their progress as they worked. Thadani et al. (2009) found that students in these classes showed more effective problem solving than other students whose teachers were less likely to provide such guidance about managing the process of problem solving, often focusing more on reminding students about specific content or suggesting that the student look at a particular piece of information.

Other research indicates that teachers often need considerable training and support to implement technology-based science activities effectively. Schwarz and White (2005) found that students' scientific problem solving improved when the computer-based activity of designing models was enhanced with an intensive classroom-based curriculum. However, it was not clear if teachers could have implemented the activities and effectively without support from university researchers. In the Technology-Enhanced Learning in Science (TELS) project, students work with inquiry modules that include visualizations for various science topics, and teachers receive extensive professional development training in effective implementation TELS activities. However, barriers such as problems with the school technology at times made it difficult for many science teachers to focus pedagogical goals such as guiding students in aspects of good problem solving (Varma, Husic & Linn, 2008).

Because prior work suggests that science teachers may need intensive support to provide students with appropriate guidance about problem solving, it is important to consider complementary approaches. One such approach is to integrate guidance about problem solving directly into the technology-based learning environment. Such guidance may include the types of suggestions and prompts about the metacognitive aspects of good problem solving that have been associated with effective teacher implementation and skilled instruction from expert human tutors. More specifically, good problem solvers do more than apply known procedures to familiar problems. Rather, they consider carefully the nature of the problem before starting to work, plan an appropriate approach, implement the plan, and continually evaluate progress towards the solution (Cooper & Sandi-Urena, 2009; Swanson, 1990). Good problem solvers also recognize that difficult problems may require time and effort to solve, and that some “moments in the dark” are to be expected during the problem solving process (Halpern, 1998). If the kinds of metacognitive guidance provided by skilled teachers could be integrated directly into simulation learning environments, then we might expect to find students adopting better strategies (Brown et al., 1998).

Prior research has investigated the potential of integrated guidance to help students improve their problem solving when working with technology-based science systems. For example, students learned math procedures better when a multimedia game included verbal instructions than when the game did not include the verbal guidance, although the prompts were more oriented to domain specific facts than metacognitive advice (Moreno & Duran, 2004). Stark et al. (1999) found that students performed better in a simulation environment for economics when they could access hints from experts along with worked examples. Benefits linked to the availability of worked examples were also reported by Yaman, Nerdel and Bayrhuber (2008) for students who learned about respiration in a simulation environment.

Other studies have investigated the potential of integrated guidance that addresses the more metacognitive aspects of problem solving, as an alternative to providing domain-specific hints and worked examples. Hollingworth and

McLoughlin (2001) designed a suite of prompts to encourage metacognitive reflection on problem solving for university students working in a distance learning environment for chemistry, however, the system was not fully evaluated. Toth, Suthers and Lesgold (2002) found that prompts designed to encourage reflection on problem solving helped students when they used software for evidence mapping. Sandoval & Reiser (2004) developed a computer-based tool that helped students organize and evaluate information sources about evolutionary theory, with positive results. Yelland and Masters (2007) found that scaffolding improved the strategies used by collaborative student pairs over unguided problem solving. There was also evidence that the scaffolding addressed motivational issues: The student pairs appeared more engaged and enthusiastic when the scaffolding guidance was present than when the pairs worked without the electronic support.

Additional support is found in studies that compared the same simulation environment used by individual students with and without integrated guidance about problem solving. Kauffman, Ge, Xie and Chen (2008) provided undergraduate education students with prompts designed to promote metacognitive reflection as they worked in a Web-based instructional module about classroom management. Students viewed scenarios about teachers who were struggling with various classroom management problems, such as disruptive student behavior, and were asked to write analyses of the problem and recommendations for addressing the problems. The results indicated that students who viewed prompts (e.g., "What is the primary concern?" and "How do you know that this is a problem?") were more likely to identify the target problem accurately and outline good solutions than those who did not view prompts.

Veermans, van Joolingen, Wouter and de Jong (2006) investigated the impact of problem solving guidance integrated into a simulation designed to help students discover physics principles related to collisions. The guidance consisted of heuristic suggestions presented to students about experimental design and problem solving, such as explicitly reminding students to vary one thing at a time. One group of students used the simulation with the explicit heuristics, whereas a second group used a version that provided more implicit guidance in the form of suggestions about specific values to try in the simulation and what the expected results should be. Students in both groups improved in their knowledge of the physics content, but there were indications that those who viewed the explicit heuristic guidance also had developed a deeper understanding of the underlying principles and were able to design better experiments. The specific heuristics investigated in the study included some that addressed metacognitive knowledge (e.g., "Keep track of what you are trying"). However, other prompts focused fairly specifically on experiment design (e.g., "Try extreme values"). Chang, Chen, Lin and Sung (2008) also found that students benefited from hints and guidance about forming a hypothesis and conducting a good experiment as they worked with a simulation about optics.

The present study was designed to learn if the addition of simple message prompts designed to help students reflect on the process of problem solving would lead to improved performance in the IMMEX simulation, relative to students who simply solved the problems without the integrated guidance. The message prompts were derived from an analysis of the suggestions and strategies used by teachers who implement the activity effectively (Thadani et al., 2009) and from studies of expert tutors drawn from the intelligent tutoring systems theoretical framework (Brown et al., 1988; Lepper, Woolverton, Mumme & Gurtner, 1993; Murray & Arroyo, 2002; Wood & Wood, 1996; Woolf, 2009). These strategies typically involve suggestions about how to approach a problem, including making sure that the goals of the problem are well understood before starting to work on it, actively making a plan for solving the problem rather than simply trying various actions in turn, and monitoring progress towards the solution. In addition, expert tutors and teachers provide guidance about the motivational and emotional aspects of problem solving, such as the importance of not becoming discouraged when a problem is not immediately solved, and that learning to solve new and challenging problems takes effort and persistence. The hypothesis investigated in the study was that students who viewed messages within the IMMEX simulation would show improved problem solving, relative to students who received no guidance about the metacognitive aspects of problem solving.

METHOD

Participants

The participants were Grade 9 students in one large school district in Southern California. The district incorporates IMMEX problem sets into its high school science curriculum. In the 2006-2007 school year, students in Grade 9 Chemistry used the message-enhanced version of Duck Run. Data from 195 students who completed at least five cases were located in the IMMEX database and extracted for analysis. Comparison data from students (N = 173) who had completed at least five cases in the original version of Duck Run during the previous year (the 2005–2006 school year) were also located and extracted for analysis. Data from the end-of-year California Standards Test (CST) in Chemistry indicated that the student samples appeared to be comparable across the two years. The mean scale score for Chemistry was 369.5 in the Spring 2006 test administration and 364.8 in the Spring 2007 administration (California Department of Education Standardized Testing and Reporting).

MATERIALS

IMMEX problem set

Students in the study worked with the Duck Run scenario for high school chemistry, which begins with a prologue describing that an unknown substance has been illegally dumped into a local duck pond, possibly putting the local wildlife at risk (Stevens & Palacio-Cayetano, 2003; Stevens et al., 2004). The student's task is to identify the substance so that it can be properly removed.

After the problem is presented in the prologue, the student can move on to view 12 menu-linked information sources, including the results of various chemical tests performed on the unknown substance, its number of electrons, its state of matter (liquid, solid), and descriptions provided by witnesses who saw the substance and can provide information about some of its characteristics. The student can also access the periodic table of elements and other reference materials. All the information necessary to solve the problem is available in the case; the student's task is to search for and integrate the relevant information.

Each request by the student to view an information source in the Duck Run simulation costs the student points that are deducted from a starting total. This means that students must think carefully about the information that they really need to identify the unknown substance. When students feel they have gathered enough information, they attempt to answer the problem. In the case of Duck Run, the student chooses his or her answer (e.g., aluminum) from a list of potential unknowns and receives immediate feedback (correct, incorrect). Students have two attempts to solve the problem, meaning that they can solve it on the first attempt, the second attempt, or not at all.

The Duck Run IMMEX simulation includes 12 different versions ("cases") of the problem. Each case follows the same general scenario, but with different target elements (e.g., aluminum, tin, phosphorous, iodine, etc.) and associated information sources. Each case takes 13 minutes to solve, on average. Cases vary in difficulty (based on data from prior users, Stevens & Soller, 2005) and are presented in random orders across students.

Message prompts

For the study, the Duck Run simulation was modified so that brief text messages about the problem solving process could be presented as part of the prologue screen for each case. Messages were randomly selected for each case from a bank of 36, with the restriction that a particular message would only be shown once to an individual student. Examples include, "Sometimes students just dive in to a problem and try out lots of different things. But it's better to make sure you know what the problem is asking first. If you don't, go back and read the prologue again, or ask your teacher for help." "Try reading the prologue for this case and then list three things you've learned in your science class that might help you solve it. Thinking about what you already know will help you make progress." "IMMEX problems give you lots of information, and it's easy to get confused about what you've already looked at. On this case, try listing the resources that you want to use, and plan out the order that you'll look at them."

Some messages focused on helping students understand that it was normal to experience emotions such as uncertainty, frustration or discouragement while solving challenging problems, and provided suggestions for recognizing and handling motivational issues appropriately. Examples include: "Sometimes, students give up too fast on IMMEX problems. They think the answer should come in just a couple of minutes. But hard problems can take a lot of mental effort. Stick with it and you'll see results!" "When you work on a hard problem, sometimes you're going to feel confused. That's a sign your brain is trying out different ideas. Don't give up too soon on this case." "Sometimes, you might not be sure how to solve an IMMEX problem. That's OK – if you already knew what to do, you wouldn't really be learning anything new. Challenge gives you the chance to learn."

Scoring

In IMMEX, students' requests to view the different information sources are automatically recorded as they work on each case. In the present study, students completed five Duck Run cases. Each case was assigned a "Solve" score of 2 if the student solved the problem on the first attempt, a score of 1 if the student solved the problem on the second attempt, or a score of 0 if the problem was not solved. Thus, each student had five Solve scores, one for each completed case.

Additionally, each student received a score for each case indicating his or her strategic efficiency. As students look at various information sources, their actions are automatically recorded by the IMMEX software, and a Hidden Markov Model (HMM) is used to analyze their efficiency in reviewing and utilizing the information (whether or not the appropriate information sources are viewed, if the same information is viewed multiple times) and effectiveness (whether the problem was solved correctly in one or two attempts, or not solved at all) in relation to the difficulty of the specific case (established by prior users of IMMEX). The resulting Strategic Efficiency Index (SEI) score is automatically calculated by the IMMEX software, and provides a quantitative metric that can be used to evaluate the quality of a student's problem solving on a single case, and to compare performance across cases (for details see Stevens & Soller, 2007; Stevens & Thadani, 2007).

RESULTS

Mean Solve scores for the two groups of students are shown in Table 1. An analysis of variance was conducted on students' Solve scores, with Group (message-enhanced, no messages) as a between subjects factor and Case (1 through 5) as a within subjects factor. There was a significant effect of Group, $F(1,366) = 141.640$, $p < .001$, indicating that Solve scores were consistently higher for students who viewed messages in the Duck Run cases. No other effects were significant.

Students' problem solving was also investigated in relation to the difficulty of the cases. Not surprisingly, the case involving mercury as the unknown substance was easiest for students ($M = 1.77$ out of a maximum possible score of 2) due to its unique characteristic as a liquid metal. The case in which tin was the unknown substance was most challenging ($M = 1.23$). Table 2 shows the mean Solve scores for students in the two conditions (message-enhanced, no messages) for the 12 cases ordered by difficulty. As may be seen in Table 2, the integrated messages appeared to be most helpful on the more challenging cases, whereas the difference in Solve scores for the two conditions (message-enhanced, no-messages) was smaller on the easier cases.

Mean SEI scores are shown in Table 3. The SEI scores were analyzed in an analysis of variance with Group (message-enhanced, no-messages) as the between subjects factor and Case (1 through 5) as the within subjects factor. There was a main effect of Group, $F(1,366) = 47.651$, $p < .001$. This effect indicates that students who worked with the message-enhanced version were more likely to search through the available resources in the simulation efficiently and effectively. There was also a main effect of Case, $F(1,363) = 6.654$, $p < .001$, suggest-

TABLE 1
Mean Solve scores by Case for students using two versions of Duck Run simulation. Standard deviations are shown in parentheses.

	Case 1	**Case 2**	**Case 3**	**Case 4**	**Case 5**
Message-enhanced (N = 195)	1.67 (.47)	1.72 (.44)	1.84 (.36)	1.77 (.36)	1.75 (.42)
No messages (N = 173)	1.21 (.90)	1.26 (.88)	1.31 (.87)	1.31 (.87)	1.45 (.83)

TABLE 2
Mean Solve scores for students in two versions of Duck Run simulation for cases ranked by difficulty. Standard deviations are shown in parentheses.

Cases by Difficulty:	Message-enhanced	No Messages
Tin (most difficult)	1.76 (.43)	.70 (.87)
Iron	1.67 (.47)	0.84 (.91)
Phosphorous	1.79 (.40)	1.40 (.87)
Lead	1.68 (.46)	1.16 (.86)
Silver	1.54 (.50)	1.20 (.86)
Iodine	1.59 (.47)	1.27 (.84)
Gold	1.84 (.37)	1.19 (.96)
Magnesium	1.77 (.42)	1.37 (.87)
Copper	1.87 (.33)	1.52 (.80)
Aluminum	1.78 (.41)	1.63 (.68)
Carbon	1.88 (.35)	1.68 (.69)
Mercury (easiest)	1.83 (.38)	1.71 (.63)

TABLE 3
Mean Strategic Efficiency scores by Case. Standard deviations are shown in parentheses.

	Case 1	Case 2	Case 3	Case 4	Case 5
Message-enhanced (N = 195)	4.36 (3.29)	5.04 (3.06)	5.53 (3.25)	5.50 (3.08)	5.51 (3.17)
No messages (N = 173)	3.24 (2.30)	3.76 (2.73)	3.98 (2.59)	3.97 (2.53)	4.50 (2.45)

ing that students gradually became more strategic across the five cases that they completed. However, the interaction was not significant, indicating that both groups improved similarly.

An exploratory analysis was conducted to investigate the impact of integrated messages on students with relatively weak problem solving skills, because prior research indicated that these students tended to stabilize with and retain poor problem solving strategies (Stevens et al., 2004). More specifically, students who did not solve their first case on the first attempt might reasonably be considered to be less successful problem solvers than students who did solve their first problem. Data records for students who failed to solve their first case were located. There were 64 students who worked with the message-enhanced version of Duck Run, and 80 students who worked with the no-message version. Mean scores for these students are shown in Table 4. An analysis of variance was conducted with Group (message-enhanced, no messages) as the between subjects factor, Case (2 through 5) as the within subjects factor, and Strategic Efficiency Index scores as the outcome measure. The results indicated a main effect of Group, $F(1,142) = 10.457$, $p < .01$. A significant effect of Case, $F(3,140) = 3.828$, $p < .05$, indicated that students tended to become more strategic as they solved more problems. There was also a significant Group x Case interaction, $F(3,140) = 2.868$, $p < .05$. The interaction indicates that students who did not initially do well improved more on subsequent cases when they received scaffolding messages about good problem solving than students who did not receive any messages.

Additional analyses were conducted to learn if the messages specifically reduced the use of exhaustive strategies. Duck Run cases were identified in which the student looked at 10, 11 or all 12 of the information sources available in the case at least once. Results are presented in Table 5. On the first case, the percentage of students who used exhaustive strategies was similar for the two groups. However, exhaustive strategies declined more across cases for students who worked with the message-enhanced version of Duck Run. That is, the messages appeared to reduce students' tendency to look at all available information.

Interestingly, for the few students who persisted with exhaustive strategies on the later cases, the average Solve scores were higher when messages were present than when no messages were included. For example, as shown in Table 5, when we considered the fifth completed cases, only 4% are still associated with exhaustive strategies in the message-enhanced version, and only 7% in the no-message version. However, the use of exhaustive strategies is more likely to result in a correct solution in the message-enhanced version (mean solve score of 1.62 out of a possible 2, relative to a mean of 0.92 in the no-message version). This suggests the possibility that the messages might have encouraged better integration of the relevant information by students, even when they continued to use a relatively inefficient problem solving strategy.

TABLE 4
Mean Strategic Efficiency Scores for Cases 2–5 for students who did not solve Case 1. Standard deviations are shown in parentheses.

	Case 2	**Case 3**	**Case 4**	**Case 5**
Message-enhanced (N = 64)	5.08 (3.11)	5.79 (2.92)	4.76 (3.25)	5.72 (3.47)
No messages (N = 80)	4.33 (2.86)	3.97 (2.63)	4.22 (2.68)	4.29 (2.63)

TABLE 5
Mean Solve scores (standard deviations in parentheses) for students who used exhaustive strategies on Cases 1 – 5.

Case		Number of students using exhaustive strategies	Mean Solved Score
1	Messages	46 (23%)	1.50 (.50)
	No Messages	45 (25%))	1.06 (.93)
2	Messages	18 (10%))	1.72 (.46)
	No Messages	33 (19%)	1.03 (.95)
3	Messages	17 (9%)	1.82 (.30)
	No Messages	21 (12%)	0.76 (.83)
4	Messages	11 (6%)	1.63 (.50)
	No Messages	23 (13%)	0.61 (.89)
5	Messages	8 (4%)	1.62 (.51)
	No Messages	13 (7%)	0.92 (.95)

GENERAL DISCUSSION

The goal of the study was to evaluate the hypothesis that the integration of metacognitive scaffolding into a technology-based science problem solving environment would lead to more effective problem solving by students. The results indicated that the addition of simple text messages designed to promote reflection on problem solving was associated with better problem solving behavior by students, including higher solve rates and higher probability of success on the more challenging cases.

It is important to note that the scaffolding messages did not provide information about the science content that would help the student solve the problem. In fact, all the relevant science content information is already available in the case; the student's task is to think about which information might be most useful, and to relate one information source to another in order to converge on the solution. Thus, the scaffolding messages were designed to address problem solving as a process, and to encourage students to focus on their actions in relation to the goal of solving the problem. Prior research has shown that students tend to view simulations as opportunities to explore the available resources and to try out various features, rather than to use the resources to address a specific question. The integration of messages into the simulation appeared to help students adopt more successful problem solving strategies. The change in strategic behavior was somewhat surprising given that the messages were quite brief and presented in simple text as part of the prologue to each case. It seemed quite possible that students would simply ignore the information; however, the results of the automatically-derived assessments of strategic efficiency suggests that they did not.

One possible interpretation for the results might be that the messages might have helped simply by interrupting students' tendency to jump into the problem and start exploring the resources without clearly considering the goal and making a plan for investigating the resources to find the solution. That is, perhaps the benefits were due to the mere presence of text on the screen before the student started the problem, rather than to the message content. Although this is a plausible hypothesis, prior work with indicates that the message content is important. In a study conducted with another IMMEX problem set, university students were assigned to work with a version that included metacognitive messages, no messages, or messages that provided generic advice about good study habits. Examples of the generic academic messages included, "Keep up with your class reading" and "Successful students tend to use a daily planner to keep track of assignments." The results indicated that the generic academic messages had no benefits for students' problem solving, whereas the metacognitive messages were associated with better performance (Stevens, Beal & Sprang, 2009). It must be acknowledged that the study involved a different problem set, the students were several years older and that the university sample may have been more selective than was the case for the high school students in the present study. However, the results suggest that the content of the messages, not simply the presence of text, is important for students' improved problem solving.

The major limiting factor in the study was the comparison of students' performance across school years. It is of course possible that the students who used Duck Run in the year that it included messages were simply better students, and that the stronger results were due to sample differences rather than the addition of scaffolding messages. However, one might then have expected to see differences on other performance metrics such as the state achievement test in Chemistry, which was not the case. Also, the proportion of students who started out using exhaustive strategies in the simulation was similar in the two groups. In addition, the beneficial impacts of messages were differential rather than uniform, which might have been expected if one group of students was simply academically superior to the other. For example, when we identified students who started out doing poorly, or those who continued to use exhaustive strategies on case after case, those who viewed messages still tended to perform better than those who did not. The messages also appeared to be especially helpful for the cases that were most difficult. Thus, the overall pattern of the results suggests that the messages had a beneficial impact on students' problem solving.

Ideally, the present study should be replicated with an experimental design in which students in the same classes were randomly assigned to use the message-enhanced version or the no-message version. However, it is often difficult to implement true experiments in authentic classroom settings; school personnel may either expect that a promising intervention will be provided to all students, or decline to participate on the grounds that there is not enough evidence that the intervention might help. Comparison of students' performance across years can provide insights into the effects of interventions (Lee, Linn, Varma & Lui, 2010). Encouraging findings from delayed cohort comparisons, as in the present study, can help to make the case that true experiments are warranted.

The results of the present study add to the growing body of evidence suggesting that technology-based environments can help students build proficiency with scientific problem solving, particularly when the environment is designed to provide students with some guidance about how to proceed. In the original version of the simulation, some students' actions suggested that their goal was to explore the various resources that were available, rather than to locate and use relevant information in a focused effort to solve the problem (Stevens et al., 2004; Stevens & Thadani, 2007). The integration of messages based on the implementation strategies of effective teachers and tutors into the simulation appeared to improve students' problem solving, defined in terms of an increased probability of finding the correct solution, and a reduction in the use of exhaustive and unproductive search strategies (Heider & Frensch, 1996).

The results also provide an illustration of the potential of technology-based learning environments to assess students' problem solving performance and their progress (Nirmalakhandan, 2007; Quellmalz & Pellegrino, Stevens & Thadani, 2007). By tracking students' actions and comparing their behavior to performance models developed on the basis of prior users, it is possible to determine with some precision the impact of interventions such as the example in the present study (Spillane, Reiser & Reimer, 2002). The integration of interactive instruction with real time assessment information automatically captured by technology-based environments will offer teachers, students and researchers new insights about how best to support students as they master problem solving skills in science domains.

ACKNOWLEGMENTS

We would like to thank the teachers and students who participated in the study, and the staff and administrators at the Placentia-Yorba Linda School District in California. Special thanks are due to Dr. Marcia Sprang for her assistance with the integration of IMMEX into the science curriculum and with teacher professional development. The research was supported by a grant from the U.S. Department of Education's Institute of Education Sciences (R305H050052). The views expressed are not necessarily those of the sponsoring agency.

REFERENCES

Augustine, N. R. (2005). National Academies Committee on Prospering in the Global Economy of the 21st Century. Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future. National Academies Press, Washington DC.

Brown, A. L., Ellery, S., & Campione, J. (1998). Creating Zones of Proximal Development electronically. In J. Greeno & S. Goldman (Eds.), Thinking practices: A symposium in mathematics and science education, pp. 341–368. Hillsdale NJ: Erlbaum.

California Department of Education Standardized Testing and Reporting. http://star.cde.ca/gov/ Data for Placentia-Yorba Linda Unified School District Grade 9 Chemistry retrieved Oct. 14, 2009.

California Department of Education. (1998). Science Content Standards for California Public Schools. Retrieved Nov. 23, 2009 from http://www.cde.ca.gov/be/st/ss/

Chang, K., Chen, Y., Lin, H., & Sung, Y. (2008). Effects of learning support in simulation-based physics learning. Computers and Education, 51, 1486–1498.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. Science Education, 86, 175–218.

Cooper, M. M., & Sandi-Urena, S. (2009). Design and validation of an instrument to assess metacognitive skillfulness in chemistry problem solving. Journal of Chemical Education, 86, 240.

Duschl, R. A. (2008). Science education in three part harmony: Balancing conceptual, epistemic and social learning goals. Review of Research in Education, 32, 268–291.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Dispositions, skills, structure thinking, and metacognitive monitoring. American Psychologist, 53, 449–455.

Heider, H., & Frensch, P. A. (1996). The role of information reduction in skill acquisition. Cognitive Psychology, 30, 304–337.

Hollingworth, R. W., & McLoughlin, C. (2001). Developing science students' metacognitive problem solving skills online. Australian Journal of Educational Technology, 17(1), 50–63. http://www.ascilite.org.au/ajet/ajet17/hollingworth.html

Kauffman, D. F., Ge, X., Xie, K., & Chen, H. (2008). Prompting in web-based environments: Supporting self-monitoring and problem solving skills in college students. Journal of Educational Computing Research, 38, 115–137.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential and inquiry-based teaching. Educational Psychologist, 41, 75–86.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction. Psychological Science, 15, 661–667.

Krajcik, J., Blumenfeld, P. C., Marx, R. W., Bass, K. M., & Fredricks, J. (1998). Inquiry in project based science classrooms: Initial attempts by middle school students. Journal of the Learning Sciences, 7(3/4), 313–350.

Kuhn, D. (2002). What is scientific thinking and how does it develop? In U. Goswami (Ed.), Blackwell Handbook of Childhood Cognitive Development, pp. 371–393. MA: Blackwell Publishing.

Lee, H., Linn, M. C., Varma, K., & Lui, O. L. (2010). How do technology-enhanced inquiry science units impact classroom learning? Journal of Research in Science Teaching, 47, 79–90.

Lepper, M. R., Woolverton, M., Mumme, D., & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. P. Lajoie & S. J. Derry (Eds.), Computers as cognitive tools (pp. 75–105). Hillsdale NJ: Erlbaum.

Mayer, R. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. American Psychologist, 59, 14–19.

Moreno, R. (2004). Decreasing cognitive load in novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. Instructional Science, 32, 99–113.

Moreno, R., & Duran, R. (2004). Do multiple representations need explanation? The role of verbal guidance and individual differences in multimedia mathematics learning. Journal of Educational Psychology, 96, 492–503.

Murray, T., & Arroyo, I. (2002). Toward measuring and maintaining the Zone of Proximal Development in adaptive instructional systems. Proceedings of the 6th International Conference on Intelligent Tutoring Systems. In S. A. Cerri, G. Gouardères & F. Paraguaçu (Eds.), Lecture Notes in Computer Science 2363, pp. 133–145. Springer Berlin.

National Academies Press. (1996). National Science Education Standards. Washington DC: National Academies Press.

National Research Council [NRC]. (2006). America's lab report: Investigations in high school science. Washington DC: National Academy Press.

Newell, A., & Simon, H. (1972). Human problem solving. Englewood Cliffs NJ: Prentice Hall.

Nirmalakhandan, N. (2007). Computerized adaptive tutorials to improve and assess problem solving skills. Computers and Education, 49, 1321–1329.

O'Connor, A. M., Mulley, A. G., & Wennberg, J. E. (2003). Standard consultations are not enough to ensure decision quality regarding preference-sensitive options. Journal of the National Cancer Institute, 95, 570–571.

OECD. (2006). PISA 2006 Science Competencies for Tomorrow's World. C. Shrewbridge & A. Schleicher (Eds.), Programme for International Student Assessment, Organization for Economic Cooperation and Development, Paris France.

Quellmalz, E. S., & Pellegrino, J. W. (2009). Technology and testing. Science, 323, 75–79.

Sandoval, W., & Reiser, B. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. Science Education, 88, 345–372.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. Developmental Psychology, 32, 102–119.

Schauble, L. (1990). Belief revision in children. Journal of Experimental Child Psychology, 49, 31–57.

Schauble, L., Klopfer, L. E., & Raghavan, K. (2006). Students' transition from an engineering model to a science model of experimention. Journal of Research in Science Teaching, 28, 859–882.

Schwartz, C. V., & White, B. Y. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. Cognition and Instruction, 23, 165–205.

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. Review of Educational Research, 72, 387–431.

Soller, A. & Stevens, R. H. (2007). Applications of stochastic analysis for collaborative learning and cognitive assessment. In G. Hancock & K. Samuelson (Eds.), Advances in Latent Variable Mixture Models. Information Age Publishing.

Stark, R., Mandl, H., Gruber, H., & Renkl, A. (1999). Instructional means to overcome transfer problems in the domain of economics: Empirical studies. International Journal of Educational Research, 31, 591–609.

Stevens, R. H., Beal, C. R., & Sprang, M. (2009). Tracking the development of problem solving skills with learning trajectories. In Kong, S. C., Ogata, H., Arnseth, H. C., Chan, C. K. K., Hirashima, T., Klett, F., Lee, J. H. M., Liu, C. C., Looi, C. K., Milrad, M., Mitrovic, A., Nakabayashi, K., Wong, S. L., & Yang, S. J. H. (Eds.) (2009). Proceedings of the 17th International Conference on Computers in Education [CDROM]. Hong Kong: Asia-Pacific Society for Computers in Education.

Stevens, R. H, & Palacio-Cayetano, J. (2003). Design and performance frameworks for constructing problem-solving simulations. Cell Biology Education, 2, 162–179.

Stevens, R. H., Soller, A., Cooper, M., & Sprang, M. (2004). Modeling the development of problem solving skills in chemistry with a web-based tutor. Artificial Intelligence in Education, 58–591. Berlin: Springer Verlag.

Swanson, H. L. (1990). Influence of metacognitive knowledge and aptitude on problem solving. Journal of Educational Psychology, 82, 306–314.

Stevens, R. H., & Soller, A. (2005). Machine learning models of problem space navigation. Computer Science and Information Systems, 2(2), 83–98.

Stevens, R. H. & Thadani, V. (2007). Quantifying students' scientific problem solving efficiency and effectiveness. Technology, Instruction, Cognition and Learning, 5, 325–337.

Thadani, V., Stevens, R. H., & Tao, A. (2009). Measuring complex features of science instruction: Developing tools to investigate the link between teaching and learning. Journal of the Learning Sciences, 18, 285–322.

Toth, E. E., Suthers, D. D., & Lesgold, A. (2002). Mapping to know: The effects of representational guidance and reflective assessment on scientific inquiry. Science Education, 86, 264–286.

Varma, K., Husic, F., & Linn, M. C. (2008). Targeted support for using technology-enhanced science inquiry modules. Journal of Science Education and Technology, 17, 1573–1839.

Veermans, K., van Joolingen, W., & de Jong, T. (2006). Use of heuristics to facilitate scientific discovery learning in a simulation learning environment in a physics domain. International Journal of Science Education, 28, 341–361.

Wood, D., & Wood, H. (1996). Vygotsky, tutoring and learning. Oxford Review of Education, 22, 5–16.

Woolf, B. P. (2009). Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning. Burlington MA: Morgan Kaufman Publishers.

Yaman, M., Nerdel, C., & Bayrhuber, H. (2008). The effects of instructional support and learner interests when learning using computer simulations. Computers and Education, 51, 1784–1794.

Yelland, N., & Masters, J. (2007). Rethinking scaffolding in the information age. Computers and Education, 48, 362–382.

Zimmerman, C. (2000). The development of scientific reasoning skills. Developmental Review, 20, 99–149.