

EDUCATION TECHNOLOGY

Computer-Guided Inquiry to Improve Science Learning

Marcia C. Linn^{1*}, Libby Gerard¹, Kihyun Ryoo², Kevin McElhaney³, Ou Lydia Liu⁴, Anna N. Rafferty⁵

Engaging students in inquiry practices is known to motivate them to persist in science, technology, engineering, and mathematics (STEM) fields and to create lifelong learners (1, 2). In inquiry, students initiate investigations, gather data, critique evidence, and make sophisticated drawings or write coherent essays to explain complex phenomena. Yet, most instruction relies on lectures that transmit information and multiple-choice tests that determine which details students recall. Massive Open Online Courses (MOOCs) mostly offer more of the same. But new cyber-learning tools may change all this, by taking advantage of new algorithms to automatically score student essays and drawings and offer personalized guidance.

Empowering Learners, Aiding Instructors

Inquiring students make predictions, gather new ideas (from investigations, visualizations, or observations); use evidence to distinguish among their predictions and ideas; and create a coherent explanation (1, 3, 4). When they write explanations, students learn more than when they select among multiple-choice answers or read explanations (5). Experimenting with visualizations and making drawings to illustrate ideas can develop students' spatial-reasoning skills (6). And, when students analyze resources to develop an explanation, they appreciate the beauty and complexity of new fields (7).

Analyzing students' essays or drawings and using the results to guide them can improve outcomes (8), but this requires more human capital than precollege and college instructors have. Precollege instructors often have five or six classes of 30 to 40 students, and college instructors may have hundreds or thousands of students in required courses or MOOCs.

However, advances in computer technologies may help offset limits in instructor

time and effort. Immediate, personalized, computer-generated guidance can motivate students to deepen their understanding of complex materials. Instructors can review the automated scores to identify students who continue to flounder. Because these systems can assign guidance to every student, even students who are reluctant to ask for help can progress.

Automated guidance on essays and drawings can improve learning in precollege and college courses.

(WISE) streamline this process by recording student ideas and supporting experiments for which researchers can randomly assign diverse forms of guidance.

These experiments show that guidance encouraging students to distinguish among their predictions and new evidence helps students to better integrate their ideas about the topic (9). Once the links between student

Automated guidance was as effective as guidance provided by an expert teacher.

To make sure that inquiry activities lead to new insights and not to erroneous or superficial conjectures, experienced teachers monitor student progress and regularly add hints to keep students on the right track (3, 4, 9) (see the first figure). To figure out what hints will help students explore a complex problem, researchers and teachers (often collaboratively in professional development programs) analyze large numbers of student essays or drawings and try out alternative approaches (3). Online learning environments like the Web-based Inquiry Science Environment

responses and effective guidance are established, environments like WISE can score student essays or drawings and automatically assign guidance designed to help students develop coherent explanations.

Guiding Writers, Drawers, and Teachers

Advances in natural language processing now enable computer-based learning environments to use scored answers to create systems for scoring future responses. For instance, the Educational Testing Service's (ETS) "c-rater" tool used human-rated responses to develop

Students were asked to use stamps to represent the chemical reaction between two methane molecules (CH_4) and oxygen (O_2) to yield carbon dioxide (CO_2) and water (H_2O). The drawing interface allowed students to create multiple "frames" in their drawing, with the first frame representing the reactants and the second frame representing the products.

Good start. You have correctly created 2 frames that represent the reactants and products of the methane combustion reaction.

Can atoms in the reaction be spontaneously CREATED OR DESTROYED?

Reread the direction and revisit steps 3.6–3.8. Then improve your drawings.

Computerized hints to improve understanding. Student response to the chemical reactions item that WISE can automatically score. Automated guidance was as effective as guidance provided by an expert teacher. [Adapted figure, based on WISE]

¹Graduate School of Education, University of California Berkeley, Berkeley, CA 94720, USA. ²School of Education, University of North Carolina, Chapel Hill, Chapel Hill, NC 27599, USA. ³SRI International, Menlo Park, CA 94025, USA. ⁴Educational Testing Service, Princeton, NJ, 08541, USA. ⁵Computer Science Division, University of California Berkeley, Berkeley, CA 94720, USA. *Corresponding author. mclinn@berkeley.edu

QUESTION: Burning coal to produce electricity has increased the carbon dioxide content of the atmosphere. What possible effect could the increased amount of carbon dioxide have on our planet? Explain your answer.

“ I think that it has a lower relative humidity because the carbon dioxide takes up the space of the air molecules and so there is less humidity. **c-rater score: 1**

“ The burning of coal will make a warmer climate, because the coal will make a smog that keeps in the warm air. **c-rater score: 2**

“ This is just like the car because of the green house effect the extra carbon dioxide will trap the sun's rays in our atmosphere. This is called global warming. **c-rater score: 3**

Computer evaluation of written answers. Student response to the coal item that c-rater can automatically score. Guidance that encourages students to look at relevant evidence and revise their ideas can be associated with these scores. [Figure based on c-rater]

an automated scoring system for an activity aimed at figuring out how burning coal to produce electricity affects the environment (10) (see the second figure). The c-rater system successfully scored new responses [inter-rater agreement reflected by κ coefficient = 0.87 (10)]. WISE can use these scores to assign personalized guidance while students are learning. Similarly, Writing-Pal (11) analyzes writing quality and recommends strategies for strengthening the essay (e.g., “think about the quality of your evidence”) rather than identifying distinct errors. Students using Writing-Pal substantially improved their essays by elaborating content, improving paragraph structure, and using more precise vocabulary (11).

AutoTutor prompts students to explain their reasoning about how to solve a physics problem through a written dialogue with a computer avatar. The computer analyzes students' explanations for misconceptions and asks questions designed to elicit better explanations. This guidance was more effective than studying a well-written text on conceptual physics for improving physics problem-solving (12). Although automated scoring algorithms for written responses must be tailored to individual questions, this process can be as simple as providing the system with about 500 to 1000 existing human-scored responses and desired guidance messages.

Computers can analyze human-scored drawings to create systems that can then score subsequent drawings and provide feedback automatically (13) (see the first figure) (Fig. 2). For instance, in a classroom study, automated guidance was highly accurate and as effective as teacher-provided guidance for stimulating understanding of chemical reactions (13). Similarly, ASSISTments research-

ers found that proficient students could take advantage of receiving annotated solutions to math problems, designed by an expert instructor but chosen and offered by a computer, to improve their performance on future problems. Less-proficient students benefited from automated guidance that emulated an expert teacher by asking questions but not giving answers (14). In addition, when using Betty's Brain, students constructed more accurate concept maps (diagrams showing relations between concepts) when they received conceptual guidance than when they received explicit directions (15). These studies support the value of guidance that encourages students to reconsider their ideas rather than telling them the right answer.

Teachers report spending up to 10 min to compose detailed guidance for each work group doing an inquiry activity and at least 2 min per work group assigning premade comments (13). When guidance is automated, teachers can instead concentrate on students the computer identifies as not making progress. Teachers can also review automated scores to gauge overall class progress or to design activities that address specific conceptual difficulties. Instructors can use student responses immediately to adjust their instruction and annually to revise courses for the next group of students.

The Need for Design and New Research

Inquiry activities in which students grapple with open-ended questions and come up with novel solutions are now feasible in large classes and MOOCs by taking advantage of automated scoring and guidance. Of course, instructors still need to design the inquiry activities, scoring rubrics, and guidance. The computer only assigns the guidance.

These technologies open up exciting opportunities. Writing explanations for complex phenomena can help students develop greater understanding in STEM and language arts. As advances increase the precision of scoring technologies, assessments featuring written work and drawings can replace detail-oriented multiple-choice questions. Adding drawing activities can help students who lack proficiency in reading and writing to explain their ideas. These technologies also raise crucial research questions including how best to design guidance and whether automated guidance works for all students. Research can help determine ways to speed up the process of creating new inquiry activities and associated automated guidance.

These findings illustrate the need for open-source learning environments that support inquiry and can be widely used in typical courses. For example, automated guidance could improve performance and sustain participation in MOOCs. Adding inquiry features and guidance to support learners' investigations should be a high priority for designers of learning environments.

References and Notes

1. National Research Council, *Taking Science to School*, R. A. Duschl, H. A. Schweingruber, A. W. Shouse, Eds. (National Academies Press, Washington, DC, 2007).
2. Next Generation Science Standards (Achieve, Inc., Washington, DC, 2013); www.nextgenscience.org/.
3. M. C. Linn, B.-S. Eylon, *Science Learning and Instruction* (Lawrence Erlbaum, Mahwah, NJ, 2011).
4. J. Osborne, *Science* **328**, 463–466 (2010).
5. K. Ryo, M. C. Linn, *J. Res. Sci. Teach.* **51**, 147–174 (2014).
6. D. H. Uttal, D. I. Miller, N. S. Newcombe, *Curr. Dir. Psychol. Sci.* **22**, 367–373 (2013).
7. S. Wineburg, M. Smith, J. Breakstone, *Soc. Educ.* **76** (6), 288–291 (2012).
8. K. VanLehn, Ed. *Psychol. (Bern)* **46**, 197–221 (2011).
9. S. Sisk-Hilton, *Teaching and Learning In Public: Professional Development Through Shared Inquiry* (Teachers College Press, New York, 2009).
10. O. L. Liu, et al., *Educ. Measure. Issues Pract.* 10.1111/emip.12028 (2014).
11. R. D. Roscoe, D. S. McNamara, *J. Ed. Psychol.* 10.1037/a0032340 (2013).
12. A. C. Graesser, et al., *Behav. Res. Methods Instrum. Comput.* **36**, 180–192 (2004).
13. A. N. Rafferty, L. F. Gerard, K. McElhany, M. C. Linn, in *Proceedings of the Workshops at the 16th International Conference on Artificial Intelligence in Education*, vol. 8, E. Walker, C.-K. Looi, Eds., Memphis, TN, 13 Jul 2013 (AIEd, 2013), pp. 612–619.
14. L. M. Razzaq, N. T. Heffernan, in *Proceedings of the 2009 Artificial Intelligence in Education Conference*, V. Dimitrova, R. Mizoguchi, B. du Boulay, A. C. Graesser, Eds., Brighton, UK, 6 to 10 July 2009 (IOS Press, Amsterdam, 2009), pp. 457–464.
15. J. R. Segedy, J. S. Kinnebrew, G. Biswas, *Educ. Technol. Res. Dev.* **61**, 71–89 (2013).

Acknowledgments: This material is based on work supported by the NSF under grant no. DRL-1119670. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the NSF.

10.1126/science.1245980