JrnlID 11412_ArtID 9298_Proof# 1 - 18/04/2019

International Journal of Computer-Supported Collaborative Learning
https://doi.org/10.1007/s11412-019-09298-y

Guiding collaborative revision of science explanations

Q1

Q2/Q3 **Libby Gerard**[1] • **Ady Kidron**[1] • **Marcia C. Linn**[1]

Received: 26 July 2018 / Accepted: 16 April 2019

Abstract
This paper illustrates how the combination of teacher and computer guidance can strengthen collaborative revision and identifies opportunities for teacher guidance in a computer-supported collaborative learning environment. We took advantage of natural language processing tools embedded in an online, collaborative environment to automatically score student responses using human-designed knowledge integration rubrics. We used the automated explanation scores to assign adaptive guidance to the students and to provide real-time information to the teacher on students' learning. We study how one teacher customizes the automated guidance tools and incorporates it with her in-class monitoring system to guide 98 student pairs in meaningful revision of two science explanations embedded in an online plate tectonics unit. Our study draws on video and audio recordings of teacher-student interactions during instruction as well as on student responses to pretest, embedded and posttest assessments. The findings reveal five distinct strategies the teacher used to guide student pairs in collaborative revision. The teacher's strategies draw on the automated guidance to personalize guidance of student ideas. The teacher's guidance system supported all pairs to engage in two rounds of revision for the two explanations in the unit. Students made more substantial revisions on posttest than on pretest yet the percentage of students who engaged in revision overall remained small. Results can inform the design of teacher professional development for guiding student pairs in collaborative revision in a computer-supported environment.

Keywords Technology · Knowledge integration · Automated scoring · Adaptive guidance · Assessment · Teaching Q5

✉ Libby Gerard
libbygerard@berkeley.edu; libby.gerard@gmail.com

Ady Kidron
ady.kidron@berkeley.edu

Marcia C. Linn
mclinn@berkeley.edu

Q4 [1] Graduate School of Education, University of California, Berkeley, CA, USA

Introduction

Q6

Computer-supported learning environments featuring powerful scientific models can offer students multiple opportunities to engage in meaningful, collaborative revision of explanations. Revision of scientific explanations is central to doing and learning science. Revision is a vital and ubiquitous practice in science careers, science learning, technical occupations, and scientific writing (Brownell et al. 2013; Perin et al. 2016; Thagard 1992). Many scientists view their work as generating, testing, and revising their ideas (Isaacson 2017; Feynman et al. 1985). Researchers have characterized students' meaningful engagement in revision as using evidence to distinguish among alternative viewpoints, and as clarifying the mechanistic explanation for an audience (Berland et al. 2016). Working with a partner in a computer-supported learning environment may encourage student use of these revision processes as the pair works toward a shared understanding. For example, each student may offer an alternative viewpoint to widen the pool of ideas for consideration (Matuk and Linn 2015).

Guidance can help students engage in these practices as they work in a computer-supported learning environment. Research shows that even when prompted, few student pairs work collaboratively to make meaningful revisions to their science explanations or models (Tansomboon et al. 2017; Sun et al. 2016; Sinha et al. 2015; Zheng et al. 2015). Rather, students are more likely to make superficial changes, paraphrase their initial view, or, add new but disconnected ideas (Crawford et al. 2008; Gerard et al. 2016; Sinha et al. 2015; Zheng et al. 2015). A recent meta-analysis found that teacher guidance had no significant impact on students' collaborative learning outcomes (Chen et al. 2018).

In this research we collaborate with a teacher to investigate how to customize guidance by taking advantage of automated explanation scoring to improve students' collaborative revision process. The results offer concrete strategies teachers can use to effectively guide collaborative revision in a computer-supported learning environment. They reveal how partnering with a teacher to customize the learning environment tools prior to implementation can impact learning.

Collaboration and revision: opportunities and challenges

Collaborating with a partner has the possibility of engaging students in the behaviors characteristic of meaningful revision. Berland et al. (2016) showed how peers can serve as an audience for one another, encouraging each other to clarify their explanations (see also Cohen and Riel 1989). A partner can add a wider repertoire of ideas to the mix for consideration, as well as articulate an idea using vocabulary that is accessible to their peer (Songer 1996). Making a wider repertoire of ideas visible may push students to attend to complexity in their explanation that they might otherwise overlook (Reiser 2012). Harrison et al. (2018) demonstrated that student pairs who critiqued another group's response and then revised their explanation, made greater revision gains than student pairs who revisited evidence in the unit and responded to questions prior to revising their explanation. For example, critique of a peer's response led more students to distinguish between phenotype and genotype in their explanations and connect these ideas to a mechanistic explanation.

Major challenges students face in revision include confirmation bias and a focus on completion and correctness over refinement. Students often ignore contrasting evidence presented by a peer and restate their own perspective (Clark and Chase 1972; Berland and

Reiser 2011), or strengthen and reiterate their initial view rather than revising their perspective (Mercier and Sperber 2011). Likewise, in our prior work, a majority of student pairs added disconnected ideas to their explanations when using automated guidance to revise - often in an attempt to answer the hint in the guidance rather than reconcile ideas suggested by the hint with their initial views (Gerard et al. 2016; Harrison et al. 2018). A recent study of student revision when using collaborative Google Docs found that peers rarely recognized gaps in one another's reasoning. Rather the majority of peer feedback given to one another's essays focused exclusively on the writing mechanics as opposed to the content or argument structure (Zheng et al. 2015).

In addition students often focus on "getting the lesson done" as opposed to meaningfully engaging in the science practices integral to collaborative revision (Jimenenez-Aleixandre et al. 2000). For example Sun et al. (2016) designed a CSCL environment to teach diffusion and osmosis. One activity was designed to engage student pairs in using evidence to revise their initial models. Analysis of online and face-to-face discussions, revealed that task-oriented talk such as clarifying procedures or work division took up the highest average proportion (43%) of peer discussions. Assessment-oriented talk, or providing constructive comments on peer's initial models of osmosis, took up the least amount of peer-talk (13%). Students were primarily concerned with specifying procedures and managing the division of labor to complete the tasks as opposed to focusing on the use of evidence to refine their explanations.

Leveraging CSCL features to support revision

Researchers have documented features of CSCL environments that can be drawn upon to strengthen guidance for collaborative revision (Chen et al. 2018). Matuk and Linn (2015) found, for example, that students benefited from guidance in an online class discussion that prompted them to seek an idea that differed from their own rather than selecting ideas congruent with their own ideas. When students intentionally selected ideas that differed from their own they wrote more coherent and normative explanations within the unit and on a posttest compared to students who selected congruent ideas. Ryoo et al. (2018) found that the frequency of knowledge-oriented peer collaboration was greatest when student pairs were guided to investigate an interactive, dynamic visualization compared to other activities within the unit. The visualizations widened the repertoire of available ideas for students' negotiation and provided a shared language for forming mechanistic explanations. The benefits of dynamic visualizations and online discussions for collaborative learning depend on how the teacher motivates and supports students' interactions.

In this case study we examine how one teacher customizes automated explanation guidance and her in-class guidance strategies to help student pairs revise their explanations. Natural language processing models are used to automatically diagnose student pairs' written explanations about convection, plate movement, and geological landforms embedded within a Web-Based Inquiry Science Environment (WISE) investigation (Liu et al. 2016; Vitale et al. 2015). The automated explanation scores are used to assign adaptive guidance to the student pairs in real-time. The adaptive guidance, designed based on the knowledge integration framework, prompts the student to consider an idea that was missing or inaccurate in their response and suggests a (linked) dynamic visualization from earlier in the unit for the student to revisit in order to strengthen their understanding (Gerard et al. 2015). The teacher in this study reviewed the automated scoring rubrics and customized the automated guidance to align with her teaching strategies, prior to implementation. She also created an in-class monitoring system to take advantage of the automated scores and guidance in supporting students' collaborative revision. We used teacher interviews, classroom audio and video recording, and logged data to capture the teacher customization of instruction, guidance strategies, and student revision processes. We used pretest, embedded assessment, and posttest data to document how the teacher's customized guidance influenced students' disciplinary learning, frequency of revision, and revision quality.

Knowledge integration and guidance

Science investigations call for students to posit predictions and questions and investigate those by exploring forms of evidence. Students often add new ideas, based on their review of the evidence, to their multiple and in many cases already conflicting views. As a result, student's ideas remain disconnected and isolated (diSessa 2006). Thus, in collaborative situations, instruction that emphasizes integration of diverse ideas has value (Furberg 2016; Matuk and Linn 2018). Guidance in science instruction could strengthen the process of knowledge integration by broadening the pool of ideas, helping students use evidence to distinguish among viewpoints and consolidate ideas into a coherent explanation (Williams et al. 2004). Furberg (2016) found, for example, that even though students worked in a well-scaffolded, computer-supported collaborative learning environment, they needed substantial teacher guidance to link results from the lab experiment with the mechanistic science ideas. This finding was extended by Ingulfsen et al. (2018) who documented the considerable teacher guidance needed to support student dyads in connecting evidence from real-time digital graphs with underlying science principles.

In a series of studies informed by the knowledge integration framework, teachers elicited students' reasoning about the topic, probed further with questions that built on or challenged the students' ideas, and then used the students' ideas to customize their guidance for next steps (Gerard and Linn 2016; Linn and Eylon 2011; Zertuche et al. 2012). The teachers personalized the guidance depending on students' ideas and level of understanding, even while maintaining overall class progress. Guidance that encourages students to make connections between their prior knowledge and the new ideas presented by their peers or instruction results in significantly greater student learning gains than procedural guidance (e.g. reread the instructions), or guidance pointing out incorrect ideas and supplying the correct information (Williams et al. 2004; Ruiz-Primo and Furtak 2007).

Most teachers are challenged to provide personalized guidance for all students during instruction. This is due in part to large class sizes. Middle school teachers often have five or six classes of 30 to 40 students each. Further, teachers may also lack experience with the wide range of student ideas they are likely to encounter in investigation of a science topic (Lakkala et al. 2005). Ruiz-Primo and Furtak (2007) conducted a study of four teachers' formative assessment strategies during a science inquiry unit. The majority (71%) of the teachers' assessment conversations did not draw on students' ideas to adapt guidance. Rather the conversations involved eliciting students' ideas, a student response, and teacher recognition of the students' viewpoint. This often meant rephrasing the student's response or providing an evaluative response. In some cases, the teachers only elicited students' ideas. A very small percentage of the teachers' guidance involved asking students to relate evidence to explanations, evaluate the quality of evidence, or to compare and contrast others' ideas. Of the questions asked in the conceptual domain, the most common were those that asked students for definitions. Another study investigated teachers' written comments for elementary and middle school students' scientific work. The vast majority of the comments given were grades or a numerical evaluation (61%); only 33% contained conceptual related comments (Ruiz-Primo and Li 2013). Collaboration is most successful when teachers encourage student teams to explain and sort out their ideas and is often undermined when guidance gives students the right answer (Hamalainen and Vahasantanen 2011).

Taking advantage of automated explanation scoring and adaptive guidance

Researchers in the computer-supported collaborative learning field have called for the use of technologies to provide teachers with real-time information on student learning that can inform the teacher's pedagogical moves (Sharples 2013). There is agreement in the field that supporting teacher agency in using automated assessment tools is paramount to making these tools successful (Roshcelle et al. 2013). This means the tools need to be flexible so the teacher can adapt them to their goals, and modify the tools in real-time to respond to unpredictable classroom events. Yet there is limited empirical work on how teachers use automated student response information to adapt instruction. Earlier work examined teachers' use of "clickers" in large, post-secondary courses. Students responded to multiple choice questions during a lesson, and the responses were aggregated and displayed to the teacher in real-time. Research identified value of clickers for providing teachers with insights into students' range of ideas, and particularly students' alternative views about the topic. The auto-scored assessments however, did not provide teachers insights into student's explanations or the reasoning underlying their multiple-choice selection. In such, the aggregated information often encouraged teachers to provide direct instruction about a commonly held idea, rather than guide students to gather evidence to investigate their views.

Tissenbaum et al. (2012) provide empirical work on using aggregated student responses for physics problems to help the teacher and students guide inquiry in real-time. They created a classroom, wall display of student progress in solving the physics problems and created a teacher report. They also provided the teacher a hand-held device during one design iteration. They observed the teacher use the wall display while circling the classroom to identify groups with which to intervene, and to jumpstart his conversation with a group on how to refine or elaborate their explanations. The teacher used the student data report to modify his lesson in between days teaching. Somewhat surprisingly, the teacher found the hand-held device distracting and stopped using it after a short time. Students in classes where the teacher had the wall display made greater learning gains than in classes where the teacher did not have the display as it supported the teacher to engage in quick and meaningful interactions with pairs. The findings suggest promise for flexible automated scoring tools that make student's reasoning accessible to the teacher in real-time.

As evidenced in the Tissenbaum et al. (2012) study, automated scoring and adaptive guidance technologies may support teachers to provide personalized guidance during instruction that promotes student pairs to engage in knowledge integration processes as they revise their explanations. The automated guidance in this study resulted from researcher analysis of over 1000 student responses from multiple teachers. To determine effective guidance, the research team distinguished the key student ideas at each level of the knowledge integration rubric. Then the team designed and tested this guidance to be sure that it is effective. Thus the computer guidance is based on substantial expertise about likely student responses. And the computer guidance has been refined based on review of how well it works for multiple students (Gerard et al. 2015). Teachers do develop this form of expertise from interacting with their students. However, they must build it up over time (Sisk-Hilton 2009). Computer guidance may give teachers a head start by modeling some approaches that have worked in the past.

The automated scoring technologies in this study are used to provide the teacher a quick diagnosis of a student pair's joint understanding as well as a hint to help the teacher target her questions in eliciting each of the individual student's views. Additionally the automated guidance may help the teacher identify relevant evidence in the unit for students to review. The teacher can direct students to use this evidence to sort out their views rather than providing the missing information. Knowing where students could find and analyze relevant evidence supports the teacher to promote knowledge integration during revision by encouraging students to distinguish their ideas from those presented in the unit.

Methods

We conducted a case study of a sixth-grade teacher to explore: How does a teacher customize instruction using a learning environment that includes automated explanation scoring and adaptive guidance to guide students in collaborative revision of explanations? Video and audio recorded class observations and logged data provide insights into how the teacher adapts her guidance to support revision for each pair. Embedded and pre/post assessments demonstrate the impact of the teacher guidance on students' explanation revisions and knowledge integration.

Curriculum: WISE plate tectonics

This research used the Web Based Inquiry Science Environment unit "Plate Tectonics: What Causes Mountains, Earthquakes and Volcanoes?" (http://wise.berkeley.edu/project/18661#/vle/) to investigate how automated scoring of student written explanations can strengthen teacher guidance. WISE is an online authoring and instructional delivery system. The units target topic areas that are aligned with state (CA) and national science standards (NGSS) and that benefit from dynamic visualizations. Topics are those that research has demonstrated are challenging to teach, hard to illustrate with static pictures, and difficult to explore with laboratory experiments (Donnelly et al. 2014). The units and assessments are designed following knowledge integration design principles (Kali et al. 2008) and are collaboratively used, typically in groups of 2–3 students. Extensive research demonstrates significantly greater knowledge integration on target science concepts when student teams use WISE units than when they learn through traditional textbook instruction (e.g., Clark and Sampson 2008; Donnelly et al. 2014; Raes et al. 2013). Students typically study each WISE unit, led by their regular classroom teacher, for 6–8 class periods (50 min each).

The Plate Tectonics unit engages students in exploration of a complex problem and includes features designed to promote knowledge integration as students explore this problem. Students investigate why are there more mountains, earthquakes, and volcanoes on the West Coast (where this study takes place) than on the East Coast of the United States. It addresses the

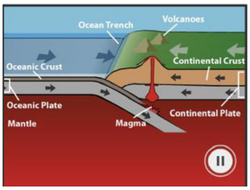
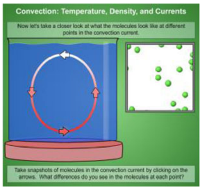
NGSS performance expectations MS-ESS2–2 and MS-ESS2–3. Students work in pairs, using one shared computer, throughout the unit.

The unit elicits the pair's ideas by guiding students to explore maps of earthquakes, mountains, and volcanoes in the United States and within California specifically. Students make observations about where these events occur, and articulate their ideas about why the events may be clustered in such a way. Pairs collaboratively add ideas about the plate tectonics processes inside the Earth by viewing dynamic visualizations of plate boundaries, magma convection currents, and resulting geological features (see Table 1). To help student pairs distinguish and sort out ideas, the students use matching steps to categorize the features (density, mass) of the different plate types. Student pairs then annotate images of Earth's interior and interpret graphs to distinguish the relationship between magma and temperature relative to surrounding material, and the proximity of magma to Earth's core. The unit helps student to make connections among ideas by collaboratively generating explanations (Table 1) that encourage student pairs to sort through and make connections among their interpretations of evidence gathered from across the unit, to explain the entire geological process.

Automated scoring of explanations

We developed natural language processing models for two select explanation prompts embedded in the Plate Tectonics unit to diagnose student pair's knowledge integration and assign adaptive guidance (Table 1). The first question, "Mountain", calls for students to connect ideas about plate type and density, plate interactions, and the resulting geological landforms. The second question, "LavaLamp", asks students to link ideas about density, temperature, and movement, to explain how a lava lamp works and how this is similar to what is happening inside of the Earth.

Table 1 The two automatically scored explanations embedded in plate tectonics WISE Unit

	explanation 1 - Mountain	**explanation 2 – Lava Lamp**
Explanation Prompts	The diagram shows a cross section of the edge of a continent. There is a section of oceanic crust and of continental crust. Both are gradually moving towards each other. Explain in detail how the mountain range near the seacoast on this continent was probably formed mountain.	Lava lamps are special lamps full of fluid. Every so often, a blob of colored fluid will go up to the top of the lamp, then go back down again. How do you think lava lamps work? Using what you know about HEAT and DENSITY, explain how you think lava lamps work.
Disciplinary Context	Explain how the density of Earth's plates affects their interaction and the resulting landform	Contrast the upward and downward movement of a blob in a lava lamp due to the changes in temperature and density to explain convection.
Sample Visualizations for Gathering Evidence		

Automated scoring

We integrated c-raterML™, a natural language processing tool developed by the Educational Testing Service, to score the explanations in the WISE unit (Liu et al. 2016). The c-raterML™ system scores each student explanation based on a 5-point knowledge integration rubric that rewards students for using evidence to make links among scientifically normative ideas (example in Table 2). It works by building a model of the linguistic features evident in student responses at each knowledge integration score level, based on the analysis of the human scoring of at least 1000 student generated responses to the same question. Both c-raterML™ scoring kappa models demonstrated satisfactory human-machine agreement using knowledge integration scoring rubrics (Kappa: *Mountain* k = .75, *LavaLamp* k = .81).

Adaptive knowledge integration guidance

After a student explanation is scored by c-raterML™, WISE instantaneously assigns the pair automated, adaptive knowledge integration guidance based on the score level (Fig. 1). The guidance is designed to help students move up one level in the scoring rubric. The guidance for each score level includes three parts, each addressing a key knowledge integration process (Gerard et al. 2015).

- *Add ideas*: Ask a question about the key missing or non-normative concept in the student's response
- *Distinguish ideas*: Direct student to revisit evidence in a relevant part of the unit illustrating the missing or non-normative concept in the student's response
- *Integrate ideas*: Ask students to use the evidence they've gathered to generate an improved response

In this study, all students were able to receive two rounds of automated, adaptive knowledge integration guidance. We also incorporated automated teacher alerts based on the explanation score (alerts further described in Gerard and Linn 2016). For the second revision, students who scored at or below a threshold (set by the teacher) received the following teacher alert:

> *StudentName, TALK TO YOUR TEACHER to help you take your answer further. The top bar of your screen is now red so your teacher knows to come talk with you. This animation may help you and your teacher discuss. You can move on in the project until your teacher comes over to help. Score: 2.*

Students who scored above the set threshold received a second round of adaptive knowledge integration guidance. The algorithm for assigning guidance based on the knowledge integration score was arranged to assign a unique second round of knowledge integration guidance, even if the student's assigned score did not change. This ensured that student pairs who revised (but did not improve or decrease in score) did not receive the same guidance twice.

Participants

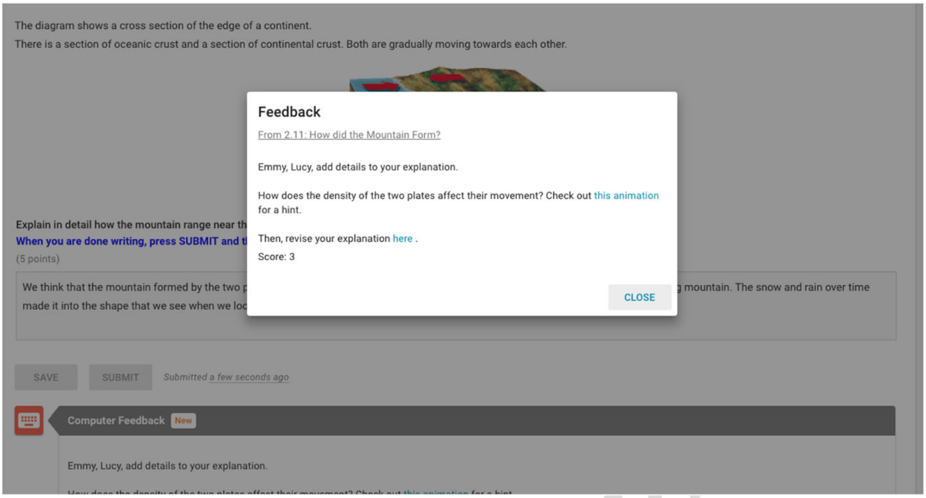
One sixth grade teacher in a public middle school and her 201 students participated in this study [98 student pairs]. Students were distributed across six class periods. The

JrnlID 11412_ArtID 9298_Proof# 1 - 18/04/2019

t2.1 **Table 2** Knowledge integration scoring rubric and automated, adaptive guidance for Lava Lamp embedded explanation

t2.2 ***Lava Lamp explanation - Key Ideas*:**
Ideas about heat and density
Ideas about heat and molecular movement
Ideas about heat and movement
Ideas about density and movement

	KI Score	Level	Example Student Response	Automated KI Guidance
t2.4	1. Off Task	Student writes but it does not answer the question being asked.	IDK	
t2.5	2. No links	Only alternative or vague idea(s) stated. Linked normative and non-normative ideas. Repeats question.	When you make a lava lamp it has oil, and the oil doesn't mix to the chemicals inside the lamp, so the blob just moves around which is the oil. The chemicals are getting pressure from the heat and makes movement and reacts, the heat makes the blob go up and down, the pressure is the density.	<Student names>, how does the temperature of the blob affect its movement? Check out <here> for a hint. Then, redo your explanation.
t2.6	3. Partial	Idea(s) within one key idea category. Ideas in multiple categories but isolated.	A lava lamp works because when it is cold the stuff in side is a soiled and when the **lava lamp gets hot the stuff inside goes to the top**. **The density decreases when it gets to the top**.	<Student names>, when does the density of the blob decrease? Check out <here> for a hint. Then, revise your explanation
t2.7	4. One link	Links two ideas. Links two ideas in at least 3 key idea categories for one direction (up or down).	Well, when objects heat up the rise, so that may be the reasoning behind how the liquid floats to the top, and once it cools down it floats back down. **When the blob heats up, it becomes less dense and floats to the top**, then when it cools down it thickens, causing it to sink.	<Student names>, when does the density of the blob increase? Check <here> to get more information. Then, expand your explanation..
t2.8	5. Complex Links	Two or more links. Links ideas in at least three categories for one direction (up or down), with ideas in at least two categories for the other direction.	The blobs change density and move from top to bottom because it becomes more dense than less dense. **When the blob is at the bottom it heats up and becomes less dense, so it rises to the top of the lamp where it is not as warm and becomes more dense and falls back to the bottom**. When this process repeats it makes the blobs fall and rise frequently. Creating the illusion of a lava lamp.	<Student names>, use your ideas about heat and density to elaborate why the blob goes all the way up, and, all the way down. Check *<here>* to get more information. Then, expand your explanation.



Q8 **Fig. 1** Example of adaptive, knowledge integration guidance for a student-team written embedded Mountain explanation

teacher used the same general instructional approach with each class period. The students are from diverse backgrounds and are distributed across six class periods depending on their overall schedule. Forty-seven percent of the participating students report that their parents speak a language other than English at home; 10% of students are labeled English Learners. The school population is 53% Non-White and 34% receive a free/reduced price lunch.

Students worked in pairs assigned by the teacher while studying the unit. The teacher assigned pairs based on who she thought would work well together taking into account multiple factors including each student's academic focus, work habits, friendships, and performance in the science class. Each student completed the pretest and posttest individually.

The teacher has used WISE over the past two years to teach Global Climate Change and Solar Ovens. This was the teacher's first time teaching the WISE Plate Tectonics unit and her first time using a WISE unit with automatically scored embedded explanations.

Teacher customization of the automated guidance

To support the teacher in guiding students' collaborative revision, we partnered with the teacher to customize the automated guidance system for her classroom use. Prior to implementing the unit, the teacher reviewed the full Plate Tectonics unit. This was followed by a two-hour meeting in which the teacher and researcher reviewed the knowledge integration rubrics used to automatically score the Mountain and Lava Lamp explanations embedded in the unit. The rubrics included sample student explanations for each level of the rubric and the assigned knowledge integration guidance for each level (Table 2). The teacher reviewed the visualizations within the unit that the automated guidance directed students to revisit. She tested the automated scoring technology by generating explanations that included the ideas she anticipated her students would express. The teacher also reviewed the WISE teacher grading and commenting tool, learning how to assign comments and identify students who received a teacher alert during class time. After reviewing the unit and explanation revision activities, the researcher collaborated with the teacher to customize the automated scoring tools to support her guidance. The teacher also customized her in-class guidance strategy to monitor student progress in revision during instruction.

The teacher reflected aloud on her guidance, the automated guidance, and students' learning as she reviewed each pair's essays and their revisions in the grading tool. She did this after class on several days during implementation of the unit (researcher audio-recorded). After student pairs had completed two essay revisions, the teacher wrote comments to each student team (Fig. 2). The comment included a final score and grade for each student pair's essay. Since this step of the teacher guidance was not performed in real-time, we excluded the data analysis in this paper.

Data collection and analysis

Teacher guidance To capture the teacher's guidance for collaborative revision we audio and video recorded the teacher-student conversations with student pairs as she guided them in revision of the short essays. We collected 37 recordings of the teacher interacting with thirteen student pairs, as students were using the automated guidance to revise their essays. The student pairs were selected for recording based on completion of the student assent and parental consent forms. The thirteen pairs included all pairs in which each of the two students in the pair returned their audio/video parental consent and student assent forms. We recorded all teacher interactions with these 13 student pairs when the pairs were working on the Mountain or Lava Lamp explanation writing and revision in the WISE Plate Tectonics unit. The pairs were distributed across the teacher's six class periods. All audio and video recordings were transcribed including the teacher's guidance statements and the student responses.

We developed a coding scheme that was informed both by our inductive analysis of the teacher guidance, and the knowledge integration framework on learning. To develop the

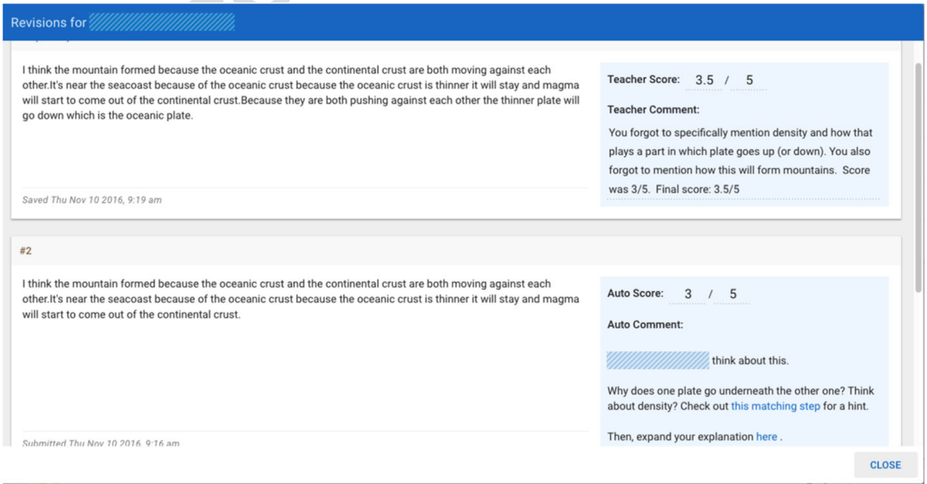


Fig. 2 The teacher's interface for viewing and assigning guidance and student revisions in response to guidance. The bottom bar shows a student pair's response with their assigned automated score and guidance on the right. The top shows the same pair's revision in response to the automated guidance and the teacher's guidance for their revision and an updated score

coding scheme, one researcher read through the teacher-pair transcripts multiple times noting different types of teacher guidance moves. The researcher then formed initial categories and reviewed the categories, with criteria and examples, with the authors of this study and two outside researchers. Together the team reformulated the categories to better capture the intent of each teacher guidance strategy in the context of students' collaborative revision process. After several meetings, a set of agreed upon categories with criteria and examples was established.

To ensure reliability of the coding using these categories, a research team of seven people (including five researchers who were not directly involved in the study and two of the authors of this paper) coded 33 teacher guidance strategies, or 19% of the full teacher guidance data set. The 33 teacher guidance statements were a part of her conversation with three different student pairs. The team used each teacher guidance strategy during a teacher-student pair interaction as the grain size for coding. A strategy consisted of one to two teacher statements, and focused on eliciting one kind of collaborative revision action on the part of the students. Each teacher guidance strategy was coded for only one category. The location of the teacher guidance strategy within the teacher-pair interaction was considered when coding to determine the intention of the teacher guidance strategy. Researchers worked in three pairs and one individually to independently code the teacher guidance data set [see sample coded teacher strategies in Tables 8 and 9]. We then compared codes, identified disagreements, and discussed disagreements until reaching consensus. To determine consensus, the team revisited the context of the guidance strategy within the teacher-pair interaction, how the teacher's use of the guidance strategy related to the teacher's surrounding guidance moves in the interaction, and the characterization of each guidance strategy within the coding rubric.

Of the 33 teacher guidance strategies coded, the four independent coders (three pairs and one individual) reached 76% agreement, disagreeing initially on codes for eight teacher guidance strategy moves. For those eight guidance strategies, coders were deciding between one of two codes. After distinguishing which code captured the essence of the teacher guidance statement, we refined and elaborated the coding rubric to reflect the criteria raised in our discussion for each category, as shown in Table 3. One researcher then coded all of the data, consisting of 171 teacher guidance strategies, using the updated rubric.

To calculate the teacher's frequency of use of each guidance strategy, we counted the number of times the teacher used the strategy across the data set. We then computed the frequency as a percentage of the whole.

To investigate how the teacher adapted her guidance strategies to support pairs at varied levels of understanding, we examined the teacher's strategies for student pairs who demonstrated different levels of understanding on their initial Mountain or Lava Lamp explanation. For this analysis, we divided the 13 audio-recorded pairs into those who demonstrated vague or correct but disconnected ideas on their initial embedded essay (KI score of 1, 2, 3),and those who expressed at least one link between two accurate ideas (KI score 4 or 5). We computed the frequency with which the teacher used each of the guidance strategies described in Table 3, during her interaction with low/partial versus high pairs.

Field notes We gathered detailed field notes while in the classroom for four of the seven class periods for each day of unit implementation. These were used to supplement interpretation of the audio and video files.

t3.1 **Table 3** Rubric for coding teacher strategies for guiding collaborative revision

	Strategy	Examples of teacher guidance and student response
t3.2	Strategy	Examples of teacher guidance and student response
t3.3	**Establish a shared understanding of progress** Ask students to read their computer guidance, and/ or response aloud, or to check where they are in the process or revising based on the guidance	Teacher: So you just got your first round of feedback. What was your score? Student 1: Four out of five. Teacher: Ok, can you read what the computer suggested you do: Student 1: Manuel, Rane, good reasoning. Now, think about this. Could any other type of landform develop at this boundary - why or why not? Check out for a hint. Then, expand your explanation
t3.4	**Ask students to assess their progress in revision and determine next step** Prompt students to reflect on the quality of their explanation, or to evaluate if they responded to the hint in the guidance, and decide the next step	Teacher: Do you feel like you've answered every part of their question? Student 1: Ya Teacher: Okay Student 2, go ahead and submit. Read the guidance aloud. Student 2: Submits response and reads guidance aloud Teacher: Do you guys think you did that (referring to automated guidance)? Student 2: Um <quiet>
t3.5	**Elicit details about each student's perspective about a specific concept** Surface the range of student ideas about a specific concept targeted by the automated guidance	Teacher: Ok, why? What does subduct mean? Stu2: sink Stu1: It means to go down Teacher: Sink or go down, ok. Why does it sink or go down?
t3.6	**Recommend students use a revision strategy** Prompt students to use a strategy to elaborate or reconcile their two ideas such as revisit animation, or clarify what elaborate means	Teacher: You put, the plate that is less dense sinks. Are you sure less dense goes down? Stu 1: Mm Teacher: Oh you know what we should do, that would be good to check out the animation. Stu 2: More dense Teacher: More dense, it would be good to still check the animation and then you can fix your answer.
t3.7	**Suggest a new idea to consider** Present a new idea to the pair to extend the students thinking about a specific concept	Stu 1: A [Lava Lamp] is similar to how there are convection currents inside the mantle. Teacher: Well in general, to have convection you need a heat source.

Interviews We conducted and audio-recorded two teacher interviews. One was conducted as the teacher customized the automated scoring system, and one was conducted after the teacher reviewed the student pair's second round of revisions in the grading tool. The interviews captured the teacher's customization decisions and her reflections on how she and the computer supported student pairs to collaboratively revise their explanations.

Student knowledge integration and revision

We documented how the teacher's guidance for collaborative revision influenced students, both within a pair and as an individual, to integrate ideas in plate tectonics and use guidance to revise explanations using logged data, pretest, embedded assessments, and posttest data.

Embedded assessments All students were prompted to write an initial response to each explanation prompt, and had two opportunities to revise. We used students' initial and final revisions on both the Mountain and Lava Lamp explanations to measure learning gains. We used all of the students' logged explanation revisions to examine their revisions relative to teacher and computer guidance. The log files [csv files] enabled researchers to distinguish each revision time point relative to the student pair's interaction with the automated or teacher guidance.

Pretest/posttest revision item The assessment item (Table 4), which was the same for the pretest and posttest, was designed to measure students' knowledge integration and student ability to use guidance to revise. The item calls for students to integrate multiple ideas taught in the respective units into a coherent explanation. Studies show that questions designed to measure knowledge integration validly assess students' conceptual understanding (Liu et al. 2008; Liu et al. 2011). Each student responded to the pre and posttest item individually.

The item included one round of real-time automated guidance, giving the students the opportunity to use the guidance to revise their initial response. The guidance was more general than the guidance given during instruction to measure student ability to transfer what they had learned from revising in the unit to revision on the posttest. This novel item format captured both the students' ability to use guidance to revise.

Students' initial and revised posttest explanations were used in the analysis to capture students' disciplinary learning from the unit. Students' revision gains on the pretest compared to their revision gains on the posttest were used to capture students' learning of how to revise.

t4.1 **Table 4** The Pre/Post assessment item: mount hood, plate tectonics unit

	Prompt		Sample Guidance
t4.3		This is Mount Hood. It is a part of the mountain range called the Cascades on the West Coast in Oregon. Write a story to explain how the mountain formed. Be sure to describe what happens inside of the Earth and on the outside. After you are done writing, press "Check Answer". You will have 1 chance to get feedback and revise your story	[KI Score = 3] Sara & Mario, expand your story. Think about: What is happening inside Earth's mantle?

All student responses to the embedded assessments and the pre/post item were scored by both the c-raterML system and a human scorer. In this study we reconciled c-raterML scores with human scores. To resolve rare disagreements, the researcher reviewed other student responses to locate similar answers. Then the researcher assigned the response to the category with greatest similarity. Both the c-raterML system and the human scorer used the 5-point knowledge integration rubrics (example in Table 2). Coding rubrics for the c-raterML system were established in prior research (see Liu et al. 2015). The c-raterML scoring models for these items were validated in comparison to human scores (see Liu et al. 2016).

To examine how the teacher's guidance for pairs with differing levels of initial understanding may have influenced the individual student pre to posttest gains, we compared the individual pre/posttest scores for students who were in pairs that had demonstrated linked understanding (score 4,5) on their initial embedded essay, to the pre/posttest scores of students who were in pairs that had expressed disconnected or vague ideas (score 1, 2,3) on their initial embedded essay. We used student pair's initial essay score on the first embedded essay (Mountain) as the prior knowledge indicator. Since students were working in pairs, we assigned each student in the pair the same initial score.

Results

Student knowledge integration and revision

We analyzed how student pairs revised their explanations during instruction as well as how each student responded individually to the unit on pretests and posttests. The embedded assessment outcomes reflect most directly the teacher customization of guidance strategies before and during instruction to support each student in the pair to engage in collaborative revision of their explanations. The student's pre to posttest improvement in explaining plate tectonics can be attributed to the entire unit including the activities, computer and teacher guidance.

Embedded explanations: Collaborative revision The combination of teacher and computer guidance supported student pairs to improve the quality of their explanations during instruction. It also increased the frequency of student revision relative to prior studies. Overall the student pairs significantly improved the coherence and accuracy of their explanations (see Table 5). Initially, the collaborators had reasonably sophisticated responses, as reflected in mean scores above 3. A score of three shows that the response had one idea relevant to the question and that the idea was not linked to evidence. The

Table 5 Collaborative embedded explanation: knowledge integration scores for student pairs

Explanation	N Pairs	Frequency of revision	KI Score 1	KI Score Final	Improvement
Mountain	98	99%	3.36(.80)	4.30(.81)	.94(.81)**
Lava Lamp	96	100%	3.53(.82)	3.91(.67)	.38(.93)*

**Mountain t(97) = 11.47, $p < .0001$

*LavaLamp t(95) = 3.94, $p < .001$

students worked together to fill gaps in their explanation and to modify inaccurate ideas. Their revised explanations had a mean close to or above four, indicating that through the revision process, the student pairs added a link to evidence. All but one of the collaborating pairs made a revision to their explanation. The participation of every collaborating pair in revision provides strong evidence for the value of the combination of teacher and automated guidance. Prior research shows that even when prompted, a minority of students make substantive revisions to their explanations in science class (Gerard et al. 2016; Sinha et al. 2015; Sun et al. 2016; Tansomboon et al. 2017).

Pre and post assessment In addition to improving their collaborative explanations, students also made significant individual pretest-posttest gains. The pre and posttest included an item where students wrote an explanation, submitted their explanation, received general guidance, and had the opportunity to revise.

Students made significant improvement from pretest to posttest demonstrating that students gained robust knowledge of plate tectonics, as shown in Table 6 (Pre to Posttest Gain $M = .92$, $SD = .13$, $t(191) = 14.62$, $p < .0001$). Specifically, on the pretest, the mean revised scores were around 2.3, indicating that most students held vague or unsubstantiated ideas about how mountains form. By the revised posttest, the mean score was around 3.2, indicating that, on average, individual students had one relevant idea.

We examined the individual pre/post test scores of students from pairs who expressed vague or disconnected ideas on their initial embedded explanation, to the pre/posttest scores from students who were in pairs that had expressed linked understanding on their initial embedded explanation. This gives insights into how the teacher's different guidance approach, for pairs who expressed differing levels of initial understanding, influenced the individual student learning. These two groups started the unit with similar pretest scores (Pretest score: vague/disconnected $n = 115$, m(sd) = 2.24(.51); linked $n = 78$, m(sd) = 2.32(.57)). Interestingly, the students from pairs that demonstrated one link on the embedded explanation made significantly greater pre to posttest gains, than students from pairs that demonstrated vague or disconnected ideas on the embedded explanation (Pre-Post gain, vague/disconnected m(sd) = .71(.83); linked m(sd) = 1.23(.85), $t(190) = 4.23$, $p = .000$). Likewise, the students from pairs that had expressed vague or disconnected ideas on the embedded explanation made smaller posttest revision gains (Posttest Revision gain, vague/disconnected m(sd) = .17(.39); linked m(sd) = .30(.49), $t(191) = 2.14$, $p = .04$). This suggests that the teacher guidance for pairs who expressed a linked understanding on their initial essay may have supported them to add and integrate new ideas during revision, whereas for the pairs who started with vague or disconnected ideas, the guidance may have supported students to add ideas but not necessarily integrate new ideas.

On the posttest, we assessed students' ability to use guidance to revise, without teacher or peer assistance. All students revised on the pre and posttest, due in part to a constraint in the

Table 6 Individual Pre/Post test Mt. Hood: knowledge integration scores for individual students

Pretest $N = 193$ individual students			Posttest N = 193		
Initial	Revised	Number of students who improved KI score in revision	Initial	Revised	Number of students who improved KI score in revision
2.18(.47)	2.27(.53)	**19 (10%)**	2.96(.87)	3.20(.95)	**43 (22%)**

learning environment that required students to make a change to their explanation after receiving guidance before they could advance. Students were more likely to improve their explanation by a full knowledge integration level when using the guidance to revise on the posttest, than on the pretest. Nineteen out of 193 students, or 10%, improved by a knowledge integration score when revising their explanation on the pretest. Fourteen of those nineteen moved from a knowledge integration score of a two to a three, meaning that they added a valid idea to their explanation. In comparison, forty-three out of 193 students, or 22%, improved by a knowledge integration score when revising their explanation on the posttest. Of those forty-three students, thirty added *and integrated* a new, valid idea in their explanation (12 moved from a score of a 2 to 3 adding an idea; 19 moved from a 3 to 4, and 11 went from a 4 to a 5, all adding and integrating an idea). The greater number of students who improved their explanations on the posttest is likely due in part to students holding a wider range of ideas to draw upon from instruction, and an improved ability to connect those ideas with the ones expressed in their initial response from guided revision. The individual gains are consistent with the collaborative gains during instruction, suggesting that both members of the pair benefitted from collaborative revision.

To explore this further, we analyzed a subset of student pairs' individual posttest responses to investigate how the two students from a pair performed, after the shared revision experience during instruction. All pairs had improved on the embedded revision activity during instruction. The results suggest that while the students who worked in the pair demonstrated active contributions to the revision activity during instruction and improved their responses substantially as a pair, the students individually integrated different insights from the experience. We focused on the same subset of thirteen pairs who were audio-recorded. Of the thirteen pairs, 10 pairs included partners who generated responses receiving the same knowledge integration score (average posttest score 3.5) or scores that differed by one point. Three pairs included two partners who generated responses receiving scores that were two points away from each other. The three pairs who generated responses two points different from each other on posttest started the embedded essay writing and revision activity with an initial score of four, compared to the average of 3.6. This may suggest that the revision work was more representative of the work of one student in the pair. One was leading the two in revision during instruction while the other partner contributed but integrated fewer of the ideas surfaced during the revision experience. Nevertheless, all six students revised their responses on the posttest.

Teacher guidance to facilitate collaborative revision

The teacher customized the automated guidance system prior to instruction and refined her guidance strategies to support each student to engage in knowledge integration as they worked together to revise their explanation. We examine how the teacher customized her guidance and how it influenced student pairs' revision process.

Customizing the automated guidance system

During the planning meeting with the researcher, the teacher customized the automated scoring system. She modified the automated alerts threshold to a score level of two in order to catch student pairs who did not demonstrate a normative idea after one round of revision. The teacher also modified the guidance to display the automated score to each pair, below the adaptive hint. The teacher thought that displaying the score in each round of guidance would increase the student pair's motivation to improve their explanation through revision. Building on these customizations, at the start of class, the teacher emphasized the importance of revision as students began the unit. She explained to the whole class that for two explanations they would write in the unit, they could receive two rounds of feedback with a score from the computer. She expected the student pairs to use the feedback to revise each response, at least two times. She expected all pairs to improve their score with each revision, and by the end, have a complete and accurate explanation. She would continue to review each pair's final explanation after the two revisions and reward continued refinement with a higher score.

The teacher customized her in-class monitoring strategy to take advantage of the automated explanation scores to guide collaborative revision during instruction. Prior to the start of the unit, she customized her monitoring system to track each pair's progress in explanation revision, and, to keep track of who she had assisted in explanation revision, as she circled the classroom. She used a clipboard with a paper listing each student pair in each class period (printed from WISE teacher tools), and a column for her to add two scores for each of the embedded explanations (Mountain and Lava Lamp). She left a blank column for notes. The teacher circled the classroom and checked in with each group to record their automated explanation scores and to probe the thinking of each student about their revision. By recording scores, the teacher ensured she checked-in with each pair at least two times as they revised each explanation. The teacher reflected after instruction on how this process worked:

> The automated feedback allows them to evaluate their own work. It might involve some teacher probing....When intercepting the students between submission 1 and submission 2, checking what was the first feedback, what are you going to add, why are you going to say that, are you really answering the question...I keep probing besides the computer feedback, I think we get there [understanding] through conversation.

The teacher was able to meet with each pair during class for an extended conversation because the other pairs would continue to work at their own pace using the automated guidance in the unit until she came to meet with them. As the teacher described, this system held each pair accountable for working together to improve their explanation. Further this system made it clear to all students that checking on each pair was a part of the teacher's routine. She did not single out students based on their scores.

Customizing in-class guidance strategies

The analysis of the teacher's in-class interactions with student pairs demonstrates how the teacher took advantage of the automated guidance in this monitoring approach to adapt her guidance strategies to support each student to contribute to the revision process. The analysis is based on students' logged explanation revisions and 37 audio/video recordings as the teacher guides 13 different student-pairs to revise their explanations. The teacher's guidance was coded for five distinct strategies, as described in Table 3.

Teacher's frequency of use of guidance strategies Figure 3 presents the overall frequency that the teacher used each of the five guidance strategies described in Table 3, across the thirty-seven teacher-pair interactions with thirteen different pairs. Critical to supporting collaborative revision, the analysis demonstrates the integral role of the automated guidance in the teacher's guidance approach. Her first move with each pair was to establish a

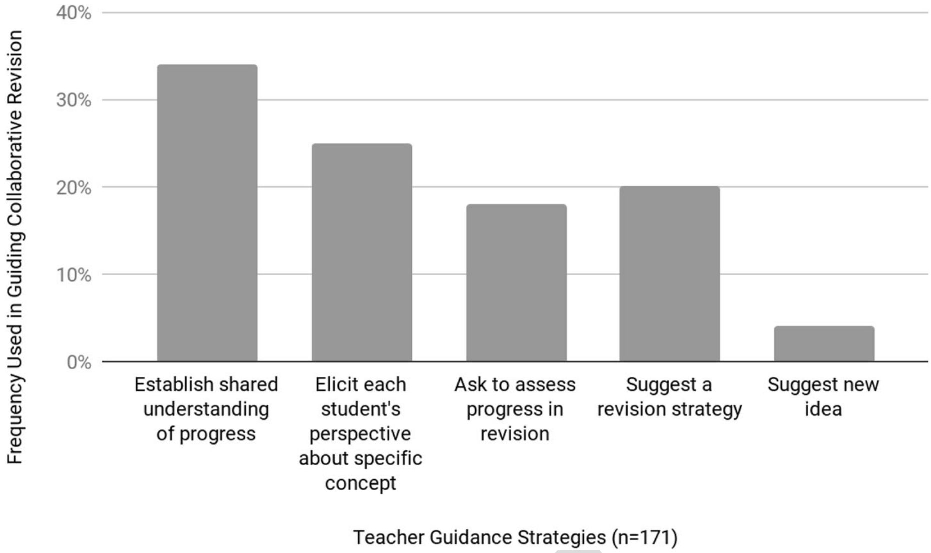


Fig. 3 Frequency of type of teacher guidance statements (out of all recorded teacher guidance statements) to support collaborative revision presented in the sequence the teacher most commonly used them

shared understanding of the pair's progress, her most frequently used guidance strategy. She typically did this by prompting one student in the pair to read their automated guidance, score, and initial explanation aloud. The teacher then used the automated guidance to personalize her subsequent guidance moves. Next, the teacher either elicited each student's ideas about a specific concept in their initial explanation that was highlighted as missing or inaccurate in their response by the automated guidance. Or, she asked each student to assess their revision progress. This involved prompting the pair to reflect on their explanation and distinguish what idea is missing or needs elaboration. Each of these strategies served to surface the two student's different ideas or to elucidate a shared gap in their views. The teacher then recommended a revision strategy the pair might try to reconcile or elaborate their views. The teacher frequently encouraged students to revisit evidence from earlier in the unit, or she clarified a revision process such as how to incorporate details into a response. Notably, the teacher rarely suggested new ideas for the student pairs to consider. Rather she encouraged the students to put forward and elaborate each of their ideas for the pair's consideration.

Teacher's sequence of guidance strategies Using her checklist approach, the teacher circled the classroom to work with each pair approximately two times during the revision process, for each of the two explanations. During each interaction, the teacher frequently used a combination of several of the strategies identified in Table 3. The sequence of strategies was naturally temporal (starting with establish progress and ending by recommending a revision strategy) and involved some variation in the middle based on the pair's expressed initial understanding.

The teacher most frequently began a conversation by *establishing a shared understanding of the student pair's progress*. She asked one of the two students to read their assigned automated guidance and score aloud. In some cases she also asked the pair to read their written explanation aloud. This gave the teacher quick insight into how to diagnose the student pair's ideas in terms of understanding plate tectonics, and established a shared understanding among the pair and teacher of what ideas were missing or inaccurate in their initial response.

Then, particularly for pairs who demonstrated vague or disconnected ideas in their initial response, the teacher frequently *elicited details about each student's perspective* on an idea that was highlighted by the automated guidance. The targeted idea was typically one that was vague or inaccurate in the pair's initial explanation. The teacher would pose a question that built on the pair's initial statement, and encouraged each student to extend their reasoning (e.g. "You mentioned great things like convection currents. You said it moved the blobs up and down. But what you should tell me s, is moves it up because <pause>.") Many students' approach to revision involved "answering" the question posed by the automated guidance, rather than integrating the new information prompted by the hint with the ideas they expressed in their initial explanation. The teacher's prompt served in some cases to raise disagreement between the two students in the pair. In other cases it gave the students a wider pool of related ideas to draw upon when elaborating and connecting their views. The teacher rarely responded by expressing judgement on the accuracy of either student's expressed idea. Rather she followed-up with another discipline relevant but general hint that encouraged the students to take ownership for elaborating, or reconciling, their views (e.g. *"If something is hot, where does it go?...Why does it go up?")*.

For student pairs who expressed at least one correct and accurate link in their initial essay, the teacher often started by asking the student pair *to assess their progress in revision*. This involved more general questions that prompted each student to evaluate their shared explanation and distinguish what idea, if any, they think might strengthen their response (e.g. "Do you think you should add anything else?"; "What else are you going to say?"; "What are your first thoughts about what are you going to write?"). These questions helped the students make their ideas visible to one another and often revealed a disagreement or a shared confusion. The teacher also interleaved these guidance prompts albeit less frequently, for pairs who demonstrated a lower initial understanding. Asking each student to reflect on their response and distinguish what idea may address a gap often revealed disagreement between the two students or shared uncertainty. The teacher used this to motivate the student pair to pursue exploration of an idea.

The teacher ended most conversations with a student pair by *recommending they try a concrete revision strategy*. Revision strategies included revisiting evidence in a dynamic visualization suggested by the automated guidance (e.g. "check the animation"*)*, or clarifying revision strategies suggested by the automated guidance (e.g. "elaborate means to add some more details").

The teacher reminded each of the student pairs at the end of each interaction that she would circle back to check on their work later during the class period, after they had a chance to revise. This held each pair accountable for making progress.

Due to the teacher's involved monitoring approach, only two groups received an automated teacher alert. The teacher noted that these two pairs included students with an Individualized Education Program (IEP) to assist with special needs. The teacher appreciated that the alert enabled her to provide these student pairs just in time assistance:

> The alerts are helpful for identifying especially my resource or ELL students who need my help. It tells me they need my help right now. Like that group [who had received an alert] I knew I needed to help them translate what they could speak from up here, into writing.

Overall the teacher reported that guiding collaborative revision using the automated scores and adaptive knowledge integration guidance was a positive experience. She reported two main challenges. First, by adding her own off-line scoring system she had difficulty reconciling it with the WISE scoring. Thus, she said,

> I am trying to go in [to the grading tool] daily to review student revisions, because they [the pairs] are at a slightly different spot...I can keep track of who I have responded to on my own paper but it would be helpful to figure out in WISE how to keep track of which responses I have responded to already, and which I have yet to review. [NOTE: WISE has a tablet tool that could be customized for this use.]

The teacher also noted that her students sometimes questioned the accuracy of the automated guidance. She highlighted this as a strength:

> Some kids said 'the feedback said I needed to mention density but hey look I already mentioned density.' So even them [the student pair] really looking at the feedback, and then evaluating their own work, is good. Then we can decide well did you really mention density, or explain it?

Overall, the teacher's guidance strategies and reflections illustrate the ways a teacher and computer can work synchronously to effectively guide students in collaborative revision in a learning environment.

Customizing guidance strategies to pairs needs

The teacher adapted her strategies to align with the needs of each pair. This led to differences in her use of guidance strategies for student pairs depending on their initial ideas (see Table 7). For pairs who expressed at least one complete and accurate connection between two ideas the teacher more frequently prompted them to assess their progress in revision. She called for each student to distinguish a gap in their response and articulate what idea might ameliorate the gap. Relatedly, when recommending a revision strategy, she emphasized how to incorporate additional details into their initial explanation rather than gathering more information (Fig. 4).

For pairs who expressed vague ideas or one correct but isolated idea, the teacher spent more time eliciting each of the student's views about a targeted idea. When recommending a revision strategy, she most frequently suggested the pair revisit evidence to elaborate their response. In sum, the teacher guided pairs who had linked ideas to distinguish the gap in their explanation and to incorporate a new idea. For pairs with a partial understanding, she focused on helping the pair gather relevant evidence to connect with their initial idea. This suggests that the focus

Table 7 Frequency of teacher guidance strategies when facilitating Low/Med versus high prior knowledge student pairs to revise explanations

	Est. shared understanding of progress	Elicit student's perspective	Recommend a revision strategy	Ask students to assess progress	Suggest new idea
Vague/Disconnected $N = 48$ teacher strategies	33%	30%	25%	**8%**	4%
Linked $N = 128$ teacher strategies	34%	22%	17%	**22%**	4%

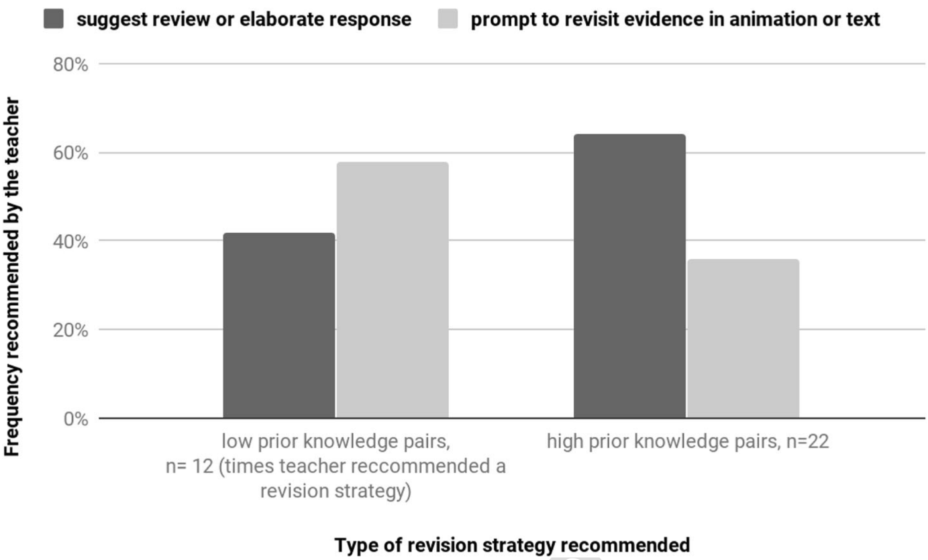


Fig. 4 Type and frequency of revision strategies recommended by the teacher when facilitating low/partial versus high prior knowledge student pairs to revise

of the collaborative learning for student pairs starting with partial understanding, from the teacher's perspective, was to revisit evidence and determine what evidence to draw on to extend the idea expressed in their initial explanation. Whereas, for students starting with a linked understanding, the focus of the collaborative learning is on evaluating their explanation to identify what is missing and to incorporate details or links to elaborate their view.

Examples of guidance for students collaborative revision

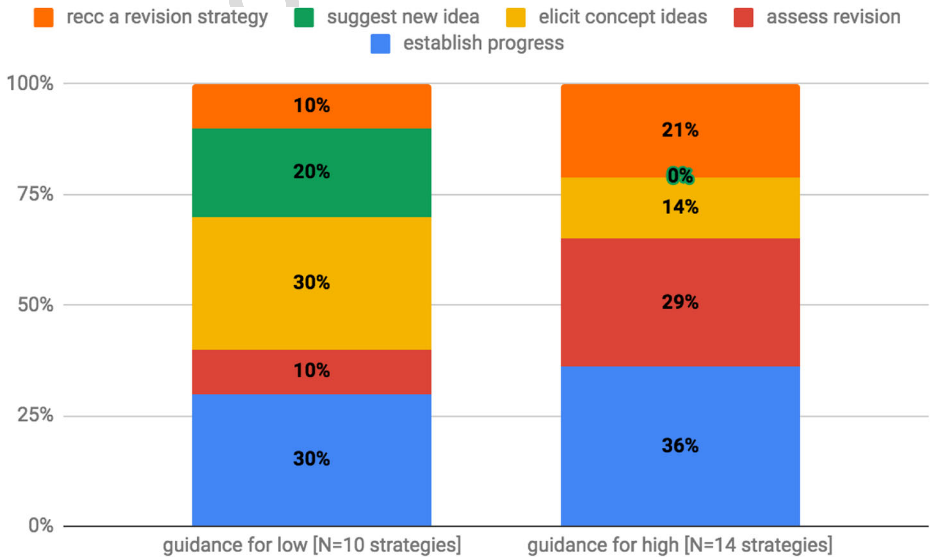
A cross case comparison illustrates how the teacher took advantage of the automated guidance and customized her guidance strategies to support student pairs in a collaborative revision process. We selected these two pairs because they were most illustrative of how the teacher guided a pair demonstrating partial versus linked understanding on their initial essay, out of the data set of teacher guidance for 13 pairs. The case is meant to give insights into how the teacher customized guidance in support of collaborative revision for two different pairs; it is not meant to be representative of the whole.

In the first example the teacher moves to elicit each student's perspective about the role of density, without giving them any new information to consider. One of the students responds by paraphrasing their initial response and attempting to connect it to plate density. The teacher recognizes an idea to build on in the student's response and presses each student to say more. Each student gives a different elaboration, adding to each other's view. The teacher builds on their shared perspective by prompting the students to distinguish the link between plate density and movement. With this question, each student gives a conflicting idea about density. The teacher affirms one student's view and suggests the students revisit the evidence in the unit to ensure they both agree. She then leaves the pair to begin their revision, promising to check back.

In the second example alternatively, the teacher prompts each student to distinguish a gap in their explanation and how they would address it. This surfaces shared confusion by the students about how to approach revision of their explanation. Rather than raising conflict in this case, one student puts forth a vague idea and the other a more targeted idea that extends an idea in their initial response. The teacher recognizes this as a promising idea to pursue and that it connects to an earlier class warm up activity. By connecting the idea to the class warm up, she attempts to make the idea more accessible to the partner.

These cases illustrate how the automated guidance supports the teacher to efficiently personalize guidance to elicit each student's ideas on how to augment their shared explanation. Evident in these two examples is the teacher's differential use of the guidance strategies. As shown in Fig. 5, the teacher places greater emphasis on eliciting each student's ideas about a specific concept and less emphasis on prompting the students to assess their progress in revision. She also offers a new idea for the pair to consider. In contrast, for the pair starting with a more complete understanding, the teacher more frequently prompts the students to assess their progress in revision. This reveals how she adapts her guidance strategies to the student pair needs: one pair needs further help in surfacing a specific idea to link to and complete their partial idea; the other pair needs further assistance in identifying a gap in their response and distinguishing what idea would address it. This calls for locating a gap within the response and determining what information would elaborate their view.

For both pairs, it is evident that one student holds a more robust understanding of the topic than the other student in the pair. In the first example, S1 leads the elaboration of subduction in mountain formation. In the second example, S1 guides elaboration of density and movement in convection. In both examples, the teacher makes visible each student's ideas, with a goal that the two students, ultimately, will integrate or reconcile their views so both have a more coherent and accurate understanding. In both of these examples, it appears that S2 will gain new ideas from hearing S1's perspective, and both students will gain from working together to elaborate and integrate this idea in their explanation. Their revised explanations suggest this is what occurs, as both pairs integrate a new idea into their initial explanation, creating a more comprehensive final explanation.



Q11

Fig. 5 Teacher's frequency of use of guidance strategies for a pair starting with partial understanding versus complete understanding on their initial explanation

The analysis of the pair's individual work on the posttest suggests each student gained more from the collaborative revision experience. Yet their gains are incremental for the individual, as opposed to reflective of the pair's progress in the revision activity. For the pair who began with partial understanding for instance (Tables 8 and 9), on the posttest, one of the students articulated a single idea to explain mountain formation ("two plates collide") and added a vague statement in revision ("It's going under and coming back up"). The second statement is most likely in response to the guidance on the posttest asking about what is happening in Earth's mantle. The other student from the same pair expressed, on the posttest, one idea ("plates collide and go up..") and connected it to another partial idea about convection during revision ("convection currents going up because they are less dense push the plates together"). Neither response illustrates the integrated view expressed as a pair about subduction. Yet each student expresses a wider repertoire of accurate and relevant ideas than they did on the pretest, and the second student links a partial idea.

Likewise, for the pair who began the embedded activity with linked ideas on posttest, the two students demonstrate different levels of individual understanding on the posttest. One of the students expressed a linked explanation of mountain formation:

> The two plates (oceanic and continental) were moving towards each other at a very slow rate...They were being moved by convection currents in the mantle...the oceanic plate subducted under the continental plate because it is denser. The magma below was so hot that it melted the plate. This created a subduction zone and built up a lot of pressure. It became so great that the magma created a volcano or mountain…

They incorporated an idea about how convection works:

> Convection currents are caused by the core heating up liquid rock in the mantle. When this happens the density of that rock decreases causing it to rise. When the magma reaches the top it cools down, making its density increase. The magma slowly sinks back towards the mantle and the process happens over again.

The other student expressed a single idea ("2 plates are moving toward each other over millions of years the plates eventually form a mountain.") and links a partial new idea about convection during revision ("Inside earth's mantle...the force from density in the mantle moves the plates together").

Analysis of the individual responses provides a platform for speculation on how the teacher's guidance for the partnered work during instruction differentially influenced individual learning. The individual posttest responses and revisions suggest that the teacher's guidance during revision of the for the pair with initially linked understanding may have better supported the two students to integrate new ideas. This would be consistent with the pre/post outcomes for students who demonstrated linked understanding on the collaborative explanation, versus disconnected ideas. The teacher's emphasis on guiding the students to evaluate their explanation and direct their revision process may have supported the students to consider and link a new idea with their initial views and to gain insights into a revision approach of integrating ideas. For the pair who started with partial understanding, the teacher's emphasis on eliciting elaboration of a specific idea may have supported the individual students to add a new idea to their repertoire. The students may not have linked this idea with their other views leaving them with a wider repertoire of accurate but fragmented ideas.

t8.1 **Table 8** Excerpt of teacher guidance for collaborative revision of Mountain explanation, with a pair demonstrating initially partial understanding

		Text	Code
t8.3	T	This is your first try. D, can you read what you first typed?	Establish shared understanding of progress
t8.4	S1	<reads aloud> The mountains form when the continental hits each other but when there is also oceanic crust. The continental spreads letting the oceanic go under, letting the magma slip through making a volcano.	
t8.5	T	Ok, so are you ready to hit submit? Alright! What's your score?	Establish
t8.6	S1	3	
t8.7	T	Can you read your feedback?	Establish
t8.8	S2	<reads aloud> J, D, add details to your explanation. How does the density of the two plates affect their movement? Check here for a hint. Then, revise your explanation."	
t8.9	T	OK so you have to mention something about density. Let's reread your answer. Did you use the word dense or density at all?	Assess progress in revision
t8.10	S1.S2	No	
t8.11	T	Do you think density is involved?	Elicit idea about a specific concept
t8.12	S1.S2	Yes	
t8.13	T	How do you think so? By thinking first?	Elicit
t8.14	S1	How the ocean is dense so it lets it slip through.	
t8.15	T	What do you mean slip through?	Elicit
t8.16	S1	Like when it goes here <gesturing under with hands>	
t8.17	S2	The magma	
t8.18	S1	It goes under.	
t8.19	T	Oh it goes under. So if it goes under, does it mean its more dense or less dense?	Suggest new idea
t8.20	S1	More dense	
t8.21	S2	Less dense	
t8.22	T	More dense, he's right. So that's good, that's kind of what we're looking for, you're guys are right.	Suggest new idea
t8.23	T	But you'll still definitely want to click on the animation to double check it and then fix it.	Recommend a revision strategy
t8.24	S1.S2 revised explanation	<after revising one time > The mountains form when the continental hits each other but when there is also oceanic crust. The continental spreads letting the oceanic go under because the water is more dense then the continental so the oceanic crust will go through because it is more dense letting the magma slip through making a volcano.	

JrnlID 11412_ArtID 9298_Proof# 1 - 18/04/2019

t9.1 **Table 9** Teacher guidance for collaborative revision of Lava Lamp explanation, with a pair demonstrating initially high level of understanding

		Text	Code
t9.2			
t9.3	T	Ok, you are on 4.2, This is the first tim and you have not hit submit yet. Read what you have.	Establish
t9.4	S1	<reads aloud> We believe that lava lamps work using convection currents. The heater at the bottom of the lava lamp heats up the fluid inside. It rises to the top of the container and cools down. It gets more dense and slowly sinks to the bottom. This process keeps on happening over and over again. A lava lamp is like the Earth's mantle because the lamp in the lava lamp is basically the Earth's core. The core heats up the fluid in the mantle and it slowly rises to the crust. It also slowly cools down and slowly drifts down closer and closer to the core. This process takes millions of years to happen. The lava lamp only takes around 30 s to happen.	
t9.5	T	Do you guys feel good about what you've written?	Assess progress
t9.6	S1	ya	
t9.7	T	OK, dare you to hit submit. OK, What's your score S2?	Establish
t9.8	S2	C, C, elaborate. Use your ideas about heat and density to elaborate why the blob goes all the way up, and, all the way down. Check this animation to get more information. Then, add details to your explanation here .	
t9.9	T	Do you guys think you did that (referring to the auto guidance)?	Assess progress
t9.10	S2	Um (quiet)	
t9.11	T	Because elaborate just means to add details.	Revision strategy
t9.12	T	Do you guys think you did mention why it goes all the way up?	Assess progress
t9.13	S1	No	
t9.14	S2	No	
t9.15	T	Well let's scroll up here <to their explanation> and see what you guys said. Where do you talk about the blob or lava lamp?	Establish
t9.16	S1	Here. From here to there <pointing to text in their explanation>	
t9.17	T	OK, you put <reads aloud>	Establish
t9.18	T	How do you think you could elaborate or add more details?	Elicit
t9.19	S2	I think we should add the details maybe, inside the lava lamp, and what's happening	
t9.20	T	What do you think S1	Elicit
t9.21	S1	I think we should add more on how the molecules come further apart from each other and ummm	
t9.22	T	Ah, like in the warm up activity today right?	Establish
t9.23	S1	Ya	
t9.24	T	OK, so don't delete anything. But, maybe describing the molecules, that would be elaborating right? OK, I'll check back.	Revision strategy
t9.25	S1.S2 revised explanation	We believe that lava lamps and Earth's mantle work using convection currents. The heater at the bottom of the lava lamp heats up the fluid inside. The molecules spread apart as they get hotter. It rises to the top of the container and cools down. The molecules slowly come back together as they cool down and the blob sinks to the bottom. This process keeps on happening over and over again. This process keeps on happening over and over again. A lava lamp is like the Earth's mantle because the lamp in the lava lamp is basically the Earth's core. The core heats up the fluid in the mantle and it slowly rises to the crust. It also slowly cools down and slowly drifts down closer and closer to the core. This process takes millions of years to happen. The lava lamp only takes around 30 s to happen.	

Limitations

This study captures how a teacher takes advantage of automated explanation scoring and adaptive guidance to guide students in collaborative revision during science inquiry. Although the combination of teacher and computer guidance appears to support students in both disciplinary learning and revision, this case study does not demonstrate that the automated and teacher guidance caused the improved outcomes. Rather the results are an outcome of the learning environment activities and the teacher and computer guidance. We cannot disentangle the particular influence of each in this study. Further, this study examined the teacher's guidance while interacting with a pair of students. It did not investigate the student pair's activities when interacting by themselves. This provides insight into how a teacher can guide collaborative revision and how that guidance may differ for pairs with differing levels of demonstrated initial understanding. Future work may investigate how the pairs collaborated with each other and the teacher.

Discussion and conclusion

Computer-supported collaborative learning contexts have the potential to support meaningful revision yet even when working in pairs and prompted by guidance few pairs make meaningful revisions to their explanations (Gerard and Linn 2016; Sun et al. 2016; Sinha et al. 2015; Tansomboon et al. 2017). This research took advantage of a WISE unit that used scientific visualizations and automated, adaptive guidance in instruction designed to promote knowledge integration about plate tectonics processes. We studied how the teacher customized and leveraged the automated guidance to engage pairs of students in making revisions to their explanations of complex phenomena such as how mountains form. We analyzed the teacher's guidance strategies and how students revised in response to the teacher and automated guidance.

Customizing guidance for collaborative revision

The teacher customized her own in-class monitoring approach and the automated guidance tools to combine both in a system promoting student pairs' knowledge integration through explanation revision. She modified the automated guidance to reveal a score for each student pair's revision and to set the teacher alerts threshold at a knowledge integration score of 2, in order to catch student pairs who held inaccurate or vague ideas after one round of revision. The teacher then created a system to guide pairs' during explanation writing. She used the WISE teacher tools to create a checklist so she could keep track of each collaborating pair and record their automated score with each round of revision. The teacher made these changes after review of the unit and the guidance for each level of the knowledge integration scoring rubric. The teacher's careful review of the guidance gave the teacher insights into common student ideas at each level, and the key visualizations within the unit to help student pairs use evidence to move up one level in the rubric. At the start of class, the teacher framed explanation revision as a goal as a goal of the Plate Tectonics inquiry unit. She expressed a clear expectation that all student pairs should revise their explanations of mountain formation and convection at least two times using evidence, increasing their score with each revision. Revision, she emphasized was a part of doing science.

During class, the teacher used the checklist to make sure she visited each group two times during the process of revising each of two explanations. She circled the classroom, reading over student pair's shoulders, and probing for understanding. She recorded the explanation score as she visited each group. Analysis of audio/video records showed that she used five main guidance strategies to help collaborating students to revise. The teacher was able to engage in such an extended conversation with each pair during class because the other pairs continued to work independently, using the automated guidance in the unit until she came to meet with them.

The strategies align well with the goals of knowledge integration. When meeting with each pair, the teacher created a shared understanding of the student pair's progress by prompting them to read their automated guidance and response aloud. For students who demonstrated one complete idea in their initial explanation, she was more likely to next, prompt them to assess their progress in revision and determine the next step. For students who expressed a vague or inaccurate idea in their initial response, she focused her question on a particular concept highlighted as missing or inaccurate in the pair's joint response by the automated guidance. She guided each student in the pair to elaborate their thinking about this idea. Guiding each student to distinguish the next idea to pursue, or, to elaborate an idea in their initial response made each collaborator's ideas accessible to one another and broadened their shared repertoire of ideas for consideration. This often led the two students in each pair to realize they held conflicting or incomplete ideas. The teacher then, encouraged the students to return to evidence presented earlier in the unit to elaborate or reconcile their views.

The teacher's press for students to elaborate their idea in the presence of their partner was a key support for effective collaboration. Matuk and Linn (2015) found that online discussion was most beneficial for individual student learning when students first selected an idea of their own to share with their classmates, and next, identified a peer's idea in the discussion that was different from their own. This process required students to generate a well-formed idea before they looked to their peers' ideas. This is valuable because it increased the likelihood that students were analyzing well-formed science ideas, as opposed to superficial or social comments, when they compared the peers' ideas to their own. Similarly the teacher in this study guided each student to express a well-formed idea about the targeted concept. While some students initially responded to the automated guidance by expressing a partial idea, in an attempt to "answer" the question presented by the automated guidance, the teacher elicited each student's reasoning to help them articulate a more complete science idea. The teacher's strategy gave credibility to the voice of each student and in doing so encouraged each partner to consider the other's view. The teacher did not require the partner to accept or reject the other's idea, but called for each to consider the other's idea relative to their perspective. In combination with the Matuk & Linn findings, these studies suggest that a key for supporting collaborative learning may be support for students to articulate a well-formed idea in the presence of the other. This is markedly different from typical online discussion activities or instructor prompts that call for students to participate in a discussion, without necessarily guiding them to formulate a complete idea to contribute first.

In this study, the teacher's guidance strategies aimed to support the pairs in what Berland et al. (2016) characterized as meaningful revision, as they interacted with the teacher and one another. The teacher elicited each student's ideas, making the pool of ideas for the pair's consideration apparent and accessible to the two students. Making the pool of ideas apparent often revealed gaps in the pair's views, or a conflict in their expressed understanding. This gave the student pair reason to revisit evidence. The teacher guided the students to revisit a particular visualization in the unit to clarify their two views. Notably the teacher then left the group to work reminding them she would check back to follow-up on how they have progressed. This placed ownership on the pair for deciding how to proceed, while also holding the two students accountable for making progress.

Students' engagement in revision during instruction gave them a foundation to draw upon when revising their explanations on the posttest, without a partner or teacher support. On the posttest, compared to the pretest, twice as many students revised and improved their explanation of mountain formation by one knowledge integration score. This is likely due both to students' improved ability to revise and their larger repertoire of relevant ideas to draw upon after instruction. While the percentage of students who revised and improved was greater than the percentage of students at pretest, it remains a low percentage of students overall (22%). This may be partially related to the degree of improvement needed in a revision to move up one level in the knowledge integration rubric. Students who began with a score level of three (partial understanding) for example and added a partial idea to their response through revision, remained at a score level three because moving up a level calls for adding and connecting an idea. In addition, the analysis of student revisions on pretest and posttest, as illustrated in the case study, demonstrate the incremental and individualized progress students made when learning to revise. One student moved from not revising their explanation at all on the pretest to adding a partial idea to their initial explanation in revision on the posttest. The other student moved from adding a partial idea to their initial explanation on pretest, to linking an idea on posttest. While this case is not generalizable it helps characterize the type and degree of improvement students make in revision of written explanations in science. The findings suggest students may benefit from additional guidance focused on how to integrate ideas in revision and why this matters for learning, in contrast with adding more but disconnected ideas.

Teacher and computer as partners in guiding CSCL

The study reported here demonstrates how a teacher can guide students in successful collaborative learning in a computer supported environment. In contrast to a recent meta-analysis on CSCL that found teachers had limited impact on students' collaborative learning, this study reveals ways teachers can have impact. Among multiple moderators investigated, the teacher was one of few that did not yield a significant positive effect (Chen et al. 2018). The teacher followed in our study presents strategies that can be used to strengthen students' collaborative learning in a CSCL. The teacher and computer worked together in this study, in that the computer provided the teacher with an efficient diagnosis of the student pair's ideas and a hint for advancing the pair's response. The teacher helped each student to interpret the computer hint and apply the hint to their initial view. The interactions between the teacher and the students revealed difficulties students had interpreting the evidence in the unit. The teacher strengthened learning by guiding each student to articulate a complete idea, and in doing so narrowing their focus and deepening their interpretation of the evidence. Thus, the teacher highlighted the elements of the automated guidance that were most relevant to each student while also giving the student pair confidence to use the guidance. This adds to the growing knowledge of how teachers guide students to develop integrated understanding in computer-supported collaborative learning environments (Furberg 2016; Ingulfsen et al. 2018).

The teacher's strategies identified in this paper reveal how a teacher can involve student pairs in high cognitive engagement as they work in a CSCL environment, consistent with Sinha et al. (2015). High cognitive engagement is characterized by a group's thoughtful and deliberate uptake of the affordances offered by the computer-based learning environment. Students' deliberate use of a simulation in which they make predictions, test them and reflect on the results, or, when students revise models using ideas from peers, were given as exemplars of high cognitive engagement. Low cognitive engagement alternatively was characterized by a focus on superficial features in a computer supported learning environment such as neatness or color in a simulation. Among the 10 student groups Sinha et al. (2015) studied, the mean cognitive engagement score was low; only one group demonstrated high cognitive engagement in spite of the technological resources available for collaborative learning. Students' social and behavioral engagement scores were much higher. These and related results suggest that guidance on how to collaborate can help students tend go beyond operating at a social level or alternating individual contributions (Cohen 1994a, b). The teacher enables students to make connections between the evidence presented by dynamic visualizations or graphical representations and the underlying science principles. This frequently involves the teacher pointing out salient features in the digital evidence that extend or challenge the student pair's ideas (Furberg 2016; Ingulfsen et al. 2018). The teacher's customized monitoring system and guidance strategies presented in this study, and the teacher's differentiated use of the guidance strategies depending on students' initial level of understanding, supported each student to express their view, recognize the ideas of the other, and determine how to connect their views in revision.

The teacher's combined use of the automated guidance with her monitoring system benefited students' overall in terms of their engagement in revision and knowledge integration. Prior research studying how peer's revise their models or explanations has demonstrated that in most cases students working in groups focus more on task procedures such as division of labor than critique and refinement (Sun et al. 2016). Or, that when prompted to revise, about 60–70% revise at all and only about 20–30% of those students engage in meaningful revision, in which they evaluate and modify their initial explanation to integrate new evidence or reasoning (Tansomboon et al. 2017; Sun et al. 2016). In this study, all of the student pairs engaged in two rounds of revision. They improved their knowledge integration scores on the embedded explanation by moving from partial understanding to a linked understanding, making a connection between two key ideas.

In the study reported here, the teacher created a checklist system that essentially assigned an alert to every pair. Results from an earlier study of teacher alerts with two different teachers revealed benefits for alerts. In the earlier study the two teachers set alerts for a specific score level using automated essay guidance and were prompted to guide the subset of the student pairs receiving alerts (Gerard and Linn 2016). The alerts increased the teachers' opportunities to talk with their students about the unit. Both of the teachers in the earlier study reflected that they gained insight into student ideas that they had not previously anticipated about photosynthesis. During the process of working with students, each of the participating teachers developed new guidance strategies to respond to each student's developing ideas. For example, one teacher found that many of her student pairs needed more support interpreting the evidence presented by a dynamic visualization of photosynthesis inside a chloroplast. The teacher began by directing each student to return to the visualization suggested by the automated guidance. She then prompted each student in the pair to articulate step-by-step, what they observed in the photosynthesis process as they advanced through the visualization. When a student was struggling to articulate what they were seeing, she encouraged the student to ask their partner or to ask another nearby pair for assistance. By calling on nearby pairs, she held multiple groups accountable for active learning at the same time. Thus in these studies the teacher also directed students to the visualization to gather additional evidence. As in the study reported here, the automated scoring tools motivated the teacher to talk with the students. These conversations identified science ideas that needed further probing.

In both the prior work and this study, it is evident that the automated explanation scoring tools can support teachers to personalize guidance for students in real-time. This includes teachers who are new to computer-supported instruction or new to teaching a certain topic, as the teacher in this study. This stands in contrast to prior work demonstrating teacher's limited use of content-oriented guidance that responds to and extends students' ideas (Black and Wiliam 1995; Ruiz-Primo and Furtak 2007; Ruiz-Primo and Li 2013). It also extends the work documenting expert teacher's guidance (e.g. van Zee and Minstrell 1997) to a typical teacher without prior experience teaching the WISE unit. These findings suggest that with limited professional development focused on customization, a teacher can promote students' collaborative knowledge integration when taking advantage of the automated scoring tools.

Teacher professional development for guiding CSCL

The five documented strategies as well as the customization process used to ensure that the teacher supported each pair can inform the design of professional development. These insights can be used to help other teachers prepare for implementation of CSCL in their classroom. The time the teacher in this study spent planning customizations to the automated guidance, prior to the start instruction, was essential to her development of these guidance strategies. She ultimately made few customizations to the automated guidance, yet the process involved the teacher in careful review of the explanation prompts, the range of likely student responses, possible hints and evidence to support students to advance their understanding at each level, and a deliberate reflection on how to monitor student progress during implementation. This planning time supported the teacher to connect and augment her own monitoring approach to monitoring student progress in real time with the WISE tools.

Teacher professional development for guiding collaborative learning had previously emphasized how to organize student groups and design tasks so that students need each other to succeed and students were responsible for guiding one another - as opposed to relying on the teacher's direct instruction (Cohen, 1994a, b). CSCL environments have shifted the teacher's role from these earlier studies and hence the needed focus in teacher professional development for teaching with a CSCL environment (Tissenbaum et al. 2012). Environments such as WISE typically provide rich tasks such as writing explanations that call for connecting two to three evidence-based ideas, or investigating interactive, dynamic models. Further, the environments enable student pairs to direct their inquiry (Donnelly et al. 2014). The teacher's role in this context is focused less on Q15 designing complex activities or guiding class inquiry, and more on linking the goals of the CSCL to their instructional goals and practices (Roshcelle et al. 2013). This calls for determining goals for collaborative inquiry, how to monitor student pair's collaboration as they progress through the investigation, and planning when and how to intervene. As evidenced by this study, working in partnership with teachers to customize the CSCL tools for their class can give teachers' ownership for facilitating student learning in the CSCL, and, reason to plan ahead. Teachers might identify steps to monitor, anticipate student ideas that need probing, and develop possible strategies for eliciting the ideas held by each member of a pair.

The research presented here demonstrates the potential for advanced technologies to engage teachers and students in complex scientific activities including generating and revising science explanations in a typical middle school classroom. The combination of technologies, designed and integrated to promote and capture knowledge integration, including natural language processing, dynamic visualizations, a web-based learning environment, and data logging – were all necessary to provide the students and teacher meaningful guidance opportunities. The flexibility of the automated scoring and guidance system and the transparency of automated scoring rubrics enabled the teacher to combine the CSCL with her own teaching and assessment practices. The teacher augmented the automated guidance for her students, and the automated guidance supported the teacher, to personalize her guidance in real time. The combination of technologies was also necessary for researchers to gain detailed empirical insights into students' revisions and developing understanding in plate tectonics. The open-source technologies used in this research can be used in future web-based curriculum and assessment materials to support teacher guidance for classroom learning.

Acknowledgements This material is based upon work supported by the National Science Foundation under grant Nos. DRL-1119670, DRL-1418423, and DRL-01451604. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Berland, L. K., & Reiser, B. J. (2011). Classroom communities' adaptations of the practice of scientific argumentation. *Science Education, 95*(2), 191–216.

Berland, L. K., Schwarz, C. V., Krist, C., Kenyon, L., Lo, A. S., & Reiser, B. J. (2016). Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching, 53*(7), 1082–1112. https://doi.org/10.1002/tea.21257.

Black, P., & Wiliam, D. (1995). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal, 22*(5), 537–548.

Brownell, S. E., Price, J. V., & Steinman, L. (2013). Science communication to the general public: Why we need to teach undergraduate and graduate students this skill as part of their Formal Scientific Training. *Journal of Undergraduate Neuroscience Education, 12*(1), E6–E10.

Chen, J., Wang, M., Kirschner, P. A., & Chin-Chung, T. (2018). The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: A meta-analysis. *Review of Educational Research, 88*, 1082–1112. https://doi.org/10.3102/0034654318791584.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3*(3), 472–517. https://doi.org/10.1016/0010-0285(72)90019-9.

Clark, D. B., & Sampson, V. (2008). Assessing dialogic argumentation in online environments to relate structure, grounds, and conceptual quality. *Journal of Research in Science Teaching, 45*(3), 293–321.

Cohen, E. G. (1994a). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64*(1), 1–35.

Cohen, E. G. (1994b). *Designing Groupwork strategies for Heterogenous classrooms*. New York: Teachers College Press.

Cohen, M., & Riel, M. (1989). The effect of distant audiences on Students' writing. *American Educational Research Journal, 26*(2), 143–159.

Crawford, L., Lloyd, S., & Knoth, K. (2008). Analysis of student revisions on a state writing test. *Assessment for Effective Intervention, 33*(2), 108–119. https://doi.org/10.1177/1534508407311403.

diSessa, A. A. (2006). A history of conceptual change research: Threads and fault lines. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 265–282). New York: Cambridge University Press.

Donnelly, D. F., Linn, M. C., & Ludvigsen, S. (2014). Impacts and characteristics of computer-based science inquiry learning environments for precollege students. *Review of Educational Research, 20*(10), 1–37. https://doi.org/10.3102/0034654314546954.

Feynman, R. P., Leighton, R., & Hutchings, E. (1985). *"Surely you're joking, Mr. Feynman!" : Adventures of a curious character*. New York: W.W. Norton.

Furberg, A. (2016). Teacher support in computer-supported lab work: Bridging the gap between lab experiments and students' conceptual understanding. *International Journal of Computer-Supported Collaborative Learning, 11*(1), 89–113.

Gerard, L. F., & Linn, M. C. (2016). Using automated scores of student essays to support teacher guidance in classroom inquiry. *Journal of Science Teacher Education, 27*(1), 111–129. https://doi.org/10.1007/s10972-016-9455-6.

Gerard, L. F., Ryoo, K., McElhaney, K., Liu, L., Rafferty, A. N., & Linn, M. C. (2015). Automated guidance for student inquiry. *Journal of Educational Psychology, 108*(1), 60–81.

Gerard, L. F., Linn, M. C., & Madhok, J. (2016). Examining the impacts of annotation and automated guidance on essay revision and science learning. In C.-K. Looi, J. Polman, U. Cress, & P. Reimann (Eds.), *International conference of the learning sciences* (Vol. 1, pp. 394–401). Singapore: International Society of the Learning Sciences.

Hamalainen, R., & Vahasantanen, K. (2011). Theoretical and pedagogical perspectives on orchestrating creativity and collaborative learning. *Educational Research Review, 6*(3), 169–184.

Harrison, E. J., Gerard, L. F., & Linn, M. C. (2018). Encouraging revision of scientific ideas with critique in an online genetics unit. Paper presented at the rethinking learning in the digital age: Making the learning sciences count, 13th international conference of the learning of the learning sciences (ICLS), London.

Herrenkohl, L. R., Tasker, T., & White, B. (2011). Pedagogical practices to support classroom cultures of scientific inquiry. *Cognition and Instruction, 29*(1), 1–44.

Ingulfsen, L., Furberg, A., & Stromme, T. (2018). Students' engagement in real-time graphs in CSCL settings: Scrutinizing the role of teacher support. *International Journal of Computer-Supported Collaborative Learning, 13*(4), 365–390.

Isaacson, W. (2017). *Leonardo da Vinci*. New York: Simon & Schuster.

Jimenenez-Aleixandre, M. P., Rodriguez, A. B., & Duschl, R. A. (2000). "Doing the lesson" or "doing science": Argument in high school genetics. *Science Education, 84*(3), 757–792.

Kali, Y., Linn, M. C., & Roseman, J. E. (Eds.). (2008). *Designing coherent science education: Implications for curriculum, instruction, and policy*. New York: Teachers College Press.

Lakkala, M., Lallimo, J., & Hakkarainen, K. (2005). Teachers' pedagogical designs for technology-supported collective inquiry: A national case study. *Computers & Education, 45*, 337–356. https://doi.org/10.1016/j.compedu.2005.04.010.

Linn, M. C., & Eylon, B.-S. (2011). *Science learning and instruction: Taking advantage of technology to promote knowledge integration*. New York: Routledge.

Liu, O. L., Lee, H. S., Hofstetter, C., & Linn, M. C. (2008). Assessing knowledge integration in science: Construct, measures and evidence. *Educational Assessment, 13*(1), 33–55.

Liu, O. L., Lee, H.-S., & Linn, M. C. (2011). Measuring knowledge integration: A four year study. *Journal of Research in Science Teaching, 48*(9), 1079–1107.

Liu, O. L., Ryoo, K., Linn, M. C., Sato, E., & Svihla, V. (2015). Measuring knowledge integration learning of energy topics: A two-year longitudinal study. *International Journal of Science Education, 37*(7), 1044–1066. https://doi.org/10.1080/09500693.2015.1016470.

Liu, O. L., Rios, J. A., Heilman, M., Gerard, L. F., & Linn, M. C. (2016). Validation of automated scoring of science assessments. *Journal of Research in Science Teaching, 53*(2), 215–233. https://doi.org/10.1002/tea.21299.

Matuk, C. F., & Linn, M. C. (2015). Examining the real and perceived impacts of a public idea repository on literacy and science inquiry. In O. Lindwall, P. Hakkinen, T. Koschmann, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference* (Vol. 1, pp. 150–157). Gothenburg: International Society of the Learning Sciences.

Matuk, C., & Linn, M. C. (2018). Why and how do middle school students exchange ideas during science inquiry. *International Journal of Computer-Supported Collaborative Learning, 13*, 1–37. https://doi.org/10.1007/s11412-018-9282-1.

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34*, 57–111. https://doi.org/10.1017/S0140525X10000968.

Perin, D., Lauterbach, M., Raufman, J., & Santikian Kalamkarian, H. (2016). Text-based writing of low-skilled adults: Relation to comprehension, self-efficacy and teacher judgments. *Reading and Writing: An Interdisciplinary Journal, 30*(4), 887–915. https://doi.org/10.1007/s11145-016-9706-0.

Raes, A., Schellens, T., & De Wever, B. (2013). Web-based collaborative inquiry to bridge gaps in secondary science education. *Journal of the Learning Sciences, 22*, 316–347. https://doi.org/10.1080/10508406.2013.836656.

Reiser, B. J. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas: National Research Council, board on science education, division of behavioral and social sciences and education*. Washington, DC: National Academies Press.

Roshcelle, J., Dimitriadis, Y., & Hoppe, U. (2013). Classroom orchestration: Synthesis. *Computers and Education, 69*, 523–526.

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching, 44*(1), 57–84.

Ruiz-Primo, M. A., & Li, M. (2013). Analyzing Teachers' feedback practices in response to Students' work in science classrooms. *Applied Measurement in Education, 26*(3), 163–175. https://doi.org/10.1080/08957347.2013.793188.

Ryoo, K., Bedell, K., & Swearingen, A. (2018). Promoting linguistically diverse students' short-term and long-term understanding of chemical phenomena using visualizations. *Journal of Science Education and Technology, 27*(6), 508–522.

Sharples, M. (2013). Shared orchestration within and beyond the classroom. *Computers and Education, 69*, 504–506.

Sinha, S., Rogat, T. K., Adams-Wiggins, K. R., & Hmelo-Silver, C. E. (2015). Collaborative group engagement in a computer-supported inquiry learning environment. *International Journal of Computer Supported Collaborative Learning, 10*(3), 273–307.

Sisk-Hilton, S. (2009). *Teaching and Learning in Public: Professional Development Through Shared Inquiry.* Columbia: Teachers College Press.

Songer, N. B. (1996). Exploring learning opportunities in coordinated network-enhanced classrooms - a case of kids as global scientists. *Journal of the Learning Sciences, 5*(4), 297–327.

Sun, D., Looi, C.-K., & Xie, W. (2016). Learning with collaborative inquiry: A science learning environment for secondary students. *Technology, Pedagogy and Education, 26*(3).

Tansomboon, C., Gerard, L. F., Vitale, J. M., & Linn, M. C. (2017). Designing automated guidance to promote productive revision of science explanations. *International Journal of Artificial Intelligence in Education, 27*(4), 729–757.

Thagard, P. (1992). *Conceptual Revolutions*. Princeton: Princeton University Press.

Tissenbaum, M., Lui, M., & Slotta, J. D. (2012). *Smart Classrooms for Knowledge Communities: Learning Across Contexts in Secondary Science* Paper presented at the American Educational Research Association annual meeting, Vancouver, BC.

van Zee, E. H., & Minstrell, J. (1997). Reflective discourse: Developing shared understandings in a physics classroom. *International Journal of Science Education, 19*(2), 209–228.

Vitale, J. M., Lai, K., & Linn, M. C. (2015). Taking advantage of automated assessment of student-constructed graphs in science. *Journal of Research in Science Teaching, 52*(10), 1426–1450. https://doi.org/10.1002/tea.21241.

Williams, M., Linn, M. C., Ammon, P., & Gearhart, M. (2004). Learning to teach inquiry science in a technology-based environment: A case study. *Journal of Science Education and Technology, 13*(2), 189–206.

Zertuche, A. N., Gerard, L. F., & Linn, M. C. (2012). How do openers contribute to student learning? *International Electronic Journal of Elementary Education, 5*(1), 79–92.

Zheng, B., Lawrence, J., Warschauer, M., & Lin, C.-H. (2015). Middle school Students' writing and feedback in a cloud-based classroom environment. *Technology, Knowledge and Learning, 20*(2), 201–229.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.