Encouraging Revision of Scientific Ideas with Critique in an Online Genetics Unit

Emily J. Harrison, Libby Gerard, and Marcia C. Linn
emilyharrison@berkeley.edu, libby.gerard@gmail.com, mclinn@berkeley.edu
University of California, Berkeley

Abstract: Encouraging students to revise their scientific ideas after encountering new evidence is essential to science learning, but has proven challenging. We investigated the merit of critique in promoting revision. 315 students participated in an online genetics unit where we investigated how critique affects the nature and frequency of revisions made to student essays. Students in the *critique* condition explain what is wrong or missing from several common non-normative student ideas regarding difficult topics in genetics. We compare this to a method used in the past to encourage students to add new ideas to their essays; students in the *revisit* condition are directed back to relevant information and interactive models rather than practicing critique. Students in the *critique* condition were more likely to revise their essays at all, especially students with low prior knowledge. Students that practiced critique were also significantly more likely to add new ideas to their revisions.

Keywords: science, genetics, critique, revision, technology, knowledge integration

Introduction

This research investigates how two types of guidance influence student revision of their scientific explanations. Revision of scientific explanations and arguments is an essential practice in learning and communicating science (Brownell et al., 2013). Revision is also stressed prominently by the NGSS. The NGSS were developed around the idea that students continually build on and revise their knowledge (NGSS, 2013). The NGSS science practices *constructing explanations, engaging in argument from evidence,* and *obtaining, evaluating, and communicating information* describe an iterative process of incorporating new ideas and evidence into continually constructed scientific knowledge. Although revision is a necessary skill for gaining the integrated understanding called for by the NGSS, there are limited opportunities for students to revise in a typical science classroom (Berland & Reiser, 2011). Furthermore, students are challenged by NGSS to make productive revisions. Bridwell (1980) analyzed 12th grade students' English essay revisions for changes in surface features, words, phrases, clauses, sentences, multiple sentences, and text. All students revised, but most revisions occurred at the word (31.2%) or surface level (24.8%). Students revised primarily by improving word choice and by correcting mechanical errors. Crawford et al. (2008) reported similar findings for 5th and 8th grade students, who focused on revising words and punctuation.

Previous studies suggest that encouraging revision is indispensable to student learning, and promoting revision in any form can be beneficial for students. For example, Tansomboon et al. (2017) revealed that students who attempted to revise their science explanations made significantly greater gains from pre to posttest than those who made no revisions at all, even if the revisions made were incorrect. In addition, students that made relevant revisions gained a more integrated understanding of the science material (Gerard et al., 2015; Linn et al., 2014). Previously, in an online unit on plate tectonics, we developed a rubric to categorize the types of revisions that students made on embedded essays. We found that students who made integrated revisions demonstrated greater gains on a posttest essay item, as compared to students that did not revise at all or simply added disconnected ideas.

This study investigates how engaging students in critique can promote productive revision of student written explanations. We conducted the study in the context of a Web-based Inquiry Science Environment (WISE) unit on the topic of genetics and simple inheritance at the middle school level. We compare critique activities to having students revisit relevant material and interactive models to gain new ideas. Both strategies promote knowledge integration by encouraging students to look for concepts that are missing from their explanations and incorporate them into their existing ideas. This unit was designed according to the Knowledge Integration (KI) framework, which encourages making valid and coherent connections between scientific concepts as well as using evidence and reasoning (Linn and Eylon, 2011). This framework *elicits* students' prior knowledge in order to build on their ideas, and promotes *adding* and *distinguishing* ideas, and finally *reflecting* on newly constructed knowledge (Linn and Eylon, 2006). These steps make this framework ideal for supporting students in revising their scientific ideas. Supporting students through challenging practices, such as critique and revision, can be difficult in the context of new and demanding disciplinary content (Scheuer et al. 2009); technology in the WISE platform is ideal for supporting these two tasks simultaneously.

Based upon the NGSS practices and a review of the literature, we decided to test the benefit of critique for encouraging students to make productive revisions to their explanations and arguments for several reasons. Critique in this context is a detailed analysis and assessment of a claim or theory; it is necessary for the process of evaluating scientific information and evidence, as well as constructing and revising explanations. Argument-based interventions have become popular, given its propensity to foster scientific literacy (Cavagnetto, 2010). However, Henderson et al. (2015) purports that argumentation without critique limits opportunities for students to engage in authentic scientific reasoning in the classroom. The construction of knowledge involves a continuous cycle of construction and critique (Ford and Forman, 2006). Critique is not simply an exercise, it is fundamental for epistemic vigilance and critical thinking (Henderson et al., 2015). Written arguments and explanations depend heavily on content knowledge, but explicitly supporting students in skills such as critique has been shown to increase the complexity of written explanations despite new and challenging content material (Berland and McNeill, 2010). This has been demonstrated repeatedly in the context of genetics (Jimenez-Aleixandre and Duschl, 1999; Zohar and Nemet, 2001), making our WISE Genetics unit an ideal context for this study. This unit covers content contained in the science standards, while also delving into aspects of genetics pertinent to general scientific literacy. Due to the growing presence of genetics in media and the public interest, developing and continually refining and evaluating one's knowledge of genetics and the mechanism of inheritance is of increasing importance.

We designed two randomly assigned conditions that student groups participated in to assess the value of critique versus revisiting. The *revisit* condition reintroduced a relevant interactive model and prompted them to answer new sets of questions designed to elicit new ideas that may not have initially been ascertained from the model. The goal of this condition was to allow students to directly test their explanations after writing them, and was modeled after past guidance used in WISE units that direct students back to relevant material to gain new insights (Donnelly et al., 2015). The *critique* condition prompted students to explain what was incorrect or missing from several non-normative statements rather than revisiting the interactive model. These non-normative statements were developed from ideas that students generally found to be the most challenging. The goal of this activity was to encourage students to consider flaws in their own reasoning, explicate them, and then incorporate these newly distinguished ideas into their own revisions. The activity was designed to help students unpack the complexity of constructing and revising scientific ideas by focusing explicitly on the practice of critique. Throughout the two-week-long unit, students encountered their assigned activity, *critique* or *revisit* (see Table 1 for outline of conditions), while the rest of the unit was identical for all students.

Table 1: Outline of a sample sequence of *critique* and *revisit* conditions

<table>
<tr><th>Critique</th><th>Revisit/ Experiment</th></tr>
<tr><td colspan="2"><u>Essay Prompt</u>: Explain how you would use a Punnett square to figure out the probability of getting a certain genotype.</td></tr>
<tr><td><u>Critique</u>
Example Critique:
Student 1: "Their 4th child will have attached earlobes."
Explain how this statement is incorrect or too vague, and how to make it more accurate.</td><td><u>Revisit</u>
Use the Punnett square model again to answer the questions below. (Model is embedded on the next page, along with the questions)
Example Question:
When both parents have the genotype EE, what is the probability of having a child with attached earlobes? Explain.</td></tr>
<tr><td colspan="2"><u>Revise</u>: Now that you've learned a bit more about Punnett squares and probability, take some time to revise or improve your answer to this question from earlier.
What is a method you would use to calculate probability of getting a certain genotype using a Punnett square?
(Students' original responses are imported automatically)</td></tr>
</table>

Methods

Participants and procedures

Three teachers from two middle schools participated in this study, with a total of 13 classes of 8th grade students (372 students, 198 student workgroups). Teacher 1 taught 4 classes (120 students) at the first school (49% non-white, 32% free/reduced lunch, 7% ELL). Teachers 2 and 3 taught at the second school (62% non-white, 22% free/reduced lunch, 12% ELL). Teacher 2 taught 4 classes (111 students), and Teacher 3 taught 5 classes (141 students). Students completed the 8 day Genetics WISE unit during 50 minute class periods. Students worked in collaborative workgroups assigned by their teachers, mostly pairs with a few students working individually or in groups of 3. Students completed the pretest one day before beginning the unit, and the posttest one day after completing the unit. Both pre and posttest were completed individually. Of the total students, 315 completed both the pre and posttest, as well as (most of) the Genetics unit. Workgroups were randomly assigned to one of the two conditions (*critique* or *revisit*) by the computer based on their WISE Workgroup ID.

Curricular materials

This study builds on a Web-based Inquiry Science Environment (WISE) unit on genetics and simple inheritance. The unit features “critique” and “revisit” activities. It is framed around 4 essential questions related to genetics: (a) Why do we look the way we do? (b) How can we predict disease? (c) How do mutations affect DNA? (d) How do we control our world with genetics? To address these essential questions, it is important to understand both the mechanism of simple inheritance and that inheritance is rarely simple. This unit does this by employing manipulatable computer models and various interactive question types to help students understand the complexity of genetics while simplifying situations into solvable problems. Throughout the unit, learning goals include independent assortment of alleles, dominance and recessivity, probability of inheriting certain traits depending on parental genotypes, tracking of alleles through several generations using a pedigree, the relationship of DNA to phenotypic expression, and how the environment and human influence leads to change in genetic expression and inheritance. The WISE platform is able to track and save student work from various assessments embedded throughout the unit.

The unit features an interactive punnett square model to help students see the effects of allele combinations on specific phenotypes (Figure 1). It also includes drag and drop questions to help students sort evidence and receive immediate feedback, and interactive graphing tools to allow students to construct visualizations of data illustrating connections between different genotype crosses. Activities were added that focus on genetic modification, both through artificial selection and engineering, as well as common mutations and their effects.

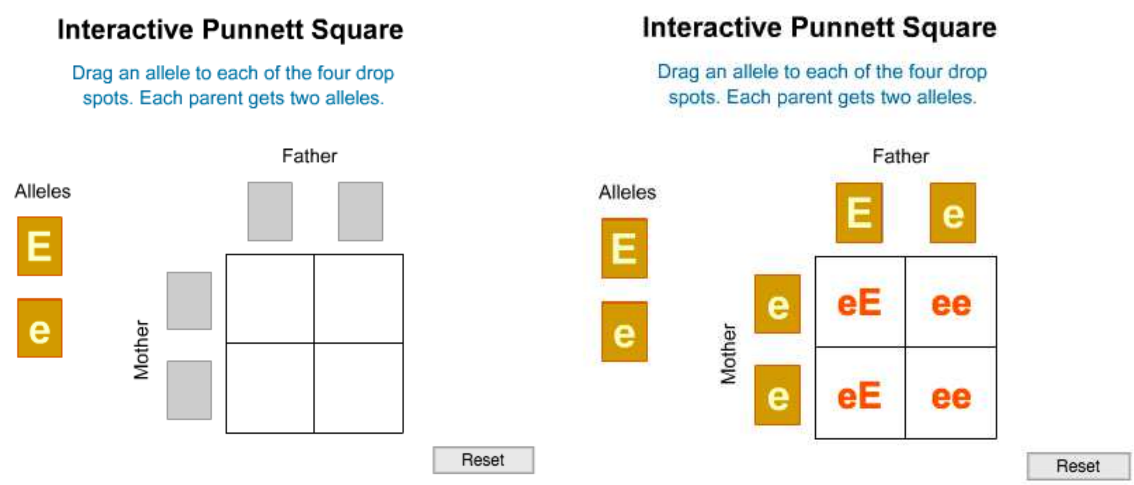

Figure 1. Screenshots of the interactive Punnett Square model; students drag different alleles to populate the boxes and test different combinations.

Assessments

The pretest and posttest were used as measures of student knowledge integration, looking at students’ prior knowledge and relative improvement. Items generally asked students for a written explanation that required synthesis of several genetics concepts. For example, one pretest question (*SiblingsPrePost)* prompted: "Siblings look similar, but not exactly the same unless they are identical twins. If they inherited their DNA from the same parents, why don't siblings look exactly the same? Explain." A critique question was also included on the pre and posttest; Students were given a completed punnett square showing two parents heterozygous for brown/blue eyes and were asked to predict the probability of having a child with blue eyes. The follow up critique question (*CritPrePost)* prompted: "Another student said: 'From the Punnett square, you can tell that the couple's 4th child will have blue eyes.' Do you agree or disagree with this statement? Explain."

We scored students' initial and revised explanations using KI rubrics to measure the value of critique versus revisit guidance. Specifically, we looked at the types of revisions students made on two embedded KI synthesis essay questions. The first question was identical to the first pre/post question regarding the mechanism of inheritance: Why do siblings from the same parents look similar but not exactly the same? (*Siblings*). The second question focused on the use of a punnett square to determine the probability of getting a certain genotype in various cases (*PunnettSquare*): "What is a method you would use to calculate probability of getting a certain genotype using a Punnett square?"

Analysis approach

Student essay responses on the pre and posttest, as well as embedded essay items, were scored using a 5-point Knowledge Integration (KI) scale (See Table 2 for sample rubric). KI scoring is designed to reward students for making connections between ideas, thereby integrating new information with their prior knowledge (Linn & Eylon, 2011; Liu et al, 2008). The rubric for the *Siblings* question demonstrates how links are scored (Table 2).

Table 2: KI Rubric Example: Why do siblings from the same parents look similar but not exactly the same?

KI Score	Description	Examples
1	No Answer	"I don't know"
2	Non-normative/irrelevant: Token mechanism only ("skips a generation") with no elaboration. Incorrect ideas: "you get different amounts of DNA from each parent"	Because you and your sibling have close genes but they are not the same genes. You inherit similar amounts of the same traits from the same parents at slightly different amounts. Because it's not exactly the same, you look a little different.
3	Partial link (one correct statement, but not connected to other scientific ideas, or student does not elaborate)	They get different parts of dna from their parents. You get a different set of genes then your sibling.
4	One full link between normative scientific ideas	Because you get half of your parents DNA but it does not specify which half you will inherit from them. This means that the half that you might get will not be the same that your sibling will get.
5	At least two full links	Siblings do not look exactly the same because they have slightly different alleles. Each child has a chance of receiving a different allele from its parents than its sibling because of probability.

Qualitative revision codes were also given to the embedded essay revisions based on how a student revised. A code was given for whether students made connected (C) or disconnected (D) revisions. Another code was given for whether students added new (N) ideas in their revision or expanded existing (E) ideas that were already present in their initial response (see Table 3 for examples of each code combination).

Table 3**:** Rubric and examples for embedded essay revisions (Student revisions underlined)

Initial Response	**Revision**	**Score (C/D) (N/E)**	
You would use the method of counting by 25's. Each square is a 25% chance.	You would use the method of counting <u>by quarters. The two letters from each parent would represent a quarter of the genotype and all the quarters combined would show what phenotype would be dominant over the other.</u>	C	N
A method I would use is that I would choose the number of squares that have a certain genotype and find out what percent of the squares have that genotype.	A method I would use <u>to calculate the probability of getting a certain genotype</u> is that I would choose the number of squares that have a certain genotype and find out what percent of the squares have that genotype.	C	E
You would put the alleles of each parent on the outside of the square and the possible alleles for their children would be in the square.	You <u>can find the dominant and recessive traits to calculate the probability out of four.</u>	D	N
The figure shows how many possible genotypes that children can have.	The figure shows how many possible genotypes that children can have. <u>D means dimples and d means no dimples.</u>	D	E

Results

All teachers implemented the unit as planned, during 50 minute class periods over the course of two weeks. Teachers intermittently reminded students of guidelines for productive collaboration with their partners throughout the project.

Pre/posttest analysis

<u>Learning gains</u>

Students began the unit with moderately low genetics knowledge, with an average pretest KI score of 2.47 (SE=0.04). Students in the *critique* and *revisit* conditions were at similar levels at the beginning of the unit [*Critique*: Pre: M=2.50 SE=0.05; *Revisit*: Pre: M=2.44, SE=0.07].

All students achieved learning gains in genetics by the posttest (*SiblingsPrePost* and *CritPrePost*) [Post: M=3.19, SE=0.05]. These items were chosen for analysis because they require the synthesis of various genetics concepts. The item *SiblingPrePost* involves independent assortment of alleles, leading to offspring from the same parents receiving different combinations of alleles, and different genotypes resulting in different phenotypes. The item *CritPrePost* asks students to explain how the probability of a certain trait can be predicted using a punnett square, and involves a critique component. There was no difference in gain between conditions from pre to posttest on these items [*Critique*: t(148)=8.42, Post: M=3.13 SE=0.07, p<0.001; *Revisit*: t(119)=9.80, Post: M=3.28, SE=0.08, p<0.001].

<u>Interactions with prior knowledge</u>

A regression analysis revealed an interaction between condition and prior knowledge on pre/post gain. We separated students into high or low prior knowledge groups based on their pretest score on each of the two items; those that scored a 1-2 were considered low, and those that scored 3-5 were considered high prior knowledge. We chose this cutoff because students must include at least one normative scientific idea to achieve a score of 3. Our analysis revealed that students with low prior knowledge scored an estimated average of 0.36 points lower in posttest gain in the critique condition (p<0.05) compared to students with high prior knowledge.

This suggests that high prior knowledge students are more likely to benefit from critique activities in terms of pre/post gains. This may be because success at critique depends on understanding the content material.

Revision on embedded assessments

Nature of student revisions and learning gains

Overall we found that critique motivated adding more new ideas to essay revisions than did revisiting relevant material.

For revisions on the *Siblings* embedded essay question, students in the *critique* condition added new ideas an estimated 1.84 times as often as students in the *revisit* condition, on average [$z(198)=2.05$, $p<0.05$]. Condition had no effect on KI score gain from initial to revised response for this item. However, students that did revise on this item scored, on average, 0.41 points higher on their revised essay than students who kept their original answer (did not revise) [$t(196)=3.40$, $p<0.001$]. This suggests that students that revised made productive changes to their essays on this item.

For the *PunnettSquare* embedded essay, students in the *critique* condition were 2.74 times as likely, on average, to add new ideas to their revisions compared to students in the *revisit* condition [$z(198)=2.72$, $p<0.01$]. In addition, students in the *critique* condition scored an estimated 0.20 points higher on this item than students in the *revisit* condition from initial to revised response [$t(196)=2.54$, $p<0.05$]. This is likely due to the addition of new relevant science ideas in their revisions.

For both embedded essay questions, we found no effect of condition on students making connected revisions.

Frequency of revision

Critique condition students were overall more likely to revise their essays, compared to students in the revisit condition. Students in the *critique* condition were 2.75 times as likely to revise at least one of the two embedded essays as compared to students in the *revisit* condition [$z(198)=3.25$, $p<0.001$].

Prior knowledge was a significant factor in deciding to revise as well; students with high prior knowledge were 2.28 times as likely to revise at least one of the two embedded essays compared to students with low prior knowledge [$z(195)=2.13$, $p<0.05$]. In order to examine the effect of prior knowledge further, we separated students again into high and low prior knowledge groups. High prior knowledge students in the *critique* condition were 2.22 times as likely to revise at least one of the essays compared to high prior knowledge students in the *revisit* condition [$z(159)=2.26$, $p<0.05$]. Students with low prior knowledge in the *critique* condition were 5.2 times as likely to revise compared to low prior knowledge students in the *revisit* condition [$z(36)=2.27$, $p<0.05$]. Therefore, not only did the *critique* condition promote revision more often, it especially encouraged students with low prior knowledge (an initial essay score of 1 or 2) to revise more often.

Qualitative critique analysis

These results suggest that both conditions were effective in promoting student understanding of genetics and simple inheritance from pre to posttest. While low prior knowledge students in the critique condition gained slightly less than high prior knowledge students, the critique activities still encouraged them to revise their responses more often than the revisit condition. For revisions, critique was helpful for adding new scientific ideas, but these ideas were not necessarily connected to their previous responses. We present examples of revisions that students in the critique condition made in order to see how ideas were being added (see Table 4).

Table 4: Revisions from students in the critique condition (revisions in **bold**)

Initial Response	**Revision after Critique**
His grandparents passed it to his parents, but the disease skipped a generation	His grandparents passed **down one recessive and one dominant gene** to his parents **and then they both gave him their recessive genes so he would have cystic fibrosis.**
Eric inherited cystic fibrosis because his grandparent had it and it could very possibly skip a generation. Eric's mother most likely recieved the dominant gene which was not	Eric inherited cystic fibrosis because his grandparent had it and **[ ]** it skipped their generation. Eric's mother most likely recieved the dominant gene **from her mother and a recessive gene from her father which was hidden and**

getting cystic fibrosis.	**allowed her** not to get cystic fibrosis.
Eric might have inherited cystic fibrosis by his grandpa's genes skipping over his parents and going straight to him.	Eric might have inherited cystic fibrosis by his grandpa's genes skipping over his parents and going straight to him. **His parents could have a heterozygous genotype which would give Eric a 25% chance of developing the disease.**
We think that siblings look similar to each other but not exactly the same because the traits of the parents are different and each child gets different traits from each parent.	We think that siblings look similar to each other but not exactly the same because **they have different combinations of alleles from their parents.**

These examples illustrate how students in the critique condition distinguished their ideas. Rather than keeping their vague answers (ex. "each child gets different traits"), they revised to include more specific scientific vocabulary and more nuanced definitions of phenomena (such as "a heterozygous genotype which could give Eric a 25% chance"). Distinguishing vague ideas to give more detailed scientific explanations is an important part of the KI framework. In this study, the critique condition showed promise in helping students achieve this difficult task.

Conclusions and implications

Revising ideas after encountering new information is an essential scientific literacy skill. This study shows the benefit of critique activities in promoting revision of scientific ideas.

Both conditions in our study were effective in helping all students achieve learning goals in genetics and simple inheritance. Further analysis revealed advantages for critique regarding motivating revision more frequently. This is likely because these students were exposed to flaws in their own thinking by analyzing common incorrect ideas, motivating them to rethink and clarify their original responses. Students in the *critique* condition also regularly added more new ideas to their essay revisions. This is likely because students had to consider their logic more carefully while critiquing, encouraging them to distinguish between ideas, whereas the models revisited in the other condition did not explicitly encourage students to think about the mechanisms of inheritance, such as that of allele movement. While we hope to create guidance and activities that encourage students to make more integrated revisions, studies have found that even attempts at revision have been shown to result in greater learning gains (Tansomboon et al., 2017). Our critique activity was successful at motivating students to at least attempt to revise their ideas more often, especially those students with low prior knowledge in the content area.

Overall, revisions that students made in this unit were highly relevant, and attempted to add value to their responses in the form of new or better-clarified ideas. This is in distinct contrast to studies, including Bridwell (1980) and Crawford et al. (2008), that found most student revisions were occurring at the word or surface level. This again promotes the practice of critique in encouraging students to revise their scientific ideas rather than just their grammar.

Future studies will look into ways to refine our critique activity to further support students with low prior knowledge in making greater gains. Improvement of our interactive models can also be made to more clearly depict the mechanism of inheritance. We will also investigate further the mechanism for critique motivating students to revise more often, and ways to encourage students to make more connected, integrated revisions of their scientific ideas.

References

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. Science Education, 94(5), 765–793. http://doi.org/10.1002/sce.20402

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

Bransford, J. D., Brown, A., & Cocking, R. (1999). How people learn: Mind, brain, experience, and school. Washington, DC: National Research Council.

Bridwell, L. (1980). Revising strategies in twelfth grade students' transactional writing. Research in the Teaching of English, 14, 197–222.

Brown, A. L., & Campione, J. C. (1994). Guided discovery in a community of learners. The MIT Press.

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.

Cavagnetto, A. R. (2010). Argument to Foster Scientific Literacy: A Review of Argument Interventions in K-12 Science Contexts. Review of Educational Research, 80(3), 336–371. http://doi.org/10.3102/0034654310376953

Crawford, L., Lloyd, S., Knoth, K.: (2008) Analysis of student revisions on a state writing test. Assessment for Effective Intervention, 33, 108–119.

Donnelly, D. F., Vitale, J. M., & Linn, M. C. (2015). Automated Guidance for Thermodynamics Essays: Critiquing Versus Revisiting. Journal of Science Education and Technology, 24(6), 861–874. doi:10.1007/s10956-015-9569-1.

Ford, M., & Forman, E. A. (2006). Redefining disciplinary learning in classroom contexts. Review of Research in Education, 30, 1-32.

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). doi:10.1037/edu0000052

Henderson, J. B., MacPherson, A., Osborne, J., & Wild, A. (2015). Beyond Construction: Five arguments for the role and value of critique in learning science. International Journal of Science Education, 37(10), 1668–1697. http://doi.org/10.1080/09500693.2015.1043598

Jiménez Aleixandre, M. P., Bugallo Rodríguez, A., & Duschl, R. A. (2000). “Doing the lesson” or ‘doing science’: Argument in high school genetics. Science Education, 84(6), 757–792. http://doi.org/10.1002/1098-237X(200011)84:6<757::AID-SCE5>3.0.CO;2-F

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

Linn, M. C., & Eylon, B.-S. (2006). Science education: Integrating views of learning and instruction. In P. A. Alexander, & P. H. Winne (Eds.), Handbook of educational psychology (2nd Ed., pp. 511-544). Mahwah, NJ: Lawrence Erlbaum Associates.

Linn, M. C., Gerard, L., Ryoo, K., McElhaney, K., Liu, O. L., & Rafferty, A. N. (2014). Computer-guided inquiry to improve science learning. Science, 344(6180), 155-156.

NGSS Lead States. (2013). Next Generation Science Standards: For States, By States. Washington, DC: The National Academies Press.

Tansomboon, C. (2017). Design of Automated Guidance to Support Effortful Revisions and Knowledge Integration in Science Learning (Doctoral dissertation, University of California, Berkeley).

Zohar, A., & Nemet, F. (2001). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. Journal of Research in Science Teaching, 39(1), 35–62. http://doi.org/10.1002/tea.10008