High School Students' Collaboration and Engagement With Scaffolding and Information as Predictors of Argument Quality During Problem-Based Learning

Brian R. Belland, Nam Ju Kim, David M. Weiss, and Jacob Piland
brian.belland@usu.edu, namju_1001@gmail.com, dmark.weiss@usu.edu, jacob.piland@aggiemail.usu.edu
Utah State University

Abstract: Strong information literacy, collaboration, and argumentation skill are essential to success in problem-based learning (PBL). Computer-based scaffolding can play a key role in helping students enhance these skills during PBL. We examined how information literacy, collaboration, and time spent in various scaffolding sections combine to predict argument quality, and qualitative analysis from the social cognitive framework perspective to explain why significant variables predicted argumentation score. Quantitative results indicated that information literacy, time spent doing individual work, and time spent on the scaffold stages *define problem* and *link evidence to claims* significantly predicted argument quality. Qualitative results suggest that there was little connection between the content of the written argument and what students wrote in the scaffold when students spent more time in individual work. Results are discussed in light of the literature.

Conceptual framework

The Next Generation Science Standards and the Common Core invite educators to enhance K-12 students' (USA students aged 5-18) problem solving and argumentation skills (Krajcik, Codere, Dahsah, Bayer, & Mun, 2014; McLaughlin & Overturf, 2012). One way to do this is to have students address authentic, ill-structured problems within the framework of problem-based learning (PBL; Hmelo-Silver, 2004). In PBL, students work in small groups to define the problem, determine, gather and synthesize needed information, make a claim about problem resolution, and link the claim to evidence (Barrows, 1985; Belland, Glazewski, & Richardson, 2008; Hmelo-Silver, 2004). Meta-analyses indicate that PBL is particularly helpful in enhancing principles-level skill - knowledge of rules that dictate the direction and strength of relationships between concepts (Gijbels, Dochy, Van den Bossche, & Segers, 2005; Sugrue, 1995), and application level skill - ability to use concept- and principles-level knowledge to address new problems (Sugrue, 1995; Walker & Leary, 2009).

Two key factors contribute to success in PBL - argumentation ability (Kuhn, 2010) and information literacy (Hakkarainen, 2009). Argumentation ability can be defined as the ability to lead an audience to accept a claim as valid by linking evidence to the claim by way of premises (Perelman & Olbrechts-Tyteca, 1958). Many K-12 students have limited argumentation skills, as evidenced by tendencies to (a) fail to provide evidence to back up claims (Belland et al., 2008) and (b) back up claims with explanations of the claim (Glassner, Weinstock, & Neuman, 2005), irrelevant evidence, and/or evidence of questionable validity (Kuhn, 2010). Argumentation ability is crucial in PBL because students need to be able to (a) provide evidence that their solutions are well reasoned, and (b) weigh other's arguments for their solutions (Belland et al., 2008). Information literacy refers to students' abilities to identify credible/relevant sources related to a topic, and weigh the credibility and relevance of gathered information when synthesizing research (Forte, 2015; Van de Vord, 2010). K-12 students often lack sufficient information literacy, and this can result in them backing claims with evidence of questionable credibility and/or relevance (Forte, 2015; Kuiper, Volman, & Terwel, 2005).

Within PBL, effective collaboration is central to student success (Arts, Gijselaers, & Segers, 2002). Effective collaboration arises when (a) shared work addresses a shared learning goal, and (b) students offer ideas to the group and engage with each other's ideas in a critical but constructive manner (Rojas-Drummond & Mercer, 2003), and (c) synergy among group members' efforts results in output superior to what would result from adding all group member's individual efforts (Schmidt, Rotgans, & Yew, 2011).

Scaffolding serves an important role in structuring and problematizing problem solving (Reiser, 2004), and can be defined in terms of strategies (e.g., enlisting interest, indicating important problem elements to consider, and questioning) and form (e.g., expert modeling, question prompts, and concept mapping; van de Pol, Volman, & Beishuizen, 2010; Wood, Bruner, & Ross, 1976). Some research suggests that students engage with scaffolding on the basis of their goals and the affordances that they perceive in the scaffolding (Belland & Drake, 2013). This depends on students' ability to be agentic (Bandura, 1986). The allowance for student agency also means that students can engage with scaffolding as a whole, and specific scaffolding elements to various degrees, and this can influence the quality by which they address the problem (Oliver & Hannafin, 2000).

Meta-analyses indicated that scaffolding in STEM education led to an average between-subjects effect of $g = 0.46$ (Belland, Walker, Kim, & Lefler, In Press) and an average pre-post effect of at least $g = 0.7$ across concept, principles, and application assessment levels, indicating that scaffolding can produce strong gains in cognitive outcomes ranging from declarative knowledge to problem solving and argumentation (Belland, Walker, & Kim, Under review; Sugrue, 1995).

Research questions

1. Do the time that students spend in each scaffolding stage, time spent working individually, time spent working collaboratively, and information literacy scores predict argument quality?
2. How and why is argument quality explained by information literacy and student engagement with a computer-based scaffold?
3. What choices do high school students make as they construct arguments in PBL with the support of a computer-based scaffold, and why?

Method

Participants and setting

The study took place in one class section of environmental sciences in a medium high school in a rural setting in the Intermountain west. Twenty-four 10th grade and 11th grade students participated. The unit centered on how to improve soil quality within the county. The teacher, with 30 years of experience teaching science, had been previously exposed to computer-assisted problem-based learning in a summer school class for credit recovery students (Belland, Gu, Weiss, & Kim, 2016). Students brought to class soil samples from their houses, which they then analyzed for nitrogen, potassium, phosphorus, calcium, copper, iron, and silt/clay consistency. Working in groups of 3-4, they needed to address what could be done to optimize soil quality for specific beneficial uses (e.g., growing vegetables) and mitigate potential problems (e.g., erosion). They then needed to argue what should be done to optimize soil quality by drawing on research on the soil samples and soil quality elements that are essential to specific beneficial uses, and how to enhance such elements. The teacher provided one-to-one, dynamic scaffolding to complement the computer-based scaffolding.

Design

We took a sequential, mixed methods approach to data analysis. Quantitative analysis in the form of Bayesian regression came first, and qualitative analysis was used to explain mechanisms underlying the relationships uncovered in Bayesian regression (Onwuegbuzie, Slate, Leech, & Collins, 2009).

Materials

The *Connection Log* is a database-driven web application designed to scaffold middle and high school students' construction of evidence-based arguments during PBL (Belland, Gu, Armbrust, & Cook, 2015). The *Connection Log* invites students to define the problem, determine needed information, find and organize needed information, make a claim, and link evidence to claim (Belland et al., 2015). In each stage (e.g., make claim), students (a) articulate ideas individually and (b) come to consensus with groupmates to maximize perspectives and enhance collaboration. Consensus entries were entered by a group member designated scribe. The *Connection Log's* support can be classified as strategic and metacognitive (Hannafin, Land, & Oliver, 1999), and served to structure and problematize the problem solving and argumentation process (Reiser, 2004). As with CSCL tools more generally, the *Connection Log* invited and supported students in "engaging in productive processes" and "engaging in co-construction" but also allowed students to monitor and regulate their groupmates' and their own work, and allowed the teacher to do the same (Jeong & Hmelo-Silver, 2016, p. 250).

In studies in 7th grade science, lower-achieving students who used the *Connection Log* gained significantly more from pre to post on an argument evaluation test than matched controls (*ES* = 0.93; Belland et al., 2015), lower-achieving experimental students performed significantly better in argument evaluation (*ES* = 0.61) than matched controls (Belland, Glazewski, & Richardson, 2011) and average-achieving experimental students performed significantly better in argument evaluation (*ES* = 0.62) than matched controls (Belland, 2010), and groups who used the *Connection Log* engaged with data, synthesized evidence from multiple sources, and adhered to stakeholder positions, while control groups largely tried to find out online if the river was polluted and labeled water quality elements dichotomously (Belland, Gu, Kim, & Turner, 2016). Across several studies, students used scaffolds in diverse ways in response to their needs (Belland, 2010; Belland et al., 2011, 2015). In a mixed method study among 6th grade students at risk of academic failure, students had greater science interest when they saw activities as authentic, and their epistemic aims were more sophisticated when their science interest was higher (Gu, Belland, Weiss, Kim, & Piland, 2015).

Data collection

Screencasting data

Everything that students (a) did on their computers and (b) verbalized was recorded using Screencastify. This was used to indicate how students searched for information, how quickly they read and interacted with sources (e.g., scrolled rapidly), and how they interacted with the *Connection Log*. All student discourse was transcribed.

Interviews

At the end of the unit, students engaged in 30-minute interviews covering how they used information to find solutions to the problem. Sample questions included "How did you judge the accuracy of information?" and "How did you decide on search terms while performing searches?" The interviews were transcribed verbatim.

Log files

For each student, time spent on individual pages and text written in response to prompts was collected. We summed time spent on each stage, resulting in the following variables: time - define the problem, time - determine needed information, time - find and organize information, time - develop claims, and time - link evidence to claims. We also counted the number of words written in response to each prompt, and summed that across pages within each stage, resulting in the following variables - word count - define the problem, word count - determine needed information, word count - find and organize needed information, word count - develop claims, and word count - link evidence to claims. We also added the amount of time spent on individual and collaborative learning tasks, resulting in the following variables: time - groupwork, and time - individual work.

Pre and post information literacy assessment

To measure information literacy, we used Tools for Real-Time Assessment of Information Literacy Skills. The 25 items cover (a) development and implementation of search strategies b) evaluation of information, and c) ethics in information use (Kovalik, Yutzey, & Piazza, 2012). Internal consistency was 0.83 (Cronbach's alpha), which is consistent with other studies ($\alpha = 0.8$ to 0.82; Arnone, Small, & Reynolds, 2010; Salem, 2014).

Essay rating

Student argument quality was assessed by rating their essays, in which students needed to make a claim for addressing soil quality, provide evidence to support their claims, and link evidence to claims. Two raters scored students' essays with a rubric designed to rate the structural soundness of the argument and the work that it does in presenting a solution to the stated problem. The rubric has six subcategories to assess argument quality - a) claim, b) relatedness of claim to topic, c) evidence, d) relatedness of evidence to topic, e) articulation of connection of evidence to claim, and f) relatedness of articulation of connection of evidence to claim to the topic. A maximum score was 12 and minimum score was 0. Each rater scored students' essays independently and then came to consensus. The initial inter-rater reliability between two raters with 10 samples was 0.87.

Data analysis

RQ1: Do the time that students spend in each scaffolding stage, time spent working individually, time spent working collaboratively, and information literacy scores predict argument quality?

We used Bayesian linear regression analysis to estimate the relationship between predictor variables (pre and post information literacy scores, and time spent (a) in each stage of the *Connection Log*, (b) working individually, and (c) working collaboratively) and essay rating, which represents argument quality. We used uniform distribution as the non-informative prior distribution and generated the posterior distribution by Metropolis-Hastings MCMC algorithms. In contrast to classic linear regression, Bayesian regression does not generate coefficients of determination (i.e., R-squared) and F values to evaluate model results because the observed data was not included in the posterior distribution and each predictor in Bayesian regression model has its own parameter at the population level (Kaweski & Nickeson, 1997). But, t-tests can still be used to identify the significance of coefficients from each predictor included within a Bayesian regression model.

RQ2: How and why is argument quality explained by information literacy and student engagement with a computer-based scaffold? and RQ3: What choices do high school students make as they construct arguments in PBL with the support of a computer-based scaffold, and why?

Theoretical framework. Social cognitive theory suggests that knowledge is acquired when students choose to observe and connect with others as models of learning behavior through social interactions and experiences. Student choices can be organized into the following categories: intentionality, forethought, self-reactiveness, and self-reflectiveness. Intentionality refers to a students' proactive commitment to carry out a course of action.

Forethought occurs when students motivate themselves to guide their own actions toward a contemplated future goal. Self-reactiveness is when students exercise self-monitoring of their choices and actions by evaluating their values and judge their choices against actual and potential outcomes. Agentic perspective is evidenced through the choices a student makes that result in learning experiences contributing to knowledge building.

Process. We followed the process of a) data reduction, b) data display, c) data transformation, d) data correlation, e) data consolidation, f) data comparison and g) data integration. An initial set of codes were derived from reading transcripts, post unit interviews, student soil quality essays and viewing screen-capture video to account for patterns in the data. An evolving coding scheme was further applied and themes clarified and strengthened. Coding categories included information literacy topics in the Tool for Real-time Assessment of Information Literacy: develop use and revise search strategies, identify potential sources, and evaluate sources. Additional codes from Bandura's social cognitive theory included intentionality, forethought, self-reactiveness, and self-reflectiveness.

Results

RQ1: Do the time that students spend in each scaffolding stage, time spent working individually, time spent working collaboratively, and information literacy scores predict argumentation quality?

Bayesian regression indicates that students' post information literacy, individual time, and time spent on 'Define the problem' and 'Link evidence to claims' significantly predict their argument quality (See Table 1).

Table 1: Bayesian Regression Results

Bayesian normal regression	MCMC iterations = 12,500
Random-walk Metropolis-Hastings sampling	Burn-in = 2,500
	MCMC sample size = 10,000
	Number of obs = 24
Log marginal likelihood = -153.70089	Acceptance rate = .2071

Predictor Variables	Mean	Std. Dev.	MCSE	t-value	Equal-tailed (95% Credible Interval)	
Pre information literacy	-1.736	2.486	0.542	-0.698	-6.987	2.782
Post information literacy	4.743	1.819	0.351	2.607*	1.140	8.090
Word count – Define the Problem	-3.386	2.716	0.562	-1.247	-8.781	1.710
Word count– Determine needed info	2.720	2.701	0.556	1.010	-3.115	7.525
Word count – Find and organize info	-0.748	2.034	0.299	-0.368	-4.642	3.011
Word count – Develop claims	2.179	3.227	0.440	0.675	-4.342	8.775
Word count – Link evidence to claims	2.567	1.706	0.284	1.505	-0.654	5.938
Time - Groupwork	-5.263	3.178	0.464	-1.656	-11.268	1.254
Time - Individual work	-16.178	6.043	1.580	-2.677*	-27.167	-3.486
Time – Define the Problem	11.359	3.147	0.635	3.609*	5.508	17.912
Time – Determine needed info	-2.657	6.361	1.591	-0.418	-15.500	9.900
Time - Find and organize info	5.854	4.272	1.011	1.370	-2.399	13.432
Time – Develop claims	3.302	4.147	0.858	0.796	-3.885	12.085
Time – Link evidence to claims	-13.509	3.804	0.894	-3.551*	-19.843	-5.558
Cons	49.549	10.710	2.748	4.626	30.093	69.446

The coefficient for post-information literacy was 4.74, which means that for every additional point in information literacy, students' argument quality can be expected to increase 4.74 points. Moreover, when students spent one more hour in the section 'Define Problem', their argument quality will increase by 11.36 points. On the other hand, Individual time in the *Connection Log* (β = -16.178) and time spent in the section 'Link evidence to claim' (β = -13.509) negatively affected students' argument quality.

RQ2: How and why is student argument quality explained by information literacy and student engagement with a computer-based scaffold? and RQ3: What choices do high school students make as they construct arguments in PBL with the support of a computer-based scaffold, and why?

Due to space constraints, we only present results for one group.

Quantitative data suggest that when information literacy increases, argumentation quality goes up. When students recognize a need for information beyond their current level of understanding, they need to locate, evaluate and apply information effectively. The *Connection Log* has steps that encourage students to identify what they know, what they don't know and plan for acquiring and using information. Trace data indicated that group 1 members identified topics to investigate, and Internet sources of information were listed. As a group, they decided which of these elements to include in their study, and made assignments for group members to view and report on the sources of information. A final step was to eliminate information that they didn't think they would use. From a time perspective, these steps were steps that students spent the least amount of time accomplishing. However, their word counts were higher than all other steps except linking evidence to claims. Student interviews indicated that this group had almost a nonchalant attitude about information searches

There are three key reasons why an improvement in information literacy had a positive effect on these students' argument quality. First, when conducting web searches, students view not only words but also focus on images displayed on the page. Students not only learn how others articulate thoughts on the subject but also come to understand abstract ideas as they view images and charts. Thus, their confidence in articulating their own ideas increases. Second, web searches explain ideas about the topic that the soil analysis does not. And third, during post unit interviews, students in Group 1 considered information found on the Internet to be authoritative when it was (a)repeated on several sites, or (b) consistent with information the teacher had explained. Student essays created by this group were well organized and often detailed in support of their ideas.

Quantitative data suggest that when students spend more time in "Define the Problem," argument quality goes up. The rubric used to evaluate argument quality assessed claims, evidence and linking evidence to claims. The measurement was conducted on two levels: the logic of argument in the essay and the connection of the argument to the topic of the PBL unit. Of these six rubric sections, the two on which students generally received the highest scores were those related to claims. Within the computer-based scaffold two of the five steps relate directly to claim making namely, "Define the Problem" and "Develop Claims". "Define the Problem" is the first activity in the *Connection Log*. Their motivation is high and a novelty effect exists.

For all four Group 1 members, both the topic of problem statement and the topic of the claim are identical. The similarity of both suggests an alignment in the minds of the students that was reinforcing their view of both the problem and the claim. In addition, the claim topics in group 1 essays were aligned with students' problem statements and claims in the *Connection Log*. Among the rubric scores, the claims elements were highest. In this sense, there is an alignment between their problem statement and their claim in the software and their claim in their essay. In addition to essay topics, problem statement topics and claim topics being centered around low nitrogen as the problem, low nitrogen was also a topic covered by the instructor specifically as a problem in their locale. Further, low nitrogen was a determination resulting from soil tests.

It is important to consider that the "Define the Problem" step includes both individual work and group work. The quantitative measurements suggest a negative effect for individual work and an unknown effect for group work. Yet the combination of individual work and group work in this step resulted in a positive effect with respect to argumentation. Perhaps the alignment of activities in this step, subsequent activities, their soil test results and the teacher's instruction combined to create a positive effect on argument quality.

Quantitative data suggest that when students spend more time on "Linking Evidence to Claims," argumentation quality does not improve. The "link evidence to claim" step, the fifth of the five major steps in the Connection Log, guides students through four individual sub-steps. Two of the sub-steps are completed individually (Select Evidence and Test Evidence) and two are group activities (Put it all Together and Conclusion). The individual activities request that students enter their work by themselves. When the software presents a group activity, students gather around the scribe's computer, discuss information on the screen and decide together what their entries should be. These steps are perhaps most important as they finalize their claim, assign and validate the evidence they are using to support their claim. However, students spent only 5% of their total time using the Connection Log working through this step. Further, individual work by members of Group 1 represented less that 2% of the total word count entered by the group for the entire step. Additionally, two students in Group 1 made no individual entries for this step at all. At the conclusion of this step, students were expected to apply the claims and evidence they had entered into the software to their argumentation essays.

All four essays from group 1 students suggest that low nitrogen was a problem and composting was a possible solution. However, only two of the four students in Group 1 supported their claim by referencing soil analysis tests. Only one student of the four in Group 1 referenced measurements from those tests.

There may be several explanations for the quantitative finding that spending more time on this step would result in a decrease in argument quality. First, as measured by time, word count and the review of their essays, the linking of evidence to claims step was characterized by a lack of individual effort in deference to group work, largely done by the scribe. Further, when the software instructed them to work as a group, only the scribe was tasked with articulating through writing into the software. Screen-capture video and transcripts confirm that Group 1 did discuss briefly what to enter into the software, but the actual recording of their ideas was the scribe's responsibility. Spending more time in this step would likely not contribute to individual argument quality when articulation of ideas in this step was largely the scribe's responsibility.

Second, observations indicated that the teacher (a) made specific reference to nitrogen levels being low in their locale, and (b) informed students that composting was an important solution to soil quality problems. As students spent only 5% of the total time in the software on this step, and little time individually articulating their own ideas in those steps, it is likely that they depended more on their teacher's lecture points in their essays than the work they did in the software. Essays confirm that students focused on low nitrogen and composting in their claims and evidence. Thus, spending more time on this step would not contribute to better argumentation scores.

Third, fatigue is a consideration. Given that two students made no individual entries at all, and the other student's individual entries were minimal, the desire to "finish" the unit may have overcome the need to complete this last step carefully. Argumentation rubric scores support the struggle this group had with articulating evidence and linking evidence to claims. During the post unit interview, students agreed that they stalled using the software, inducing frustration and confusion. Identifying points of evidence that require investigation into collected data, and connecting data to claims, is a higher order skill that these students seem to have resisted especially at the end of the unit.

Quantitative data suggest that when students spend more time engaged with individual activities as opposed to group activities in the software, student argumentation quality does not improve. The *Connection Log* consists of 20 separate steps, 11 of which are completed by individual students and 9 of which are completed as a group. Individual activities include a) define problem individually, b) determine needed information, c) find and organize information, d) generate claim and e) link evidence to claims. Some Group 1 members were methodical entering information step by step. Others spent time on each page but made no entries at all. Students were guided through each step of the PBL process to increase their awareness of what they already knew about the topic, what they needed to know and how they can organize this information as evidence to support a claim. As each student identified what they knew and gathered information needed to understand the problem they identified in the very first step, it was expected that what was entered into the software would make it into the final student essays. However, student essays reflected more of the teacher's lectures than the student's investigative work or interpretation of collected data from internet sources. In some cases, interview data and screen capture software confirm that students completed individual activities using Google searches that led to promising sites on the Internet, as well as visiting the web-based resource page prepared by the research team. Entries were made in the software identifying these sites as sources of evidence as they wrote their arguments. However, successive individual entries referenced the information delivered by their teacher rather than the information they searched for and read on the internet.

There are several explanations why spending more time doing individual work in the computer-based scaffold would not translate into improved argumentation scores. There seems to be a disconnect between their individual work in the scaffold and what they wrote in the essays. High school students often focus on "getting things done". In this way, they may have completed each step in the software without connecting what they did in the software to their essay writing. Spending more time would not be reflected in argumentation scores.

While some students admitted to having a hard time using the software occasionally, they perceived that it invited them to do both individual and group work, citing the pros and cons of both approaches to learning. Students experienced high motivation to write a report that would inform their parents about soil close to their place of residence. However, word counts demonstrated that group activities generated higher interaction with the software than individual activities. Individual activities might have been perceived by students as check-off activities if, in the end, the scribe would articulate answers with greater detail. This is further supported in that for some students in group 1, individual activities had no or little response recorded.

Discussion and implications

One key contribution of this paper was collection of empirical evidence of the prediction of argument quality by information literacy. This implies that the pursuit of enhanced information literacy is not important only in terms of making sure that students have the tools to engage with information effectively, but also so that they can engage with and generate effective arguments.

That time spent in define problem was a significant, positive predictor of argumentation quality, and the one that explained the most variance in argument quality makes sense, as quality of problem definition is often considered to be one of the most important contributors to problem solving ability (Chi, Feltovich, & Glaser, 1981; Jonassen, 2000). And to successfully define a problem, one needs to take time to think about it qualitatively. Problem definition is not always central to argumentation scaffolding (Scheuer, Loll, Pinkwart, & McLaren, 2010), but this result may suggest that it should be.

Time spent working individually was a significant negative predictor of argument quality, which makes sense because students who spent much time working individually and correspondingly less time working collaboratively did not put in the time to critically and constructively engage with their groupmates' ideas (Hmelo-Silver & Barrows, 2008; Rojas-Drummond & Mercer, 2003). At the same time, some students spent relatively little time working individually because they expected the scribe to answer the questions. This would similarly lead to poor collaboration and poor argument quality. Ultimately, what is needed is the right balance between individual and collaborative work time. Further research is needed.

The finding that time spent in the link evidence to claims stage was a significant negative predictor of argumentation ability was surprising. But evidence indicates that the longer students spent in this stage, the less they actually did individual work, and simply left it up to the scribe to do the work. This, of course, is not effective collaboration, and thus conditions were not set for work that leverages the collective group strengths.

References

Arnone, M. P., Small, R. V., & Reynolds, R. (2010). Supporting inquiry by identifying gaps in student confidence: Development of a measure of perceived competence. *School Libraries Worldwide*, *16*, 47.

Arts, J. A. R., Gijselaers, W. H., & Segers, M. S. R. (2002). Cognitive effects of an authentic computer-supported, problem-based learning environment. *Instructional Science*, *30*, 465–495.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Barrows, H. S. (1985). *How to design a problem-based curriculum for the preclinical years.* NY: Springer.

Belland, B. R. (2010). Portraits of middle school students constructing evidence-based arguments during problem-based learning: The impact of computer-based scaffolds. *Educational Technology Research and Development*, *58*(3), 285–309. doi:10.1007/s11423-009-9139-4

Belland, B. R., & Drake, J. (2013). Toward a framework on how affordances and motives can drive different uses of computer-based scaffolds: Theory, evidence, and design implications. *Educational Technology Research & Development*, *61*, 903–925. doi:10.1007/s11423-013-9313-6

Belland, B. R., Glazewski, K. D., & Richardson, J. C. (2008). A scaffolding framework to support the construction of evidence-based arguments among middle school students. *Educational Technology Research and Development*, *56*(4), 401–422. doi:10.1007/s11423-007-9074-1

Belland, B. R., Glazewski, K. D., & Richardson, J. C. (2011). Problem-based learning and argumentation: Testing a scaffolding framework to support middle school students' creation of evidence-based arguments. *Instructional Science*, *39*, 667–694. doi:10.1007/s11251-010-9148-z

Belland, B. R., Gu, J., Armbrust, S., & Cook, B. (2015). Scaffolding argumentation about water quality: A mixed method study in a rural middle school. *Educational Technology Research & Development*, *63*(3), 325–353. doi:10.1007/s11423-015-9373-x

Belland, B. R., Gu, J., Kim, N., & Turner, D. J. (2016). An ethnomethodological perspective on how middle school students addressed a water quality problem. *Educational Technology Research & Development*, *64*, 1135–1161. doi:10.1007/s11423-016-9451-8

Belland, B. R., Gu, J., Weiss, D. M., & Kim, N. J. (2016). An examination of credit recovery students' use of computer-based scaffolding in a problem-based, scientific inquiry unit. Presented at the Annual Meeting of the American Educational Research Association, Washington, DC, USA.

Belland, B. R., Walker, A. E., & Kim, N. J. (Under review). A Bayesian network meta-analysis to synthesize the influence of contexts of scaffolding use on cognitive outcomes in STEM education.

Belland, B. R., Walker, A. E., Kim, N., & Lefler, M. (In Press). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research*.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science: A Multidisciplinary Journal*, *5*(2), 121–152.

Forte, A. (2015). The new information literate: Open collaboration and information production in schools. *International Journal of Computer-Supported Collaborative Learning*, *10*(1), 35–51.

Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. *Review of Educational Research*, *75*, 27–61.

Glassner, A., Weinstock, M., & Neuman, Y. (2005). Pupils' evaluation and generation of evidence and explanation in argumentation. *British Journal of Educational Psychology*, *75*, 105–118.

Gu, J., Belland, B. R., Weiss, D. M., Kim, N. J., & Piland, J. (2015). Middle school students' science interest and epistemic beliefs in a technology-enhanced, problem-based, scientific inquiry unit. Presented at the Annual Meeting of the American Educational Research Association, Chicago, IL, USA.

Hakkarainen, P. (2009). Designing and implementing a PBL course on educational digital video production: lessons learned from a design-based research. *Educational Technology Research & Development*, *57*(2), 211–228. doi:10.1007/s11423-007-9039-4

Hannafin, M., Land, S., & Oliver, K. (1999). Open-ended learning environments: Foundations, methods, and models. In C. M. Reigeluth (Ed.), *Instructional design theories and models: Volume II: A new paradigm of instructional theory* (pp. 115–140). Mahwah, NJ, USA: Lawrence Erlbaum Associates.
Hmelo-Silver, C. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, *16*, 235–266. doi:10.1023/B:EDPR.0000034022.16470.f3
Hmelo-Silver, C., & Barrows, H. (2008). Facilitating collaborative knowledge building. *Cognition and Instruction*, *26*, 48–94. doi:10.1080/07370000701798495
Jeong, H., & Hmelo-Silver, C. E. (2016). Seven affordances of computer-supported collaborative learning: How to support collaborative learning? How can technologies help? *Educational Psychologist*, *51*, 247–265.
Jonassen, D. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, *48*(4), 63–85. doi:10.1007/BF02300500
Kaweski, D., & Nickeson, M. (1997). *C-SHRP Bayesian modelling: A user's guide*. Ottawa, Canada: Canadian Strategic Highway Research Program. Retrieved from https://trid.trb.org/view.aspx?id=474116
Kovalik, C. L., Yutzey, S. D., & Piazza, L. M. (2012). Assessing change in high school student information literacy using the tool for real-time assessment of information literacy skills. *Contemporary Issues in Education Research (CIER)*, *5*(3), 153–166. doi:10.19030/cier.v5i3.7092
Krajcik, J., Codere, S., Dahsah, C., Bayer, R., & Mun, K. (2014). Planning instruction to meet the intent of the Next Generation Science Standards. *Journal of Science Teacher Education*, *25*(2), 157–175.
Kuhn, D. (2010). Teaching and learning science as argument. *Science Education*, *94*, 810–824.
Kuiper, E., Volman, M., & Terwel, J. (2005). The web as an information resource in K–12 education: Strategies for supporting students in searching and processing information. *Review of Educational Research*, *75*, 285–328. doi:10.3102/00346543075003285
McLaughlin, M., & Overturf, B. J. (2012). The Common Core: Insights into the K–5 standards. *The Reading Teacher*, *66*, 153–164. doi:10.1002/TRTR.01115
Oliver, K., & Hannafin, M. J. (2000). Student management of web-based hypermedia resources during open-ended problem solving. *Journal of Educational Research*, *94*, 75–92.
Onwuegbuzie, A. J., Slate, J. R., Leech, N. L., & Collins, K. M. (2009). Mixed data analysis: Advanced integration techniques. *International Journal of Multiple Research Approaches*, *3*(1), 13–33.
Perelman, C., & Olbrechts-Tyteca, L. (1958). *La nouvelle rhétorique: Traité de l'argumentation [The new rhetoric: Treatise on argumentation]*. Paris, France: Presses Universitaires de France.
Reiser, B. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences*, *13*, 273–304. doi:10.1207/s15327809jls1303_2
Rojas-Drummond, S., & Mercer, N. (2003). Scaffolding the development of effective collaboration and learning. *International Journal of Educational Research*, *39*(1–2), 99–111.
Salem, J. A. (2014). *The development and validation of all four TRAILS (tool for real-time assessment of information literacy skills) tests for K-12 students* (PhD Dissertation). Kent State University, Kent, OH.
Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. (2010). Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning*, *5*, 43–102.
Schmidt, H. G., Rotgans, J. I., & Yew, E. H. (2011). The process of problem-based learning: what works and why. *Medical Education*, *45*(8), 792–806. doi:10.1111/j.1365-2923.2011.04035.x
Sugrue, B. (1995). A theory-based framework for assessing domain-specific problem-solving ability. *Educational Measurement: Issues and Practice*, *14*(3), 29–35.
van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, *22*, 271–296. doi:10.1007/s10648-010-9127-6
Van de Vord, R. (2010). Distance students and online research: Promoting information literacy through media literacy. *The Internet and Higher Education*, *13*(3), 170–175. doi:10.1016/j.iheduc.2010.03.001
Walker, A., & Leary, H. (2009). A problem based learning meta analysis: Differences across problem types, implementation types, disciplines, and assessment levels. *Interdisciplinary Journal of Problem-Based Learning*, *3*(1), 12–43. doi:10.7771/1541-5015.1061
Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*, 89–100. doi:10.1111/j.1469-7610.1976.tb00381.x

Acknowledgements

This research was supported by Early CAREER Grant 0953046 awarded to the first author by the National Science Foundation. Any opinions, findings, and or conclusions are those of the authors and do not necessarily represent official positions of NSF.