# Teacher Practices that Support Students' Construction of Scientific Explanations in Middle School Classrooms

David J. Lizotte, Katherine L. McNeill, & Joseph Krajcik
University of Michigan, Center for Highly Interactive Computing in Education
610 E. University Ave., Ann Arbor, MI, 48109-1259
Tel: 734-647-4225, Fax: 734-615-5245
Emails: dlizotte@umich.edu, kmcneill@umich.edu, krajcik@umich.edu

**Abstract:** Scientific explanation is an important inquiry practice emphasized in national standards, yet research has neglected the role of the teacher in supporting students' explanation construction. The present study considers whether two teacher practices, modeling explanation and making a framework for explanation explicit to students, predict students' improvement in explanation during a middle-school chemistry unit. Six teachers enacted the unit with a total of 21 classes in urban public schools. We coded the teachers' practices during a focal lesson in which they introduced their students to scientific explanation. Regression analyses showed that teachers' scores for the two practices made unique contributions to the prediction of students' posttest scores for one component of the explanation framework, reasoning (i.e. justification for why evidence supports a claim). Furthermore, charting students' explanations across the unit showed that both teacher practices had an immediate impact on students' reasoning in the focal lesson, relative to the pretest, an effect that carried through the unit. We identify avenues for future research that build on our initial effort at mapping teacher practices to student learning outcomes for explanation.

Recent science education standards documents (American Association for the Advancement of Science, 1993; National Research Council, 1996) advocate that teachers create learning environments that support student learning of scientific inquiry and the nature of science. Engaging in reading, writing, and talking science is often difficult for middle-school students because science discourse and practice are new to them (Krajcik, Blumenfeld, Marx, Bass, & Fredericks, 1998). Without support for learning these new ways of knowing, doing, and talking science, students may not relate to science and even actively resist learning it (Lee & Fradd, 1998). Teachers need to help mediate these ways of knowing science for students by helping them make sense of these practices (Driver, Asoko, Leach, Mortmer, & Scott, 1994).

## Teacher Practice

Teachers play a key role in structuring students' scientific discourse and guiding their scientific inquiry (Reiser et al., 2001). Teacher-to-student and student-to-student dialogs provide an important opportunity for helping students learn to think critically (Hogan & Pressley, 1997). Teacher provided prompts or contextualized scaffolding can encourage a deep learning approach in students where they are able to articulate their reasoning about how and why something occurs (Chinn & Brown, 2000). Yet there are few research studies that examine features of teachers' practices, roles, and interactions with students that support learning in inquiry-oriented instruction (Flick, 2000; Keys & Bryan, 2001).

Previous research in classrooms does suggest a couple of important teacher practices that can help support student learning. One of the key characteristics of a teacher establishing an inquiry-based learning environment is that they model the behaviors of a scientist (Crawford, 2000). The teacher needs to demonstrate for students how to conduct various inquiry practices such as asking questions, collecting data, analyzing results and forming explanations. By modeling these practices and clearly pointing them out, students can see what that particular form of inquiry looks like in practice. Observing others perform inquiry practices can help students execute these practices on their own. Instruction should both facilitate students' ability to perform inquiry practices and their understanding of the logic behind the practice (Kuhn, Black, Keselman, & Kaplan, 2000). Consequently, a second key instructional characteristic is making scientific thinking strategies explicit to students to facilitate their understanding and use of the strategies (Herrenkohl, Palinscar, DeWater, & Kawasaki, 1999). For example, Chen and Klahr (1999) found that providing students with the rationale behind controlling variables in science experiments, as well as examples of unconfounded comparisons, resulted in greater learning relative to students who did not receive the explicit instruction.

## Scientific Explanation

We are interested in the role of the teacher in supporting inquiry practices. Specifically, we are interested in one inquiry practice – the construction, analysis, and communication of scientific explanations. A deep understanding of science content is characterized by the ability to explain phenomena (Barron et al., 1998). Having students engage in explanation can change or refine their image of science (Bell & Linn, 2000) and enhance their understanding of science content (Driver, Newton, & Osborne, 2000). While explanations are often cited as important for classroom science, they are frequently left out of classroom practice (Kuhn, 1993) and few research studies have examined the effectiveness of instructional practices in helping students construct explanations (Reznitskaya & Anderson, 2002). Previous research on students' construction of explanations in science has focused on written scaffolds provided in the student materials or software program (e.g. Bell & Linn, 2000; Lee, 2003; Sandoval, 2003; Zembal-Saul, Munford, Crawford, Friedrichsen, & Land, 2002). Other studies have focused on students' discussions in order to characterize their explanations (Jiménez-Aleixandre, Rodríguez, & Duschl, 2000; Meyer & Woodruff, 1997). Research appears to have neglected the role of the teacher in supporting students' explanation construction. We are interested in how different teacher practices during the enactment of the same instructional unit influence students' understanding of scientific explanations. More specifically, we are interested in whether modeling the construction and critique of scientific explanations, and providing students with an explanation framework and the rationale behind that framework will result in greater student understanding of scientific explanation.

## Instructional Context

Our research team developed a middle-school chemistry unit using a learning-goals-driven design model (Reiser, Krajcik, Moje, & Marx, 2003). The unit's learning goals target national content standards (AAAS, 1993; NRC, 1996) for three interrelated chemistry topics: substances and properties, chemical reactions, and conservation of matter. Each topic emphasizes both the macroscopic phenomenon or process (e.g. chemical reaction) and the particulate-level understanding of that phenomenon or process. Students and their teachers explore the chemistry content through the context of the driving question (Krajcik, Czerniak, & Berger, 1999), "How can I make new stuff from old stuff?" Specifically, they investigate in-depth how they can make soap from fat through a chemical reaction. During the instructional sequence, students investigate other phenomena to meet the learning goals, each time cycling back to soap and fat.

The unit's learning goals also target national inquiry standards (AAAS, 1993; NRC, 1996). The inquiry learning goals delineate the scientific practices that students develop to investigate and understand the chemistry content. A key scientific practice developed through the unit is constructing explanations for scientific phenomena. To introduce students to scientific explanations, teachers use a focal lesson. First students write an explanation to address a question about data that they collected in previous lessons, using their prior understanding of explanation in science as a guide. Then the teacher introduces the concept of "scientific explanation." The instructional materials define a framework for explanation that was adapted from Toulmin's (1958) model of argumentation. Our explanation framework includes three components: a claim (a conclusion about a problem), evidence (data that supports the claim), and reasoning (a justification, built from scientific principles, for why the evidence supports the claim). The materials provide guidance to teachers for making the framework explicit to students, including appropriate definitions of the components. Next the teacher leads a discussion about three hypothetical examples of weak and strong explanations for the phenomenon that students addressed at the start of the lesson. The instructional materials provide guidance to teachers on how to model the use of the explanation framework to evaluate the explanations for the quality of the three components. Using what they learned about the explanation framework and the models of explanations provided by the teacher, students critique their explanations from the beginning of the lesson and subsequently revise them.

After the focal lesson, students write eight more explanations over the course of the unit, thus employing the same scientific practice to understand a variety of content. Students record their explanations on their investigation sheets. These investigation sheets reinforce the explanation framework via scaffolds that address the three components of explanation. Whereas previous research on using scaffolds in science to promote students written explanations has focused on scaffolds to help students with the content they need to use across the different explanation components (e.g. Bell & Linn, 2000; Sandoval, 2003; Lee, 2003; Zembal-Saul et al., 2002), the investigation sheets in our unit include scaffolds integrated with the content that aim to help students with the explanation components themselves. For example, a scaffold for evidence stated: "Three Pieces of Evidence (Provide three pieces of data that support you claim that new substances were or were not formed.)" A separate study by the research team compared student learning gains for explanation based on whether the explanation component scaffolds provided the same high detail of support throughout the unit or faded the detail of support over time (see McNeill, Lizotte, Krajcik, & Marx, 2004, for discussion).

## Method

### Participants

Participants included 6 teachers and 619 seventh grade students from schools in the Midwest. Four of the teachers and 410 of the students in 14 classes were from public middle schools in a large urban area. The majority of these students were African American and from lower to lower-middle income families. The other two teachers and 209 students in 7 classes were from two public middle schools in a second large urban area. In one of these schools, the majority of the students were Hispanic and from low-income families. In the second school, the student population was racially diverse (44% Hispanic, 18% African American, 24% Asian/Pacific Islander, 12% Caucasian and 2% Native American) and the majority of students were from low-income families.

### Scoring Teacher Practices

In order to characterize teacher practices, we analyzed videos of the classroom enactments. The videos consisted of the focal lesson on explanation for each of the six teachers. While the instructional materials discussed scientific explanations throughout, we chose this lesson because the focus was on scientific explanation. Consequently, we chose this lesson because we believed it would capture or predict how each teacher taught scientific explanation throughout the unit. We coded each video to characterize the quality of two teacher practices, making the rationale or framework behind scientific explanations explicit to students and modeling scientific explanations. We developed the coding schemes from our theoretical framework and an iterative analysis of the data (Miles & Huberman, 1994). After finalizing the coding schemes, each lesson was coded by two independent raters. Reliability is presented separately for the two teacher practices. For coding teacher practices of making the framework behind scientific explanations explicit, the inter-rater reliability was 81%. For coding teacher practices of modeling scientific explanations, the inter-rater reliability was 83%. All disagreements were resolved through discussion.

For the first teacher practice, making the explanation framework explicit to students, we gave each teacher a score from 0 to 6 for each of the three components of scientific explanation (claim, evidence, and reasoning). The codes are described in more detail in Figure 1.

| Code | Description of Code |
|---|---|
| 0: Does not identify | The teacher did not mention the component during the focal lesson. |
| 1: Incorrect description | The teacher mentioned the component, but the description of it was inaccurate. |
| 2: No description | The teacher mentioned the component, but did not explicitly describe or define the component. |
| 3: Vague description | The teacher provided a vague definition of the component. |
| 4: Correct but incomplete description (less important part) | Included teachers who described the component correctly, but the description was incomplete. The definitions of claim, evidence, and reasoning each included two parts. One part was more important for an understanding of the component than the other part. Teacher practices received a code of 4 or 5 when they discussed one part and not the other. Code 4 was given when teachers described the less important of the two parts. |
| 5: Correct but incomplete description (more important part) | Consisted of teachers who discussed the more important of the two parts of the definition. |
| 6: Correct and complete description | The teacher provided a complete and accurate definition of the component, which included both parts. |

Figure 1. Codes for teachers' explicitness of explanation framework.

For the second teacher practice, modeling scientific explanations, we coded each teacher's discussion of the three examples of scientific explanations provided in the instructional materials for the focal lesson. We assigned each teacher a total of nine codes: claim, evidence, and reasoning codes for each of the three examples. Each code ranged from 0 to 5. The codes are described in more detail in Figure 2. Codes were averaged across examples to assign each teacher a mean score for each explanation component (claim, evidence, and reasoning).

| Code | Description of Code |
|---|---|
| 0: Incorrect identification | The teacher incorrectly identified the component in the explanation. For instance, a teacher might say that an example does not include a claim when in fact it did include a claim. |
| 1: Does not identify | The teacher did not mention whether the example included the component. |
| 2: Identifies too much | Consisted of teachers who identified more than the component in an explanation. For instance, a teacher might say that the claim in an example was "Fat and soap are different substances. Fat and soap have different colors." The second sentence is in fact part of the evidence so the teacher has identified more than the claim in this example. This score could only apply if the example included a component. |
| 3: Vague identification | Included teachers who made a vague statement that an explanation did or did not include the component, but did not explicitly address why the example did or did not include that component. For instance, a teacher might simply say that an example includes reasoning without discussing where the reasoning is in the example or why it counts as reasoning. |
| 4: Identifies too little | Consisted of teachers who explicitly identified only a portion of a component. For instance, an example explanation may include three pieces of evidence and a teacher only discusses two of those pieces of evidence. A teacher could only receive this code if a component included multiple parts (e.g. three pieces of evidence). |
| 5: Correct and complete identification | The teacher explicitly identified the component and discussed why the component was in a particular example. |

Figure 2. Codes for teachers' modeling explanations.

### Scoring Student Explanations

In order to assess student learning, we collected two types of assessment data: student investigation sheets and pre- and posttest data. For the student investigation sheets, all three components of explanation (claim, evidence, and reasoning) were scored separately. One rater scored all of the questions. A random sample of 20% of the student sheets was scored by a second independent rater. The average inter-rater reliability was above 85% for each component (claim, evidence, and reasoning) for the two explanations analyzed here (see Results). All students completed the same pretest and posttest, which consisted of 30 multiple-choice and six open-ended items. Only students who completed all parts of the test were included in the analyses. Due to high absenteeism in the urban schools and the necessity of students being in class for four days of testing, only 406 students took all parts of the pre- and posttest assessments. For this study, we focused on one of the open-ended items, which asked the students to write a scientific explanation. For the scientific explanation question, we scored the different components of explanation (claim, evidence, and reasoning) separately. Again, one rater scored all items. Twenty percent of the tests were randomly chosen and scored by a second independent rater. The average inter-rater reliability was above 85% for each component of the test question.

## Results

Our analyses address the following questions: 1) Did students' explanations improve from pre- to posttest and, if so, in which of the components (claim, evidence, reasoning)? 2) Did the two teacher practices measured during the focal lesson (making the explanation framework explicit and modeling explanations) predict students' pre-posttest progress with explanation? 3) How were the developmental patterns of students' explanation scores across the unit, assessed through their investigation sheets and tests, related to the two teaching practices?

### Students’ Pre-Posttest Explanation Improvement

Students’ explanation scores on the test item improved from pre- to posttest. Table 1 gives their pre- and posttest scores for each explanation component and the composite, calculated as the sum of their component scores. Each component score was significantly higher on the posttest relative to the pretest. The pre-post effect was comparable for claim and evidence. Students’ claim and evidence scores were higher than their reasoning scores on the posttest; however, students’ reasoning scores demonstrated the most improvement from pre- to posttest as indexed by the greater effect size for reasoning relative to claim and evidence.

Table 1. Students’ Pre- and Posttest Explanations (N=406)

| Score type | Maximum | Pretest *M* (*SD*) | Posttest *M* (*SD*) | *t* (405) [a] | Effect size [b] |
|---|---|---|---|---|---|
| Composite | 3.75 | 0.46 (0.84) | 1.31 (1.27) | 13.04$^{***}$ | 1.01 |
| Component | | | | | |
| Claim | 1.25 | 0.27 (0.49) | 0.57 (0.61) | 8.89$^{***}$ | 0.61 |
| Evidence | 1.25 | 0.18 (0.41) | 0.46 (0.56) | 8.82$^{***}$ | 0.68 |
| Reasoning | 1.25 | 0.02 (0.13) | 0.29 (0.47) | 11.54$^{***}$ | 2.07 |

[a] One-tailed paired *t*-test
[b] Effect size is the difference between pretest *M* and posttest *M* divided by pretest *SD*.
$^{***}$ $p < .001$

### Impact of Teacher Practices on Students’ Progress with Explanation

To examine whether the teacher practices that we measured during the focal lesson impacted students’ pre-posttest progress with explanation, we performed a series of least-squares linear regressions. We regressed students’ posttest explanation scores on teachers’ practice scores separately for claim, evidence, and reasoning. For each explanation component, we performed a sequential regression in which we regressed students’ posttest explanation scores for the component on their corresponding pretest scores in step 1 and then added the teachers’ scores for the appropriate component for making the explanation framework explicit and modeling explanations as predictors in step 2. To determine whether the teacher practice scores made a significant contribution beyond students’ pretest scores to the prediction of their posttest scores, we examined the change in posttest score variance explained by the regression models from step 1 to step 2. Table 2 gives the results of each sequential regression. The teacher practice scores made a significant contribution to the prediction of students’ posttest scores for reasoning, but not for claim and evidence. Furthermore, in the case of reasoning, the two teacher practices made unique significant contributions to the model; the standardized beta coefficients in full model (i.e. following step 2) for teachers’ explicitness of explanation framework ($\beta = 0.14$) and modeling explanations ($\beta = 0.18$) were both significant: $t = 2.97$, $p < .01$ for explicitness of framework and $t = 3.65$, $p < .01$ for modeling explanations. Table 2 provides part correlations as an index of the magnitude of the relationships between posttest scores and the predictors in the full models. Whereas the correlation of posttest scores and pretest scores was significant for claim, evidence, and reasoning, the correlations of posttest scores and each teacher practice were significant for reasoning only.

Table 2. Regressions of Students’ Posttest Scores on Students’ Pretest Scores and Teachers’ Practices (N = 406)

| | | | Part correlations with posttest score [c] | | |
|---|---|---|---|---|---|
| | | | | Teachers’ explicitness | Teachers’ |
| modeling | | | | | |
| Component | $\Delta R^2$ [a] | *F* (2, 402) [b] | Pretest score | of framework | explanations |
| Claim | .003 | 0.75 | .236$^{**}$ | -.053 | -.045 |
| Evidence | .007 | 1.53 | .157$^{**}$ | .084 | -.063 |
| Reasoning | .054 | 11.66$^{**}$ | .113$^{*}$ | .143$^{**}$ | .175$^{**}$ |

[a] Change in posttest score variance explained by the regression models from step 1 (predictor: pretest score) to step 2 (predictors: pretest score, teachers’ explicitness of framework, teachers’ modeling explanations).
[b] Significance test for $\Delta R^2$.
[c] Part correlation of each predictor with posttest score, calculated by comparisons to the full model (i.e. after step 2). Significance relative to $r = 0$ determined by two-tailed *t*-test.
$^{**}$ $p < .01$; $^{*}$ $p < .05$

To determine whether the two teacher practices had an additional interactive effect on students' posttest reasoning scores not captured by the previous regression model, we performed a third step in the sequential regression in which we added the product of the teachers' two practice scores for reasoning as an interaction predictor. Prior to computing the product, all predictors were centered; we performed the three-step regression on the centered scores. The additional variance in students' posttest reasoning scores explained by the interaction predictor ($\Delta R^2$ = .001) was not significant, $F$ (1, 401) = 0.64, *ns*, further suggesting that the two teacher practices made unique contributions to the prediction of students' posttest reasoning.

### Relationship between Teacher Practices and Students' Reasoning Development across the Unit

To characterize how the teachers' practices in the focal lesson influenced students' posttest reasoning, we examined patterns in the development of students' reasoning scores on explanation tasks across the unit. For the sub-sample of students having explanation data for the student investigation sheets ($n$ = 265), we examined their reasoning scores on the explanation task immediately following the focal lesson and one subsequent explanation task involving the same content area, as well as their pre- and posttest explanations for that same content area. Figures 3 and 4 chart students' mean reasoning scores on the tests and investigation sheets across the unit based on teachers' scores for their explicitness of the explanation framework and modeling explanations, respectively. For comparison purposes, we grouped teachers into low, medium, and high levels of practice based on whether their scores fell in the lower, middle, or upper third of the rating score range. Both teacher practices had an immediate impact on students' reasoning in the focal lesson, an effect that carried through the unit. Specifically, the students of teachers in the high and medium practice groups had consistently higher reasoning scores than those of teachers in the low practice group.

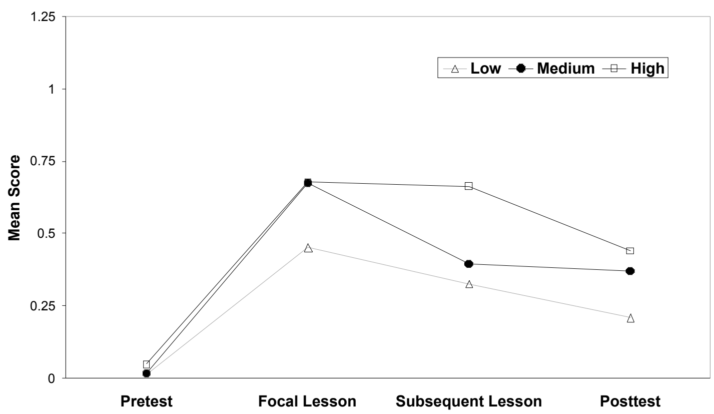


Figure 3. Reasoning scores across the unit by teachers' explicitness of explanation framework.

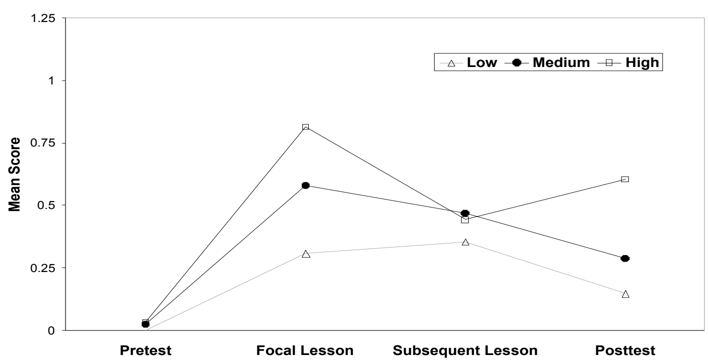


Figure 4. Reasoning scores across the unit by teachers' modeling explanation.

## Discussion

While previous studies found that modeling scientific practices (Crawford, 2000) and making scientific thinking strategies explicit (Herrenkohl et al., 1999) are important teacher practices in inquiry-oriented learning environments, we have built on these studies by using these teacher practices to predict student learning outcomes. Our results suggest that the teacher practices can play an important role in students' understanding and use of scientific inquiry practices, specifically explanation. Modeling the construction and critique of scientific explanations, and providing students with an explanation framework and the rationale behind that framework resulted in greater student understanding of the reasoning component of scientific explanations. Teachers characterized as exhibiting high quality teacher practices had students with higher reasoning on both the lesson artifacts and posttest (see Figures 3 and 4). This suggests that these students had a greater understanding of the reasoning component of explanation after the focal lesson and that their understanding remained over time. We believe that the teacher practices may have been more influential for the reasoning component compared to claim and evidence because students had the most difficultly constructing this part of scientific explanations. Students' pre- and posttest scores for reasoning were lower than either their claim or evidence scores (see Table 1). Our results are similar to other studies that found that students have difficulty with the reasoning or backing component of scientific explanations (Bell & Linn, 2000). Consequently, it is not surprising that having the teacher both explicitly describe and model reasoning resulted in greater student understanding of this aspect of explanation.

We did not find that our characterization of the quality of teacher practices for either claim or evidence had a significant effect on students' construction of claim and evidence. These components seem to be easier for students, and student-level variables, such as content knowledge, may be more powerful predictors of students' success with claims and evidence. Furthermore, our characterizations of either student understanding or teacher practice may account for the findings. We analyzed only one item on the posttest, which may not have accurately captured students' understanding of explanations. We also analyzed only one lesson to capture teacher practices. We believe that the teachers' practices throughout the unit are essential for student understanding of scientific explanation but due to time limitations we were not able to videotape or analyze all of the lessons that address explanation. If we were able to code multiple lessons from each teacher, we would have more accurately characterized the practices of each teacher, providing a more sensitive analysis.

Currently, we are studying ways to further unpack the influence of teacher practices on students' explanation construction that address some of the limitations in this study. Teachers' use of scaffolds within a curriculum as well as their own instructional scaffolds can encourage students to articulate their reasoning about how and why something occurs (Chinn & Brown, 2000). We hope that an investigation of the teacher and curriculum scaffolding in our unit will suggest further variables that relate to students' explanations, thereby adding to the limited research that has examined the effectiveness of instructional practices in helping students construct explanations (Reznitskaya & Anderson, 2002).

## References


American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Barron, B., Schwartz, D., Vye, N., Moore, A., Petrosino, A., Zech, L., Bransford, J., & The Cognition and Technology Group at Vanderbilt. (1998). Doing with understanding: Lessons from research on Problem- and Project-Based Learning. *The Journal of the Learning Sciences. 7 (3&4),* 271-311.

Bell, P., & Linn, M. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. *International Journal of Science Education. 22 (8),* 797-817.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development. 70*(5). 1098-1120.

Chin, C., & Brown, D. E. (2000). Learning in science: A comparison of deep and surface approaches. *Journal of Research in Science Teaching, 37*(2), 109-138.

Crawford, B. A. (2000). Embracing the essence of inquiry: New roles for science teachers. *Journal of Research in Science Teaching, 37*(9), 916-937.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher, 23* (7), 5-12.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education. 84 (3),* 287-312.

Flick, L. B. (2000). Cognitive scaffolding that fosters scientific inquiry in middle level science. *Journal of Science Teacher Education, 11*(2), 109-129.

Herrenkohl, L. R., Palinscar, A. S., DeWater, L. S., & Kawasaki, K. (1999). Developing scientific communities in classrooms: A sociocognitive approach. *The Journal of the Learning Sciences. 8*(3&4), 451-493.

Hogan, K., & Pressley, M. (1997). *Scaffolding student learning: Instructional approaches and issues.* Cambridge, MA: Brookline Books.

Jiménez-Aleixandre, M. P., Rodríguez, A. B., & Duschl, R. A. (2000). “Doing the lesson” or “doing science”: argument in high school genetics. *Science Education, 84*, 757-792.

Keys, C. W., & Bryan, L. A. (2001). Co-constructing inquiry-based science with teachers: Essential research for lasting reform. *Journal of Research in Science Teaching, 38*(6), 631-645.

Krajcik, J., Blumenfeld, P. C., Marx, R. W., Bass, K. M., & Fredericks, J. (1998). Inquiry in project based science classrooms: Initial attempts by middle school students. *Journal of the Learning Sciences*, 77, 317-337.

Krajcik, J. S., Czerniak, C. M., & Berger, C. (1999). *Teaching children science: A project-based approach.* Boston, MA: McGraw-Hill.

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education, 77,* 319-338.

Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction, 18*(4), 495-523.

Lee, H.-S. (2003). Scaffolding elementary students’ authentic inquiry through a written science curriculum (Doctoral dissertation, University of Michigan, 2003). *Dissertation Abstracts International, 64,* 2032A.

Lee, O., & Fradd, S. H. (1998). Science for all, including students from non-english-language backgrounds. *Educational Researcher*, 27 (4), 12-21.

McNeill, K., Lizotte, D., Krajcik, J., & Marx, R. (2004). *Supporting students’ construction of scientific explanations using scaffolded curriculum materials and assessments*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Meyer, K., & Woodruff, E. (1997). Consensually driven explanation in science teaching. *Science Education, 80*, 173-192.

Miles, M., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook (2nd edition).* Thousand Oaks, CA: Sage.

National Research Council. (1996). *National science education standards.* Washington, DC: National Academy Press.

Reiser, B. J., Krajcik, J., Moje, E. B., & Marx, R. W. (2003, March). *Design strategies for developing science instructional materials.* Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Philadelphia, PA.

Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B. K., Steinmuller, F., & Leone, A. J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 263-305). Mahwah, NJ: Erlbaum.

Reznitskaya, A., & Anderson, R. C. (2002). The argument schema and learning to reason. In C. C. Block, & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 319-334). New York: The Guilford Press.

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students’ scientific explanations. *The Journal of the Learning Sciences, 12*(1), 5-51.

Toulmin, S. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.

Zembal-Saul, C., Munford, D., Crawford, B., Friedrichsen, P., & Land, S. (2002). Scaffolding preservice science teachers’ evidence-based arguments during an investigation of natural selection. *Research in Science Education*, *32 (4),* 437-465.

## Acknowledgments

We gratefully acknowledge the collaboration of the six teachers and their students. This project is funded by the National Science Foundation under Grant F005626/ESI-0101780. Opinions expressed herein are those of the authors and not necessarily those of NSF.