Learning Science Online: What Matters for Science Teachers?

Elizabeth Rowe
Jodi Asbell-Clarke
TERC

Abstract

Online education is a rapidly growing phenomenon for science teachers. Using a sample of 40 online science courses for teachers offered during the 2004-2005 academic year, the Learning Science Online (LSO) study explores what characteristics of online science courses are most strongly associated with positive learning outcomes among science teachers, after accounting for teachers' prior science experiences and demographics. This research is unique in that it is the first aggregate study of teachers learning science online in a wide variety of educational programs. Hierarchical linear modeling points to changing roles of instructors and students in online courses, with lower perceived levels of instructor support and a supportive course design strongly associated with positive learning outcomes.

Introduction

Online education is a rapidly growing alternative to face-to-face education for virtually all audiences, including teachers seeking professional development. It addresses issues of geographical remoteness, limited offerings by institutions, and the complex lives of students (Scarafiolti & Cleveland-Innis, 2006). The growth rate of online education in 2003-2004 (18%) was a factor of 10 greater than that projected by the National Center for Education Statistics for the U.S. postsecondary student population on the whole for the same period (NCES, 2005) and has been projected to continue at similar rates (Sloan, 2004). In Fall 2005, a total of 3.2 million post-secondary students in the US studied online, which represents 17% of the post-secondary population and a growth rate of 35% from the year before (Sloan, 2006).

The teacher professional development community discovered early that online programs could offer "anytime, anywhere" education for working teachers. In the early 1990s, programs such as National Teacher Enhancement Network and Project Datastream were using the Internet for science teacher professional development, and today there are many master's degree programs for science teachers delivered online by mainstream universities such as Florida State University and University of Massachusetts, and also online universities such as Capella and Walden Universities. Teachers take online courses not only to improve their content knowledge, but also for professional certification and advancement. Irving (2006) points out that in the US, online courses are becoming particularly important for teachers as they strive to meet the NCLB highly qualified teacher requirements.

Yet even with this rampant growth, the learning that takes place online has been left relatively unstudied. There have been many evaluations of individual courses, but few studies that examine online teaching and learning in a cross-institutional perspective and in particular, there are very few studies of teacher education and even fewer of science learning online. It is clearly important to understand the nature of learning in this increasingly widespread phenomenon, in order to inform the development of new courses and programs. In addition, online courses may provide unique opportunities for teaching science as well as for research on learning and teaching because of their text-based, asynchronous, and archivable learning environments.

Learning Science Online (LSO), a research project funded by the National Science Foundation (NSF), has taken advantage of these opportunities in order to understand the nature and variety of online science courses for teachers. One central research question of the LSO study is:

- What characteristics of online science courses are most strongly associated with positive learning outcomes among science teachers, after accounting for teachers' prior science experiences and demographics?

Theoretical Framework and Related Research

Over the past few decades, the science education community has been advocating for inquiry-based and collaborative instructional methods for students and teachers (AAAS, 1993; Loucks-Horsley, Stiles, & Hewson, 1996; NRC, 1996). Research has shown that professional development programs that actively engage teachers; model appropriate inquiry; and interact with teachers as learners, rather than as information gatherers, are more effective in changing teachers' knowledge and practice (Loucks-Horsley, Hewson, Love, & Stiles, 1998). It is widely believed that learning is fundamentally social in nature (NRC, 1999) and that learning very frequently takes place in cooperative or collaborative settings, and/or in communities or communities of practice (Johnson, Johnson, & Holubec, 1998; Millis & Cottell, 1998; Scardamalia & Berieter, 1996; Wenger, 1997).

The NRC (2005) report on *How Students Learn* suggests that community-centered learning environments, in which learners are encouraged to question and use discussion to solve problems in groups, support scientific inquiry-based learning. Online environments have been advocated for fostering this type of social construction of knowledge (Duffy & Jonassen, 1992; Polin, 2004) within communities of inquiry (Garrison & Anderson, 2003; Gunawardena, Lowe, & Anderson, 1997; Henri, 1991). The textual nature of online discussion forces participants to put their thoughts into writing in a way that others can understand (Koshmann, 1996; Valacich, Dennis, & Connolly, 1994) and promotes the social construction of knowledge through reflection, integration, and reaction to other people's ideas (Gay, Sturgill, Martin, & Huttenlocher, 1999; Pena-Perez, 2000). Students may see different perspectives that may foster new meaning construction (Heller & Kearsley, 1996; Ruberg, Moore, & Taylor, 1996). Online discussion may also foster student ownership of learning and collaborative problem-solving skills, encouraging students to initiate exchanges; potentially shifting the instructor's role from main mediator of student contributions to a guide or prompter (Becker, 1992; Bender, 2003; Hawisher & Selfe, 2000; Hiltz & Goldman, 2005; Luke, 1999; Roberts, 2004; Sefton-Green, 1998).

An oft-used model for analysis of social construction of knowledge in online discussions is the Communities of Inquiry model (Garrison & Anderson, 2003). This model describes three components of a social learning experience: social presence, teaching presence, and cognitive presence; and provides a framework for studying the interaction among them.

Social Presence

The first component, social presence, is "the ability to project one's self and establish personal and purposeful relationships" (Garrison & Anderson, 2003, p. 49). The root of social presence comes from the idea of immediacy (Mehrabian, 1969) a construct that encompasses communication behaviors that enhance closeness to and nonverbal interaction with another.

Short et al. (1976) introduced social presence as an analogous measure to assess the immediacy of others perceived in an online setting through non-visual cues such as emoticons, language, and flow of conversation. In learning communities, social presence is affected by participation from both the instructors and students perceived learning (Richardson & Swan, 2003). Weaver and Albion (2005) used self-report measures of social presence and motivation and found that perceived social presence impacted students' motivation to contribute to discussions. Richardson and Swan (2003) also used a survey to assess students' perception of social presence as well as perceived learning and satisfaction with the instructor. They found that students reporting higher perceived social presence scores also perceived they learned more from the courses than students with low perceived social presence scores, and also that they were more satisfied with their instructor. Students who were most satisfied with their instructors believed they learned more from their courses than students who were less satisfied with their instructors.

Based upon this research, we hypothesize:

- Students who perceive higher levels of social presence in online courses will have more positive learning outcomes

Teaching Presence

The second component of the communities of inquiry model, teaching presence, is defined as "the design, facilitation, and direction of cognitive and social processes for the purpose of realizing personally meaningful and educationally worthwhile learning outcomes" (Garrison & Anderson, 2003, p. 66). This construct integrates:

- *Instructional design and organization* with behaviors such as setting curriculum, designing methods, establishing time parameters, utilizing the medium effectively, and establishing netiquette.
- *Facilitating discourse* with behaviors such as identifying areas of agreement/disagreement; seeking to reach consensus/understanding; encouraging, acknowledging, or reinforcing students' contributions; setting climate for learning; drawing in participants/prompting discussion; and assessing the efficacy of the process.
- *Direct instruction* with behaviors such as presenting content/questions, focus the discussion on specific issues, summarize the discussion, confirm understanding through assessment and explanatory feedback, diagnose misconceptions, inject knowledge from diverse sources, and responding to technical concerns.

Instructional design of a learning environment can refer to basic course management and materials choices but also can include very sophisticated pedagogical mechanisms. Online environments add new dimensions to even the most mundane elements of course design such as scheduling. The use of asynchronous discussion boards allows students to participate at their own optimal time and also allows built-in wait time for students to digest comments or questions before they respond (Hiltz & Goldman, 2005). Netiquette is of importance as students become familiar with how their tone and comments may be received online, and in some cases the extent to which the instructor regulates the flow of discussion may have significant impact on students' participation and perhaps their learning. As technologies advance, the choice of technologies used in online courses may shape the learning experience. Asynchronous discussion boards have shown to be effective new learning tools that improve discourse over face-to-face discussion in some cases (Harlen & Altobello, 2003) and as whiteboards and other shared workspaces become more widely used, they may have tremendous potential to support distributed learning and social knowledge construction. Specific instructional scaffolds for argumentation and knowledge integration have been used online in knowledge construction and integration environments such as Computer-Supported Intentional Learning Environments (Scardamalia & Bereiter, 1991, 1996; Scardamalia, Bereiter, & Lamon, 1994) and Knowledge Integration Environment (Bell, 1997). These environments include prompts to draw out the structures for developing scientific arguments (e.g., Bell & Linn, 1997) as well as specific interfaces that model and facilitate appropriate thinking skills for generating questions, responses, and feedback (Scardamalia, Bereiter, Mclean, Swallow, & Woodruff, 1989). CoVIS, a collaborative visualization project (Edelson & O'Neill, 1994), also incorporates scientific visualization into the learning environment as a scaffold to promote scientific understanding.

Instructor facilitation is also a key component to the teaching presence in online environments. Studies show evidence that students' perceived learning is correlated with the amount and quality of their interaction with their instructor (Jiang and Ting, 2000; Picciano, 1998; Shea, Picket, & Pelz, 2003; Swan, 2000). This is not surprising, as similar findings were reported for face-to-face discussions with students. Hogan, Nastasi, and Pressley (2000) have shown that a difference between teacher-guided discussions and peer discussions in middle school classrooms, where teacher-guided discussions were more efficient at attaining higher levels of cognitive depth – namely reasoning and higher quality explanations, but peer discussions tended to be more generative and exploratory. Nathan and Knuth (2003) found that, when teachers maintain a central social scaffolding role, the student-led discussion increased considerably but lacked the mathematical precision offered by the teacher.

Based upon this research, we hypothesize:

- Students who perceive higher levels of teaching presence in online courses will have more positive learning outcomes.

Cognitive Presence

The third component of the community of inquiry is cognitive presence, "the exploration, construction, resolution and confirmation of understanding through collaboration and reflection in a community of inquiry" (Garrison & Anderson, 2003, p. 3); a description of the social and individual ways in which learners construct knowledge within a community. This aspect, perhaps the most important to a learning community, is also the most elusive to measure.

Garrison and colleagues have operationalized cognitive presence as a cycle of practical inquiry that includes four steps:

1. *Triggering* – the initial engagement and questioning in scientific inquiry;
2. *Exploration* – where learners make hypotheses and predictions and then gather evidence through more controlled investigations;
3. *Integration* – where learners substantiate or counter their previous claims and revise hypotheses from new evidence; and
4. *Resolution* – when learners confirm or revise their working scientific theories.

These have significant overlap with many inquiry models in science, but are more generic to suit multiple disciplines. Garrison found that online discussion in non-science content areas rarely moved beyond triggering and exploration, and question if (a) online discussion is actually limited from achieving integration and resolution; or (b) integration and resolution is occurring elsewhere in the course outside of the discussion. Others question if this result is byproduct of the measurement tool itself (Fahy, 2005; Flynn & Polin, 2003).

Much research about social learning in science focuses on argumentation which includes evaluating evidence, assessing alternatives, establishing the validity of scientific claims, and addressing counterevidence (Driver, Newton, & Osborne, 2000; Kuhn, 2005; Toulmin, 1958).

Discussion and argumentation not only situate student learning in discourse relevant to practice, but also reveal preconceptions and conceptual change that are vital to the learning process. Argumentation emphasizing the coordination of evidence and theory in science requires students to make their thinking visible to themselves and others. The notion of argumentation is counter to the current state of science classrooms, in which teacher talk is valued and student talk is typically discouraged. Science teacher education is instrumental in making changes necessary for teachers to facilitate meaningful argumentation (Sadler, 2004). As an instructional method, argumentation implies having students articulate their scientific ideas, reflect upon their own scientific ideas and the ideas of others, and pose relevant questions about their peers' scientific ideas as well as analyze and interpret scientific data to justify their own ideas. Based on their work in middle school classrooms, Borko et al. (2003) labeled this set of instructional activities hands-on and minds-on instruction.

One of the few studies of science learning in online teacher professional development is Harlen and Altobello's (2003) comparison study of online science courses for teachers to their face-to-face counterparts. They studied *TryScience*, an introductory course to an online master's program for teachers focusing on scientific inquiry. Teachers performed a series of hands-on activities accompanied by extensive online discussion about the activities as well as their own conceptual development. Through questionnaires, pre- and post results on a scientific thought experiment, and content analysis of online discussion transcripts, Harlen and Altobello found that these online discussions promoted reflection and articulation about their own science learning and the process of inquiry and resulted in a greater change in science understanding in the online participants relative to participants in the face-to-face course. Though Harlen and Altobello's study only included one course, it was ground breaking in examining scientific inquiry in online science courses for teachers.

Based on the research in science education and online learning, we hypothesize that:

- Students who engage more frequently in hands-on and minds-on instructional activities will have more positive science learning outcomes.

Research Methodology

Learning Science Online (LSO) builds upon this previous research to examine the nature of teaching and learning within 40 online science courses and examines which course characteristics, if any, correlate with positive learning outcomes for science teachers. Participants at all levels were informed that the purpose of this study was to learn about a variety of online science courses for teachers rather than the evaluation of their specific course.

Participants

The first phase of LSO includes a cross-sectional analysis of 40 online science courses for teachers offered by six online course providers during the 2004-2005 academic year. To ensure a sufficient level of uniformity, courses were sought that met the following criteria:

- At least one graduate credit is offered for the course;
- Science content is the primary focus of the course (as opposed to instructional methods);
- Course requires some interaction among students and between instructor and students, which occurs primarily online.

Of the 60 courses we found meeting these criteria, 45 participated in the study and 40 of those were unique course (five were courses that were re-offered for a second time within the study), giving an overall participation rate of 75%. Two thirds of the courses were affiliated with a Master's degree program and were offered for a maximum of 3 credits. Seventy percent were semester long courses (i.e., duration of 12 weeks or more) while the remaining 30% had a duration of 4-5 weeks. Seventy percent of the courses in the study were biology or life sciences courses (n=28). Most of the rest were in environmental science and Earth/space science. Fewer than ten of the courses included content in astronomy, chemistry, and physics.

Six programs hosted the online courses in this study. Three programs were administered by educational nonprofit institutions and the other three were administered by universities–either within an academic department or through a continuing education or distance education program. Courses in any of these programs may have contributed to a master's degree, but only two programs offered online master's degrees themselves. All courses studied were offered for graduate credit, ranging from 1-4 credits depending on the duration and nature of the course. The intended audience for all LSO courses was science teachers with three-quarters of the courses (n=30) for designed specifically for high school science teachers.

A total of 296 students completed the study and received a $15 online gift certificate. Instructors reported 795 students were enrolled in these 40 courses at the beginning of the courses and 735 were enrolled at the end of the add/drop period. The overall student response rate was 40%. While this student response rate is not ideal, it exceeds the response rate of other studies of online education (Sloan, 2004). Two thirds of the students completing the pre and post questionnaires were women. Ten percent were minority.

The low student response rate suggests the perspectives of these students may be selective in unknown ways. A third of those students completing the pre- and post-questionnaires were male and two thirds were female. There was no difference in the student participation rates by gender in the study. Students who performed better in their courses were more likely to have participated in this study—45% of those who received A's completed pre- and post-questionnaires versus 39% of those receiving B's, 17% receiving C's, and 7% of those who failed, received an incomplete, or withdrew from the course. This suggests the data from students are more representative of those who performed well in the course than of those who did not.

In summary, a total of 35 instructors and 296 students from 40 unique courses in 6 institutions participated in this study.

Data Sources

The findings reported in this study rely on three data sources—(1) pre- and post-course online questionnaires of instructors and students, (2) instructor reports of student course performance, and (3) expert review of the science content in these courses.

Questionnaires

Instructors' and students' pre- and post-course questionnaires were developed and tested with focus groups and piloted in four online courses during the summer of 2004. All questionnaires were administered online. Instructor and student pre-questionnaires collected demographic information such as their highest degree earned, fields of study, teaching experience, experience with online courses, and their expectations about the course. The academic self-efficacy section of the *Teacher Preparation Quality and Capacity: Science Questionnaire* (Skeele, Walker & Klemballa, 2004) was used to measure teacher's science academic self-efficacy and the *Survey of Attitudes toward Statistics (SATS)* scale was modified to measure their perceived value of science (Schau, Stevens, Dauphinee, & Del Vecchio, 1995). Both measures were selected based on evidence of good internal consistency of their scales.

In order to measure instructional methods employed in online science courses for teachers, we used a set of survey items based on the measure of reform-based instruction in K-12 science classrooms developed by Borko et al. (2003). Although Borko's constructs were developed examine children's learning, many of their principles are compatible with the call for new methods of professional development for teachers. This framework has many parallels to NSES standards (NSES, 1996); it utilized detailed rubrics in classroom observations with moderate to high levels of inter-rater agreement. Constructs particularly relevant for measuring reform-based science instruction in online courses for teachers include:

- *Hands-On*: The extent to which learners are interacting with physical materials or models to learn science.
- *Minds-On*: The extent to which learners participate in activities that engage them in wrestling with scientific issues and developing their own understanding of scientific ideas.
- *Collaborative Grouping*: The extent to which a series of lessons uses learner groups to promote learning.

Questions about the use of common instructional materials for science (e.g., completing problem sets, reading textbooks) as well as those unique to online courses (e.g., online discussion boards, simulations) were also included so as to cover the potential range of instructional materials used in these courses.

In addition to questions about instructional methods and materials, instructor and student post-questionnaires focused on course characteristics such as:

- *Intellectual Difficulty*: Student perceptions of the intellectual difficulty of the course materials.
- *Instructor Choice of Discussion Topics:* Student and instructor reports of who chose the new topics in the online discussions (instructors vs. students).
- *Support:* Student's perceptions of support for their learning from their instructor, other students, and course design.

Support measures are indicators of social and teaching presence as defined by Garrison and Anderson (2003). That is, students in courses with high levels of social presence would be expected to report high levels of instructor and student support. Similarly, students in courses with high levels of teaching presence would be expected to report high levels of instructor and course supports. Appendices A and B summarize the items included in each course-level and student-level construct, respectively, and the internal consistency of each scale developed for this study. The internal consistencies of the self-efficacy and perceived value scales were similar to those reported by the scale developers.

Student Course Performance

After they submitted final course grades for all students, instructors were asked to report on the following measures of course performance of students participating in this study:

- final course grade
- instructor rating of their mastery of science content
- instructor rating of the quality of their participation in online discourse
- instructor rating of the frequency of their participation in online discourse

Grades were reported on a 13 point scale with 13=A+, 12=A, etc. Instructors also reported the total number of points each student earned, which allowed the students who took the course Pass/Fail to be included in these analyses through conversion of their total number of points to their probable grade. For the ratings, instructors were asked to think about all of the students in their course and indicate into which quartile (1=Bottom 0-25 percent, 2=25-50 percent, 3=50-75 percent, and 4=Top 75-100 percent) students participating in the study fell relative to other students in the course.

Instructors were also asked to report aggregate information for their courses such as the total number of students enrolled at the beginning of the course and at the end of the add/drop period—both disaggregated by gender—as well as the number of students earning specific grades (A, B, C, etc.). Finally, instructors reported how grades were determined in their course by outlining each grade component (e.g., assignments, tests, online discussion) and the percentage of the final grade assigned to that component. Online discussions accounted for between 0 and 55%of the final course grade (with one-third assigning less than 20%, one-third assigning 21-30%, and one-third assigning more than 30%of the final grade to online discussions). To account for this variation in grading strategies and the lack of representativeness of the students participating in the study described above, three course characteristics drawn from these instructor reports were included in the analyses:

- percentage of all students in the course participating in the study (response rate)
- percentage of all students in the course who received an A
- percentage of final course grade assigned to online discussions

Expert Review

Throughout the 2004-2005 academic year, LSO staff visited online course websites and printed all course materials (with the exception of online discussion transcripts). Copies of additional course readings required to complete the course but not available online were also obtained. All course materials were organized in binders by the following topics: (a) syllabi or course overview, (b) assignments/projects, and (c) readings.

Building on a rubric developed Simon and Forgette-Giroux (2001) to help teachers assess the depth, coherence, and relevance of their student assessments, LSO developed and piloted a four-point rubric focused on three dimensions of science content quality:

- *Relevance to the learning goals of the course:* The extent to which course materials address directly the learning goal(s).
- *Accuracy of the science content:* The extent to which current concepts, terms, principles, and conventions are used correctly and with clarity throughout the course materials.
- *Cognitive depth of the science content:* The extent to which course materials reflect a rich analysis of relevant and high quality references.

In addition to rating the course materials along these dimensions, reviewers were asked to provide evidence to support their ratings. In Summer 2006, LSO convened eight experts with Ph.D.s in the relevant science content area and expertise in teacher education to review these course materials. Each reviewer independently reviewed 6-15 courses, depending upon the area of their scientific expertise. Two experts reviewed each course. After independent reviews were completed, LSO staff convened a conference call between the reviewers to discuss their reviews. Reviewers were instructed that consensus was not the goal of those calls but that they were free to change their ratings based on new evidence they had not considered. Once they had completed all of their reviews, each expert was asked to re-examine their ratings to be sure they were rating courses consistently (i.e. all courses that received a specific cognitive depth rating had more in common than courses that received different cognitive depth ratings). For each rating, reviewers agreed 50% of the time. Disagreements, however, were never more than one point apart on the scale (with one exception). Correlations between ratings ranged from 0.77 for relevance to 0.87 for accuracy, suggesting that while reviewers did not always agree on the exact position of the course along the scale they agreed on the relative position (i.e. which courses were better or worse than others). Not surprisingly, the ratings were highly correlated. In the analyses presented in this paper, courses were described as being of high cognitive depth (top two ratings) or not.

Data Analyses

Hierarchical linear modeling (HLM) was used to examine the relationship between course characteristics and student outcomes while accounting for course-level and student level characteristics. While similar to standard regression analyses, HLM is more statistically precise because it accounts for the hierarchical nature of the design by separating the variation in student outcomes into within-course and between-course components (Raudenbusch & Bryk, 2002). Fully unconditional HLM models revealed that the amount of variability at the course level in these outcomes ranged from 12% (grades and quality of participation) to 20% (mastery of science content). All HLM models accounted for the following course/instructor and student characteristics:

Course-level Characteristics

- Whether or not the course is administered through a university
- Response rate (percentage of all students in the course completing the study)
- Instructor grading strategy (percentage of A's, percentage of course grade assigned to discussions)
- Instructor sex and age

Student-level Characteristics

- Whether or not the student had earned a bachelor's degree in a science field
- Students' sex, pre-course science academic self-efficacy, and pre-course perceived value of science
- Whether or not the student had 10 or more years of science teaching experience
- Other relevant measures of course performance (e.g., analyses of student's final course grades accounted for the instructor's ratings of their mastery and quality of participation, both potential contributors to their final grade)

Tables 1 and 2 present descriptive statistics and bivariate correlations for course-level and student-level characteristics.

Table 1
Descriptive statistics and bivariate correlations among all course-level variables

Course-level variables	Course-level variables (n=37)				
	Mean	SD	1	2	3
1. Pen and paper instruction	2.49	0.94	--	--	--
2. Hands-on instruction	1.64	0.72	-0.29	--	--
3. Minds-on instruction	3.64	0.57	0.12	-0.04	--
4. Collaborative instruction	2.49	1.04	0.19	-0.20	0.37*
5. Aggregated perceived instructor support	4.11	0.36	0.08	0.17	0.49**
6. Aggregated perceived student support	3.89	0.41	0.04	0.06	0.59**
7. Aggregated perceived course support	3.87	0.29	-0.17	0.16	0.21
8. High cognitive depth (1=Yes)	0.46	0.51	0.44**	-0.29	-0.03
9. Choice of discussion topics	3.54	1.26	0.20	-0.05	0.35*
10. Percentage of course grade towards discussions	25.90	13.82	-0.08	0.38*	-0.03
11. University program (1=Yes)	0.68	0.47	0.49**	-0.22	-0.05
12. Percentage of A's	60.74	24.71	0.26	-0.03	0.34*
13. Response rate	65.53	23.15	0.27	0.05	-0.31
14. Instructor sex (1=Female)	0.41	0.50	-0.03	0.16	-0.09
15. Instructor age	55.32	11.29	-0.05	-0.16	-0.06

NOTE: University program, instructor sex, and high cognitive depth are binary variables. All other variables are standardized. See Appendix A for description of pen and paper, hands-on, minds-on, and collaborative instruction constructs. Aggregated instructor, student, and course support measures are course averages of student perceptions described in Appendix B.
*p<0.05, **p<0.01

Table 1 (cont.)
Descriptive statistics and bivariate correlations among all course characteristics

Course characteristic	4	5	6	7	8	9
4. Collaborative Instruction	--	--	--	--	--	--
5. Aggregated perceived instructor support	0.18	--	--	--	--	--
6. Aggregated perceived student support	0.33*	0.45**	--	--	--	--
7. Aggregated perceived course support	0.07	0.44**	0.23	--	--	--
8. High cognitive depth (1=Yes)	0.18	0.36*	0.04	0.05	--	--
9. Choice of discussion topics	0.07	-0.23	-0.38*	-0.23	0.12	--
10. Percentage of course grade towards discussions	-0.22	0.18	0.33*	-0.07	0.08	-0.05
11. University program (1=Yes)	0.51**	-0.08	0.07	-0.34*	0.18	0.02
12. Percentage of A's	0.34*	0.16	0.45**	0.09	0.30	-0.09
13. Response rate	-0.17	0.09	0.03	0.04	0.00	-0.01
14. Instructor sex (1=Female)	-0.35*	0.13	-0.05	0.45**	-0.21	-0.23
15. Instructor age	0.02	0.06	-0.06	-0.26	0.13	0.47**

NOTE: University program, instructor sex, and high cognitive depth are binary variables. All other variables are standardized. See Appendix A for description of pen and paper, hands-on, minds-on, and collaborative instruction constructs. Aggregated instructor, student, and course support measures are course averages of student perceptions described in Appendix B.
*p<0.05, **p<0.01

Table 1 (cont.)

Descriptive statistics and bivariate correlations among all course characteristics

Course characteristic	10	11	12	13	14
10. Percentage of course grade towards discussions	--	--	--	--	--
11. University program (1=Yes)	-0.16	--	--	--	--
12. Percentage of A's	0.01	0.39*	--	--	--
13. Response rate	0.02	0.17	-0.19	--	--
14. Instructor sex (1=Female)	0.02	-0.49**	-0.26	0.30	--
15. Instructor age	0.01	-0.06	-0.06	0.03	-0.21

NOTE: University program, instructor sex, and high cognitive depth are binary variables. All other variables are standardized. See Appendix A for description of pen and paper, hands-on, minds-on, and collaborative instruction constructs. Aggregated instructor, student, and course support measures are course averages of student perceptions described in Appendix B.

*p<0.05, **p<0.01

Table 2

Descriptive statistics and bivariate correlations among all student-level characteristics

Student-level variable	Student-level variables (N=263)				
	Mean	SD	1	2	3
1. Pen and paper instruction	2.53	1.20	--	--	--
2. Hands-on instruction	1.63	0.84	-0.11	--	--
3. Minds-on instruction	3.67	0.63	0.10	0.12	--
4. Collaborative instruction	2.48	1.15	0.14*	-0.16**	0.24**
5. Perceived instructor support	4.11	0.67	0.13*	0.11	0.29**
6. Perceived student support	3.95	0.64	0.09	0.07	0.35**
7. Perceived course supports	3.92	0.57	0.01	0.20**	0.18**
8. Perceived intellectual difficulty	2.00	0.31	0.17**	-0.13*	0.09
9. Choice of discussion topics	3.76	1.28	0.13*	-0.13*	-0.13*
10. Science academic self-efficacy	4.21	0.53	-0.05	-0.07	0.00
11. Perceived value of science	4.39	0.43	-0.10	0.03	-0.03
12. Bachelor's degree in science (1=Yes)	0.67	0.47	0.05	-0.18**	-0.07
13. Student sex (1=Female)	0.65	0.48	-0.05	0.01	-0.10
14. 10+ years of teaching exp (1=Yes)	0.22	0.42	0.05	0.07	0.03
15. Final course grade	11.21	1.52	0.05	-0.03	0.13*
16. Instructor rating of mastery of content	3.26	0.94	-0.17**	-0.04	0.07
17. Instructor rating of quality of participation	3.24	0.94	-0.16*	-0.03	0.10
18. Instructor rating of freq of participation	3.27	0.90	-0.15*	0.05	0.12*

NOTE: All non-binary variables are standardized. See Appendix B for description of course perceptions.

*p<0.05, **p<0.01

Table 2 (cont.)
Descriptive statistics and bivariate correlations among all student-level characteristics

Student-level variable	4	5	6	7	8	9
4. Collaborative instruction	--	--	--	--	--	--
5. Perceived instructor support	0.10	--	--	--	--	--
6. Perceived student support	0.22**	0.50**	--	--	--	--
7. Perceived course supports	-0.04	0.53**	0.34**	--	--	--
8. Perceived intellectual difficulty	0.13*	0.10	0.08	-0.18**	--	--
9. Choice of discussion topics	-0.03	-0.14*	-0.07	-0.05	0.03	--
10. Science academic self-efficacy	0.12	0.03	0.14*	0.11	-0.19**	-0.03
11. Perceived value of science	0.11	0.10	0.16*	0.15*	-0.15*	-0.08
12. Bachelor's degree in science (1=Yes)	0.07	-0.17**	-0.11	-0.16*	-0.14*	-0.07
13. Student sex (1=Female)	-0.10	0.04	-0.09	-0.10	0.07	-0.08
14. 10+ years of teaching exp (1=Yes)	-0.04	0.08	-0.01	-0.02	-0.02	-0.18**
15. Final course grade	0.08	0.18**	0.09	0.18**	-0.16**	-0.11
16. Instructor rating of mastery of content	-0.06	-0.02	0.03	0.15*	-0.24**	0.00
17. Instructor rating of quality of participation	-0.05	0.10	0.11	0.22**	-0.22**	-0.13*
18. Instructor rating of freq of participation	-0.03	0.14*	0.18**	0.19**	-0.23**	-0.16*

NOTE: All non-binary variables are standardized. See Appendix B for description of course perceptions.
*p<0.05, **p<0.01

Table 2 (cont.)
Descriptive statistics and bivariate correlations among all student-level characteristics

Student-level variable	10	11	12	13	14	15	16	17
10. Science academic self-efficacy	--	--	--	--	--	--	--	--
11. Perceived value of science	0.57**	--	--	--	--	--	--	--
12. Bachelor's degree in science (1=Yes)	0.28**	0.11	--	--	--	--	--	--
13. Student sex (1=Female)	-0.18**	-0.08	-0.03	--	--	--	--	--
14. 10+ years of teaching (1=Yes)	0.05	0.10	0.14*	0.06	--	--	--	--
15. Final course grade	0.10	0.05	0.02	-0.13*	0.20**	--	--	--
16. Mastery of science content	0.11	-0.03	0.07	-0.09	0.12	0.61**	--	--
17. Quality of participation	0.11	0.01	0.07	-0.03	0.16**	0.54**	0.74**	--
18. Frequency of participation	0.06	-0.02	0.04	0.07	0.14*	0.55**	0.67**	0.73**

NOTE: All non-binary variables are standardized. See Appendix B for description of course perceptions.
*p<0.05, **p<0.01

In HLM analyses, all non-binary variables, including outcomes, were standardized to have a mean of zero and a standard deviation of one. Thus, all coefficients are reported in terms of effect sizes. All non-binary course-level variables were centered around the grand mean across courses and all non-binary student-level variables were centered around the group (course) mean. All binary variables were not centered. In all models, the intercept coefficient represents the estimated outcome for male students (gender=0) with less than 10 years science teaching experience (10+ years=0) and no bachelor’s degree in a science field (bachelor’s degree=0), with an average level of science academic self-efficacy and perceived value of science among their classmates (self-efficacy and perceived value=0). These students are in courses taught by male instructors of average age (instructor sex and age=0) in non-university programs, in courses with average response rates, average percentages of As given, and average percentage of the final grade allotted to online discussions. Coefficients of course-level predictors of the intercept represent the unit change in an outcome associated with a one-unit change in the course characteristic. Coefficients for the effects of student-level predictors represent the unit change in the outcome associated with a one-unit change in the course characteristic among students in their course.

Results

Final HLM models are presented for each outcome. Relative to the fully unconditional model, the model fit significantly improved for final course grades (chi-sq=326.03, 21(202) d.f., $p<.01$), mastery of science content (chi-sq=348.04, 21(203) d.f., $p<.01$), and quality of participation in online discussions (chi-sq=363.3, 21(207) d.f., $p<.01$). These models explained between 24 and 55% of the course-level variability and approximately half (47-54%) of the student-level variability in these outcomes.

Final Course Grades

Results of the final HLM model of student's final course grades are presented in Table 3. Across courses, the following course characteristics were significantly related to higher final course grades:

- More frequent usage of pen and paper instructional activities
- Less frequent usage of hands-on instructional activities
- Less frequent usage of minds-on instructional activities

Final course grades did not vary significantly across courses with different levels of collaborative instructional activities; perceived support from course materials, instructor, and other students; cognitive depth of course materials; or choice of discussion topics. There are two potential interpretations of this set of findings. One interpretation is that more common instructional activities (pen & paper) may be more comfortable/familiar for students than less common activities (hands-on and minds-on), resulting in better course performance. An alternative explanation are differences in grading strategies among instructors utilizing each of these instructional strategies—perhaps instructors who rely more on hands-on and/or minds-on instructional activities are less likely to give higher grades than instructors who rely more on pen & paper activities. This variation in grading strategies may be an idiosyncratic characteristic of the instructors or may be related to the instructional methods they chose (i.e., performance in minds-on activities may be more difficult to evaluate).

Table 3
Course-level and student-level variables as predictors of student's final course grades

Predictor	Coef.	SE	t	d.f.	P-value
Intercept	0.56	0.24	2.36	21	0.03
Course-level predictors					
Pen and paper instruction	0.17	0.09	2.05	21	0.05
Hands-on instruction	-0.16	0.07	-2.42	21	0.03
Minds-on instruction	-0.21	0.10	-2.12	21	0.05
Collaborative instruction	0.04	0.07	0.53	21	0.60
Aggregated perceived instructor support	0.08	0.07	1.15	21	0.26
Aggregated perceived student support	0.00	0.12	0.03	21	0.98
Aggregated perceived course support	0.10	0.07	1.35	21	0.19
High cognitive depth (1=Yes)	-0.29	0.22	-1.33	21	0.20
Instructor choice of discussion topics	-0.03	0.06	-0.52	21	0.61
Course-level controls					
Percentage of course grade towards discussions	-0.01	0.00	-1.37	21	0.19
University program (1=Yes)	-0.45	0.15	-2.96	21	0.01
Percentage of A's	0.01	0.00	3.41	21	0.00
Response rate	0.00	0.00	-1.41	21	0.17
Instructor sex (1=Female)	0.11	0.17	0.66	21	0.52
Instructor age	0.02	0.01	2.61	21	0.02
Student-level predictors					
Pen and paper instruction	-0.10	0.11	-0.93	202	0.35
Hands-on instruction	-0.01	0.10	-0.14	202	0.89
Minds-on instruction	0.02	0.07	0.37	202	0.71
Collaborative instruction	0.00	0.07	0.06	202	0.95
Perceived instructor support	0.16	0.07	2.19	202	0.03
Perceived student support	-0.13	0.08	-1.60	202	0.11
Perceived course supports	-0.06	0.05	-1.35	202	0.18
Perceived intellectual difficulty	0.15	0.13	1.16	202	0.25
Instructor choice of discussion topics	-0.06	0.04	-1.51	202	0.13
Student-level controls					
Instructor rating of mastery of content	0.59	0.11	5.43	202	0.00
Instructor rating of quality of participation	0.20	0.09	2.16	202	0.03
Science academic self-efficacy	0.04	0.06	0.63	202	0.53
Perceived value of science	0.00	0.05	-0.06	202	0.95
Bachelor's degree in science field (1=Yes)	-0.10	0.11	-0.91	202	0.36
Student sex (1=Female)	-0.07	0.08	-0.91	202	0.37
10+ years of teaching experience (1=Yes)	0.02	0.10	0.17	202	0.87

NOTE: Grades are standardized and all coefficients are effect sizes (i.e. units are standard deviations of grades).

Among students within the same course, students reporting higher levels of instructor support received higher final course grades than students reporting lower levels of support. The level of student or course supports students perceived and the frequency with which they participated in any of the instructional activities, however, was not significantly related to their final course grade after accounting for the instructor's ratings of their mastery of science content and the quality of their participation in online discussions and other background characteristics.

Mastery of Science Content

The final HLM model for student's mastery of science content (as reported by the instructor) is presented in Table 4. The following course characteristics were significantly related to higher levels of student mastery of science content across courses:

- Lower average levels of perceived instructor support
- Higher average levels of perceived course supports
- More student choice of discussion topics.

Student's mastery of science content ratings did not vary significantly by average levels of student support, any instructional methods, or the cognitive depth of the course materials. Instructor and course supports may counterbalance one another. That is, instructors may design courses with high levels of course support where course materials were organized so students see the connections between concepts so they expect students to rely less on the instructor to support their learning. Another explanation focuses on the meaning instructors attribute to support-seeking behavior—courses in which student seek more support may be seen as lower in mastery than courses in which students seek less support. Student choice of discussion topics may encourage learning by giving students more ownership over the course content.

Table 4
Course-level and student-level variables as predictors of student's mastery of science content (as rated by their instructor)

Predictor	Coef.	SE	t	d.f.	P-value
Intercept	-0.06	0.25	-0.24	21	0.82
Course-level predictors					
Pen and paper instruction	-0.02	0.08	-0.27	21	0.79
Hands-on instruction	-0.12	0.06	-1.88	21	0.07
Minds-on instruction	0.05	0.12	0.44	21	0.67
Collaborative instruction	0.03	0.08	0.46	21	0.65
Aggregated perceived instructor support	-0.34	0.09	-3.94	21	0.00
Aggregated perceived student support	-0.13	0.13	-1.01	21	0.33
Aggregated perceived course support	0.25	0.10	2.54	21	0.02
High cognitive depth (1=Yes)	-0.16	0.18	-0.87	21	0.40
Instructor choice of discussion topics	-0.20	0.08	-2.50	21	0.02
Course-level controls					
Percentage of course grade towards discussions	0.00	0.00	0.05	21	0.96
University program (1=Yes)	-0.07	0.17	-0.40	21	0.69
Percentage of A's	0.00	0.00	0.26	21	0.80
Response rate	-0.01	0.00	-1.70	21	0.10
Instructor sex (1=Female)	0.39	0.22	1.80	21	0.09
Instructor age	0.04	0.01	5.93	21	0.00
Student-level predictors					
Pen and paper instruction	0.02	0.07	0.30	203	0.77
Hands-on instruction	0.00	0.06	0.05	203	0.96
Minds-on instruction	0.06	0.07	0.86	203	0.39
Collaborative instruction	0.06	0.06	0.98	203	0.33
Perceived instructor support	-0.15	0.05	-2.70	203	0.01
Perceived student support	0.11	0.05	2.20	203	0.03
Perceived course supports	0.08	0.04	1.84	203	0.07
Perceived intellectual difficulty	-0.51	0.17	-2.97	203	0.00
Instructor choice of discussion topics	-0.03	0.03	-0.88	203	0.38
Student-level controls					
Final course grade	0.59	0.06	10.31	203	0.00
Science academic self-efficacy	0.06	0.08	0.80	203	0.42
Perceived value of science	-0.12	0.06	-2.18	203	0.03
Bachelor's degree in science field (1=Yes)	0.07	0.13	0.56	203	0.58
Student sex (1=Female)	-0.06	0.07	-0.78	203	0.44
10+ years of teaching experience (1=Yes)	0.24	0.11	2.16	203	0.03

NOTE: Mastery ratings are standardized and all coefficients are effect sizes (i.e. units are standard deviations of mastery ratings).

Within a course, students reporting less instructor support, more student support, and less perceived difficulty of the course materials received higher mastery ratings than their classmates after accounting for their final course grades, academic self-efficacy, perceived value of science, and other background characteristic.

Quality of Participation in Online Discussions

Table 5 presents the final HLM model for the quality of student's participation in online discussions (as rated by their instructor). Across courses, the following course characteristics were associated with higher quality of participation:

- Lower average levels of perceived instructor support
- Higher average levels of perceived course supports
- More student choice of discussion topics

Student's quality of participation ratings did not vary significantly by average levels of student support, any instructional methods, or the cognitive depth of the course materials. While the overall pattern of course-level results is identical to those found for student's mastery of science, the magnitude of the effects differs across outcome. The effect of perceived instructor support is -0.34 for student's mastery of science content and -0.25 for student's quality of participation. Similarly, the positive effect of perceived course supports is stronger for student's quality of participation (coeff=0.33) than their mastery of science content ratings (coeff=0.25). Perhaps the counterbalancing of support differs by outcome—instructors giving more importance to instructor supports in relation to mastery and more importance to course supports in relation to online discussions. Instructors may also interpret student support-seeking behaviors as reflecting more on their mastery than the quality of participation in online discussions (where, in fact, support-seeking may be expected and encouraged). Finally, while the effect of more student choice of discussion topics on student's mastery of science content ratings (coeff=-0.20) was 0.10 S.D. less than the effect on student's quality of participation ratings (coeff=-0.30). This difference in magnitude makes sense if online discussions are a primary site of science knowledge construction in online discussions—the more students feel in control of the discussion, the more they may contribute to its quality.

Table 5
Course-level and student-level variables as predictors of student's quality of participation in online discussions (as rated by their instructor)

Predictor	Coef.	SE	t	d.f.	P-value
Intercept	0.06	0.26	0.21	21	0.83
Course-level predictors					
Pen and paper instruction	-0.02	0.08	-0.20	21	0.85
Hands-on instruction	0.04	0.06	0.69	21	0.50
Minds-on instruction	0.04	0.11	0.35	21	0.73
Collaborative instruction	0.05	0.08	0.66	21	0.52
Aggregated perceived instructor support	-0.25	0.06	-4.34	21	0.00
Aggregated perceived student support	-0.17	0.12	-1.36	21	0.19
Aggregated perceived course support	0.33	0.08	4.24	21	0.00
High cognitive depth (1=Yes)	-0.13	0.15	-0.85	21	0.41
Instructor choice of discussion topics	-0.30	0.06	-4.75	21	0.00
Course-level controls					
Percentage of course grade towards discussions	0.00	0.00	-0.57	21	0.57
University program (1=Yes)	0.01	0.19	0.08	21	0.94
Percentage of A's	0.00	0.00	-1.44	21	0.16
Response rate	-0.01	0.00	-2.10	21	0.05
Instructor sex (1=Female)	0.13	0.21	0.60	21	0.55
Instructor age	0.04	0.01	6.63	21	0.00
Student-level predictors					
Pen and paper instruction	-0.05	0.08	-0.57	207	0.57
Hands-on instruction	-0.09	0.05	-1.79	207	0.07
Minds-on instruction	0.09	0.05	1.82	207	0.07
Collaborative instruction	0.09	0.05	1.97	207	0.05
Perceived instructor support	-0.06	0.05	-1.12	207	0.27
Perceived student support	0.01	0.05	0.25	207	0.81
Perceived course supports	0.10	0.05	1.95	207	0.05
Perceived intellectual difficulty	-0.12	0.13	-0.93	207	0.36
Instructor choice of discussion topics	0.00	0.03	-0.02	207	0.99
Student-level controls					
Instructor rating of frequency of participation	0.68	0.05	12.54	207	0.00
Science academic self-efficacy	0.02	0.05	0.45	207	0.66
Perceived value of science	0.00	0.06	0.06	207	0.95
Bachelor's degree in science field (1=Yes)	0.04	0.05	0.80	207	0.43
Student sex (1=Female)	-0.11	0.10	-1.15	207	0.25
10+ years of teaching experience (1=Yes)	0.36	0.12	3.05	207	0.00

NOTE: Quality of participation ratings are standardized and all coefficients are effect sizes (i.e. units are standard deviations of instructor ratings).

Among students taking the same course, students reporting greater participation in collaborative instructional activities and high levels of course support received higher quality of participation ratings than students reporting less frequent collaboration or perceived course support after accounting for differences in the frequency of their participation on online discussions.

Discussion

In the LSO study we asked:

- What characteristics of online science courses are most strongly associated with positive learning outcomes among science teachers, after accounting for teacher's prior science experiences and demographics?

Based upon previous research, the following findings were hypothesized:

- Students who perceive higher levels of social presence in online courses will have more positive learning outcomes
- Students who perceive higher levels of teaching presence in online courses will have more positive learning outcomes
- Students who engage more frequently in hands-on and minds-on instructional activities will have more positive science learning outcomes

Final Course Grades

Within each course, students who reported receiving more instructor support got higher grades than other students in their course, controlling for their mastery and quality ratings. This supports the hypothesis that instructor support contributes to positive learning outcomes, however it is surprising that neither course supports nor student supports have a significant impact on students' final course grades.

Regarding instructional methods, hands-on and minds-on activities are often thought of as consistent with reform or inquiry-based instructional-methods (Borko, Stecher, & McClam, 2003). Students who were in courses that had more frequent expectations for hands-on and minds-on activities tended to get lower grades than students in courses that did not expect students to engage as frequently in these activities. Students in courses with expectations for frequent pen and paper activities received higher grades than courses with lower frequencies of these activities. Perhaps students are unable to fulfill (or even be aware of) instructor's expectations as well when engaged in reform-based instructional activities. It also may suggest that instructors are more familiar with how to rate achievement in courses with more common activities such as pen and paper problem sets or exams than they are in more inquiry-based courses.

Mastery of Science Content

Students' mastery of science content ratings told a different story. Instructors teaching courses where students report higher average levels of instructor support tended to give lower mastery ratings compared to instructors teaching courses where students reported less average levels of instructor support. Students receiving higher mastery ratings may have relied more on course supports, as high mastery ratings correlated with high levels of course supports. This partially confirms the hypothesis in that course supports are positively correlated with positive learning outcomes, but it is counter to the hypothesis that instructor support actually has a positive impact. It may be that the instructor who does not provide much support is unaware of and tends to overrate their students' mastery level or they may interpret student requests for support as evidence of a lack of mastery. Average levels of perceived student support in a course had no significant impact on students' mastery of science content.

Within a course, students who perceived low levels of instructor support and high levels of student support received higher mastery ratings than other students in their courses. Instructors may have rated students who were independent from the instructor but felt supported by classmates as having more mastery.

The average mastery rating varies across courses depending on who chooses the majority of discussion topics. In courses where the instructor chose most of the topics, students tend to have lower mastery ratings. This may mean that students are having trouble mastering the content when it is not focused on their topics of interest, or conversely, they excel when they are able to choose the path of discussion. Or, instructors may decide to choose more of the discussion topics when they perceive their students as having less mastery of the science content.

Quality of Online Discussion Participation

Instructors also rated the quality of students' participation in online discussions. Like with the mastery rating, hypotheses regarding the relationship of increased supports with higher student outcomes were partially confirmed. With quality ratings, as with master ratings, students in courses with low levels of perceived instructor support and high levels of perceived course supports received higher quality of participation ratings. The students achieving these higher ratings (both mastery and quality) may be independent learners who are in less need of instructor support, but again, appreciate and are able to utilize well-crafted course materials.

The quality of online participation ratings were also higher for students in courses with higher frequencies of collaborative activities. This may be because the collaborative work provided fodder for rich discussion, more so than in courses that didn't provide those activities. Instructors who tend towards collaborative work may also have a different sense with which to evaluate discussions, looking for the quality of teamwork more than the quality of content.

The quality ratings of students were also higher in courses where students chose most of the discussion topics. Again, this suggests that students will invest more, and get more out of, discussions when they are allowed to discuss their related areas of interest.

Summary and Implications

This study examined three types of student outcomes: final grades, mastery of science content, and quality of participation in online discussions, and their relationship to specific course characteristics. Important findings include that:

- High levels of course supports, meaning scaffolding within the materials to help students understand and link concepts, were important for student's mastery of content and quality of participation in online discussions.
- High levels of instructor support were related to higher grades but lower mastery and quality of discussion.
- Students in courses with high frequency of hands-on and minds-on activities tended to get lower grades, while those in courses with high frequency of pen and paper activities tended to get higher grades.

Some overall implications of these findings for developers include the importance of course design for students' success. Courses with supportive structures in place through materials presented opportunities for student achievement, even when there was less instructor support. This suggests that the role of the instructor shifts in online learning to one who must pre-design the course with the student in mind, but then fade the interaction and let students play a significant role in their own learning, including choosing relevant discussion topics.

Online discussion boards were used in over 90% of LSO courses as a primary instructional tool, and students and instructors in these courses also reported high frequency of minds-on activities such as articulating and reflecting on scientific ideas, analyzing evidence and drawing conclusions, and posing scientific questions (Asbell-Clarke & Rowe, 2007). Assessment and evaluation of minds-on instructional activities associated with online discussion may not be as straightforward for instructors, as they appear to give higher grades in courses with more familiar activities such as frequent pen and paper activities and lower grades in courses with higher rates of minds-on activities. This suggests that instructors may need new methods of assessment as well as clear ways to explicate criteria and expectations to their students for online discussion activities.

The prevalence of minds-on instructional activities using online discussion in these courses may also provide a rich opportunity for research on teaching and learning, and suggests that additional research is needed to understand the nature of these discussions and the extent to which they foster useful scientific discourse. Future analyses of the transcripts collected from LSO will examine the nature of the course, instructor, and student supports within these discussions, in particular those that correlate with higher levels of mastery of science content and scientific reasoning.

References

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy, Project 2061*. New York: Oxford University Press.

Asbell-Clarke, J., & Rowe, E. (2007). Learning science online: A descriptive study of 40 online science courses for teachers. *Journal of Asynchronous Learning Networks*, *11*(*3*), 95-121.

Becker, H. (1992). A model for improving the performance of integrated learning systems: Mixed individualized/group/whole class lessons, cooperative learning, and organizing time for teacher-led remediation of small groups. *Educational Technology 32*, 6-15.

Bell, P. (1997). Using argument representations to make thinking visible for individual and groups. In R. Hall, N. Miyake, & N. Enyedy (Eds.), *Proceedings of CSCL '97: The Second International Conference on Computer Support for Collaborative Learning* (pp. 10-19). Toronto: University of Toronto Press.

Bell, P., & Linn, M. C. (1997, March). *Scientific arguments as learning artifacts: Designing for learning on the web*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Bender, T. (2003). Building a model of assessment of online education. In T. Bender (Ed.), *Discussion-based online teaching to enhance student learning* (pp.166-192). Virginia: Stylus Publishing.

Borko, H., Stecher, B. M., & McClam, S. (2003). Artifact packages for measuring instructional practice: A pilot study. *Comparative analysis of current assessment and accountability systems.* Retrieved April 28, 2005, from http://www.cse.ucla.edu/products/Reports/R615.pdf

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

Duffy, T. M., & Jonassen, D. H. (1992). *Constructivism and the technology of instruction: A conversation.* Mahwah, NJ: Lawrence Earlbaum Associates.

DuFour, R. (2005). What is a professional learning community? In R. DuFour, R. Eaker, & R. DuFour (Eds.), *On common ground* (pp. 31–43). Bloomington, IN: Solution Tree.

Edelson, D. C., & O'Neill, D. K. (1994). The CoVis collaboratory notebook: Supporting collaborative scientific inquiry. In A. Best (Ed.), *Proceedings of the 1994 National Educational Computing Conference* (pp.146-152). Eugene, OR: International Society for Technology in Education in cooperation with the National Education Computing Association.

Fahy, P. (2005). Two methods for assessing critical thinking in computer-mediated communication (CMC) transcripts. *International Journal of Instructional Technology in Distance Learning, 2(3)*, 13-28.

Flynn, T., & Polin, L. (2003, April). *Making sense of online learning: Frames, rubrics, and coding systems for analyzing asynchronous online discourse.* Paper presented at Annual Meeting of the American Educational Research Association, Chicago, IL.

Garrison, D. R., & Anderson, T. (2003). *E-Learning in the 21st century: A framework for research and practice.* New York: Routledge Falmer.

Gay, G., Sturgill, A., Martin, W., & Huttenlocher, D. (1999). Document-centered peer collaborations: An exploration of the educational uses of networked communication technologies. *Journal of Computer-Mediated Communication 4(3).*

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist*, *53*, 5–26.

Gunawardena, C., Lowe, C., & Anderson, T. (1997). Analysis of a global on-line debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research, 17(4)*, 395-429.

Harlen, W., & Altobello, C. (2003). *An investigation of "Try Science" studied on-line and face-to-face*. Cambridge, MA, TERC.

Hawisher, G., & Selfe, C. (2000) *Global literacies and the World-Wide-Web*. London: Routledge.

Heller, H., & Kearsley, G. (1996). Using a computer BBS for graduate education: Issues and outcomes. In Z. Berge & M. Collins (Ed.), *Computer-mediated communication and the online classroom.* (pp. 129-137). Creskill, NJ: Hampton Press.

Henri, F. (1991). Computer conferencing and content analysis. In A. Kaye (Ed.) *Collaborative learning through computer conferencing: The Najaden papers,* (pp.117-136). London: Springer-Verlag.

Hiltz, S. R., & Goldman, R. (2005). *Learning together online: Research on asynchronous learning networks*. Mahwah, NJ: Lawrence Earlbaum Associates.

Hogan, K., Nastasi, B. K., & Pressley, M. (2000). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction, 17(4)*, 379-432.

Irving, K. E. (2006). The impact of technology on the 21st Century classroom. In S. Rhoton (Ed.) *Teaching science in the 21st Century*. Arlington, VA, NSTA Press.

Jiang, M., & Ting, E. (2000). A study of factors influencing students' perceived learning in a web-based course environment. *Journal of Educational Telecommunications, 6(4)*, 317-338.

Johnson, D. W., Johnson, R., & Holubec, E. (1998). *Cooperation in the classroom (6th Edition)*. Edina, MN: Interaction Book Company.

Koshmann, T. (1996). *CSCL: Theory of practice of an emerging paradigm*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kuhn, D. (2005). *Education for thinking*. Cambridge, MA, Harvard University Press.

Loucks-Horsley, S., Hewson, P. W., Love, N., & Stiles, K. E. (1998). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin Press.

Loucks-Horsley, S., Stiles, K. E. & Hewson, P. W. (1996). Principles of effective professional development for mathematics and science education: A synthesis of standards. *NISE Brief 1(1)*, 1-6.

Luke, C. (1999). Cyber-schooling and technical change: Multiliteracies for new times. In B. Cope & M. Kalantzis (Eds.), *Multiliteracies: literacy learning and the design of social futures.* Melbourne: MacMillan.

Mehrabian, A. (1969). Attitudes inferred from non-immediacy of verbal communication. *Journal of Verbal Learning and Verbal Behavior, 6*, 294-295.

Millis, B. J., & Cotell, P. G., Jr. (1998). *Cooperative learning for higher education faculty.* Phoenix: Oryx Press.

Nathan, M. J., & Knuth, E. J. (2003). A study of whole classroom mathematical discourse and teacher change. *Cognition and Instruction, 21(2)*, 175-207.

National Center for Education Statistics (2005). *Digest of education statistics, 2005*. Retrieved February 21, 2007, from http://nces.ed.gov/programs/digest/d05/

National Research Council. (1996). *National Science Education Standards.* Washington, D.C., National Academy Press.

National Research Council. (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

National Research Council. (2005). *How students learn: History, mathematics, and science in the classroom*. Washington, DC: The National Academies Press.

Packer, M. J., & Goicoechea, J. (2000). Sociocultural and constructivist theories of learning: Ontology, not just epistemology. *Educational Psychology*, *35*, 227–241.

Pena-Perez, J. (2000). *Participation, interaction and meaning construction in a university-level course using a computer bulletin board as a supplement to regular class discussions. A case study.* Unpublished doctoral Dissertation, Cornell University, Ithaca, NY.

Picciano, A. (1998). Developing an asynchronous course model at a large, urban university. *Journal of Asynchronous Learning Networks, 2(1)*. Retrieved April 28, 2005, from http://www.sloan-c.org/publications/jaln/

Polin, L. (2004). Learning in dialogue with a practicing community. In T. Duffy & J. Kirkley (Eds.), *Learner centered theory and practice in distance education: Cases from higher education* (pp. 17-48). Mahwah, NJ: Lawrence Earlbaum Associates..

Richardson, J. C., & Swan, K. (2003) Examining social presence in online courses in relation to students' perceived learning and satisfaction. *Journal of Asynchronous Learning Networks, 7*(1). Retrieved April 28, 2005 from http://www.sloan-c.org/publications/jaln/

Roberts, T. S. (2004). *Online collaborative learning: Theory and practice*. Hershey, PA: Information Science Publishing.

Ruberg, L., Moore, D., & Taylor, D. (1996). Student participation, interaction, and regulation in a computer-mediated communication environment: A qualitative study. *Journal of Educational Computing Research, 14(3)*, 243-268.

Sadler, T. (2006) Promoting discourse and argumentation in science teacher education. *Journal of Science Teacher Education*, *17*, 323–346.

Scarafiolti, C., & Cleveland-Innis, M. (2006). The times they are a-changing. *Journal of Asynchronous Learning Networks, 10(2)*.

Scardamalia, M., & Bereiter, C. (1991). Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *Journal of the Learning Sciences, 1(1)*, 37-68.

Scardamalia, M., & Bereiter. C. (1996). Computer support for knowledge-building communities. In T. Koschmann (Ed.) *CSCL: Theory and practice of an emerging paradigm* (pp. 249–268). Mahwah, NJ: Lawrence Erlbaum Associates.

Scardamalia, M., Bereiter, C., & Lamon, M. (1994). The CSILE project: Trying to bring the classroom into world 3. In K. McGilly (Ed.), *Classroom practice.* (pp. 201-228). Cambridge, MA: MIT Press.

Scardamalia, M., Bereiter, C., McLean, R. S., Swallow, J., & Woodruff, E. (1989). Computer supported intentional learning environments. *Journal of Educational Computing Research*, *5*, 51-68.

Schau, C., Stevens, J., Dauphinee, T., & Del Vecchio, A. (1995). The development and validation of the Survey of Attitudes Toward Statistics. *Educational & Psychological Measurement, 55(5)*, 868-876.

Sefton-Green J. (1998). *Digital diversions: Youth culture in the age of multimedia.* London: University College London Press.

Shea, P., Pickett, A., & Pelz, W. (2003). A follow-up investigation of “teaching presence” in the

SUNY learning network. *Journal of Asynchronous Learning Networks, 7(2)*, 61-80.

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. Toronto: Wiley.

Simon, M., & Forgette-Giroux, R. (2001). A rubric for scoring postsecondary academic skills. *Practical Assessment, Research & Evaluation, 7(18)*.

Skeele, R., Walker, E., & Klemballa, C. (2004, April). *Preservice Teachers' Academic Efficacy and Attitudes toward Teaching Science, Mathematics, and Technology*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Sloan Consortium. (2004). *Entering the Mainstream: The Quality and Extent of Online Education in the United States, 2003 and 2004*. Retrieved December 20, 2006 from http://www.sloan-c.org/publications/survey/pdf/entering_mainstream.pdf

Sloan Consortium. (2006). *Making the Grade: Online Education in the United States, 2006.* Retrieved December 20, 2006 from http://www.sloan-c.org/publications/survey/pdf/making_the_grade.pdf

Swan, K. (2000). Non-print media and technology literacy standards for assessing technology integration. *Journal of Educational Computing Research*, *23(1)*, 85-100.

Swan, K., & Shea, P. (2005). The development of virtual learning communities. In. S. R. Hiltz & R. Goldman (Eds.), *Learning together online: Research on asynchronous learning networks* (pp. 239-260). Mahwah, NJ: Erlbaum.

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

Valacich, J., Dennis, A., & Connolly, T. (1994). Idea generation in computer based groups: A new ending to an old story. *Organizational Behavior and Human Decision Processes, 57*, 448-467.

Weaver, C., & Albion, P. (2005, December). *Momentum in online discussions: The effect of social presence on motivation for participation*. Paper presented at the 22nd Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education, Brisbane, Australia.

Wenger, E. (1997). *Communities of practice: Learning, meaning, and identity*. New York: Cambridge University Press.

Appendix A

Course scales, items, and alpha coefficients

Pen and Paper Instructional Activities (α=0.62)
1. Worked with pen and paper problem sets (e.g., problems from a textbook or worksheet)
2. Students expected to use books as part of this course.
3. Students expected to use calculators as part of this course.

Scale: 1=Not at all; 2=Once or twice in the course; 3=Once or twice a month; 4=Once or twice a week; 5=Three times a week or more.

Hands-On Instructional Activities (α=0.82)
1. Interacted with physical materials or models (e.g., mixing solutions, building circuits, scale models)
2. Carried out procedures of scientific investigations designed by instructors or course developers (e.g., lab exercises, kitchen experiments)
3. Designed their own scientific investigation(s) (e.g. developed hypothesis or question and procedures)
4. Carried out procedures of scientific investigations they designed (e.g. collected data, made observations)

Scale: 1=Not at all; 2=Once or twice in the course; 3=Once or twice a month; 4=Once or twice a week; 5=Three times a week or more.

Minds-On Instructional Activities (α=0.79)
1. Articulated their scientific ideas in an on-line discussion.
2. Reflected upon their earlier scientific ideas
3. Reflected upon the scientific ideas of other students.
4. Raised questions with other students about their scientific ideas.
5. Analyzed and drew conclusions from data, observations, and other forms of scientific evidence
6. Provided evidence to support their scientific ideas.

Scale: 1=Not at all; 2=Once or twice in the course; 3=Once or twice a month; 4=Once or twice a week; 5=Three times a week or more.

Collaborative Instructional Activities (α=0.87)
1. Worked as part of a team on group projects or assignment
2. Worked as part of a small student group created to discuss course content
3. Worked as part of a small student group created to complete assignments or activities
4. Worked as part of a small student group created to review each other’s work

Scale: 1=Not at all; 2=Once or twice in the course; 3=Once or twice a month; 4=Once or twice a week; 5=Three times a week or more.

Instructor Choice of Discussion Topics
1. Which of the following best describes who chose new topics (i.e. initiated a new thread) within the on-line discussions in this course?
Scale:
1=Students chose nearly all of the new topics within the on-line discussions.
2=Students chose most of the new topics within the on-line discussions. Instructors chose a few of the new topics within the on-line discussions.
3=Students and instructors were equally likely to have chosen new topics within the on-line discussions.
4=Instructors chose most new topics within the on-line discussions. Students chose a few of the new topics within the on-line discussions.
5=Instructors chose nearly all new topics within the on-line discussions.

Appendix B

Student scales, items, and alpha coefficients

Pen and Paper Instructional Activities (α=0.77)

1. Worked with pen and paper problem sets (e.g., problems from a textbook or worksheet)
2. Used books as part of this course.
3. Used calculators as part of this course.

Scale: 1=Not at all; 2=Once or twice in the course; 3=Once or twice a month; 4=Once or twice a week; 5=Three times a week or more.

Hands-On Instructional Activities (α=0.84)

1. Interacted with physical materials or models (e.g., mixing solutions, building circuits, scale models)
2. Carried out procedures of scientific investigations designed by instructors or course developers (e.g., lab exercises, kitchen experiments)
3. Designed my own scientific investigation(s) (e.g. developed hypothesis or question and procedures)
4. Carried out procedures of scientific investigations I designed (e.g. collected data, made observations)

Scale: 1=Not at all; 2=Once or twice in the course; 3=Once or twice a month; 4=Once or twice a week; 5=Three times a week or more.

Minds-On Instructional Activities (α=0.65)

1. Articulated my scientific ideas in an on-line discussion.
2. Reflected upon my earlier scientific ideas
3. Reflected upon the scientific ideas of other students.
4. Raised questions with other students about their scientific ideas.
5. Analyzed and drew conclusions from data, observations, and other forms of scientific evidence
6. Provided evidence to support my scientific ideas.

Scale: 1=Not at all; 2=Once or twice in the course; 3=Once or twice a month; 4=Once or twice a week; 5=Three times a week or more.

Collaborative Instructional Activities (α=0.74)

1. Worked as part of a team on group projects or assignment
2. Worked as part of a small student group created to discuss course content
3. Worked as part of a small student group created to complete assignments or activities
4. Worked as part of a small student group created to review each other’s work

Scale: 1=Not at all; 2=Once or twice in the course; 3=Once or twice a month; 4=Once or twice a week; 5=Three times a week or more.

Perceived Intellectual Difficulty
1. Intellectually, the course materials were…
Scale: 1=too easy for me; 2=just right for me; 3=too difficult for me

Perceived Instructor Support (α=0.88)
1. Interactions with the instructor helped me understand the course material better.
2. I felt supported by the instructor(s) as I developed my understanding of the course material.
3. I felt my contributions to the on-line discussions were valued by the instructor.
4. The instructor(s) was accessible to me.
5. I was encouraged to provide feedback to the instructors about my questions and concerns about the course.
Scale: 1=Strongly Disagree; 2=Disagree: 3=Neutral; 4=Agree; 5=Strongly Agree

Perceived Student Support (α=0.80)
1. Interactions with the other students helped me understand the course material better.
2. I felt supported by other students as I developed my understanding of the course content.
3. I felt my contributions to the on-line discussions were valued by other students.
4. The class atmosphere encouraged me to make contributions to the on-line discussions.
Scale: 1=Strongly Disagree; 2=Disagree: 3=Neutral; 4=Agree; 5=Strongly Agree

Perceived Course Supports (α=0.77)
1. I usually understood the content being taught in the course.
2. The course seemed to be designed to address multiple learning styles
3. I felt my learning style was well suited for this course.
4. Course materials were organized so that each new concept built upon previous learning.
5. Course materials were organized so it was clear how different concepts covered in this course fit together.
Scale: 1=Strongly Disagree; 2=Disagree: 3=Neutral; 4=Agree; 5=Strongly Agree

Instructor Choice of Discussion Topics
1. Which of the following best describes who chose new topics (i.e. initiated a new thread) within the on-line discussions in this course?
Scale:
1=Students chose nearly all of the new topics within the on-line discussions.
2=Students chose most of the new topics within the on-line discussions. Instructors chose a few of the new topics within the on-line discussions.
3=Students and instructors were equally likely to have chosen new topics within the on-line discussions.
4=Instructors chose most new topics within the on-line discussions. Students chose a few of the new topics within the on-line discussions.
5=Instructors chose nearly all new topics within the on-line discussions.

Acknowledgements

This study was supported by funding from the National Science Foundation, Gender in Science and Engineering program (Grant HRD-0332602).