Design-Based Research Strategies for Studying Situated Learning in a Multi-user Virtual Environment

Chris Dede, Brian Nelson, Diane Jass Ketelhut, Jody Clarke & Cassie Bowman
Graduate School of Education, Harvard University, Cambridge, MA 02138
Tel: 617-495-3839; Fax: 617-495-9268
Chris_Dede@.harvard.edu Brian_Nelson@gse.harvard.edu ketelhdi@gse.harvard.edu
clarkejo@gse.harvard.edu bowmanca@gse.harvard.edu

Abstract: This National Science Foundation funded project utilizes graphical multi-user virtual environments (MUVEs) as a vehicle to study (1) classroom-based situated learning and (2) the ways in which virtual environments may aid the transfer of learning from classroom contexts into real world settings. In the project's River City curriculum, teams of middle school students are asked to collaboratively solve a simulated 19th century city's problems with illness, through interaction with each others' 'avatars', digital artifacts, tacit clues, and computer-based 'agents' acting as mentors and colleagues in a virtual community of practice. This paper describes the design-based research strategy by which we are currently extending an educational MUVE environment and curriculum developed with prior NSF funding. We are implementing a series of studies to determine if such virtual environments can sufficiently replicate authentic contexts and multi-leveled communities of practice to provide students with classroom experiences in situated learning.

Introduction

Scientific literacy—the capabilities (1) to understand the interrelationships among the natural world, technology, and science and (2) to apply scientific knowledge and skills to personal decision-making and the analysis of societal issues—is a major goal for education in the 21st century (AAAS, 1993; NRC, 1996). Research suggests that, if all students are to become scientifically literate citizens, science instruction must convey greater engagement and meaning to them. To achieve this, we believe that science instruction in secondary schools must provide students with opportunities to explore the world; to apply scientific principles; to sample and analyze data; and to make connections among these explorations, their personal lives, and their communities. However, given the constraints of classroom settings, real world data collection is challenging to orchestrate. Laboratory experiments are also difficult to conduct due to lack of equipment and safety issues. It is no surprise that teachers report hypothesis formation and experimental design are among the most difficult challenges they face with students who have a history of low achievement in science.

With prior NSF funding, we have created and studied graphical multi-user virtual environments (MUVEs) that use digitized museum resources to enhance middle school students' motivation and learning about science and society (http://www.gse.harvard.edu/~dedech/muvees/). Multi-user virtual environments enable multiple simultaneous participants to access virtual contexts, to interact with digital artifacts, to represent themselves through "avatars," to communicate with other participants and with computer-based agents, and to enact collaborative learning activities of various types. Our "River City" MUVE is centered on skills of hypothesis formation and experimental design, as well as on content related to national standards and assessments in biology and ecology.

Because MUVE research is still in its infancy, we are employing a design-based research (DBR) approach to the development of the River City environment and curriculum. Design-based research was first detailed by Brown (1992) and Collins (1992) as a way to study new learning environments in the context of the classroom. "Design experiments were developed as a way to carry out formative research to test and refine educational designs based on principles derived from prior research" (Collins, Joseph, & Bielaczyc, in press p. 4). We are implementing a series of studies to determine if virtual environments can sufficiently replicate authentic contexts and multi-leveled communities of practice to provide students with classroom experiences in situated learning. In this paper, we conduct a meta-level reflection of our design-based methodology and aims.

By offering a 'glass-box' view into our research strategy, we hope to provide a guide to others interested in conducting design-based research.

Design-based Research

The scholarly community is still defining design-based research. The purpose of creating such a methodology was to move educational research from laboratories into classrooms in order to gain insight into how, when, and why innovations work in practice (The Design-Based Research Collective, 2003). Yet the focus of DBR is not on specific designs and curricula, but on how the strengths and limits of a design inform theories of learning. DBR combines quantitative and qualitative methods to view how designs work in the crucible of practice and to gain insights into how students learn in typical school contexts. The DBR Collective has identified 5 key characteristics of good design-based research:

- Goals of designing learning environments and developing theories or 'prototheories' of learning are intertwined.
- Development and research take place through continuous cycles of design, enactment, analysis, and redesign.
- Research leads to theories that communicate relevant implications to practitioners and other designers.
- Research accounts for how designs function in authentic settings, not only documenting success or failure, but also focusing on interactions that refine our understanding of the learning issues involved.
- Research relies on methods that can document and connect process of enactment to outcomes of interest.

Moreover, Dede (2004) has articulated four challenges that good DBR should address:

- Collect a manageable amount of data that is focused on evaluating the consequences of key design choices.
- Separate design from conditions for success by ensuring that the latter are present in the pilot implementation.
- Determine what constitute reasonable criteria for "success" in declaring a design finished.
- Develop designs that meet the needs of practitioners and policymakers in the context of the No Child Left Behind initiative.

Our research focuses on strategies to meet these characteristics and challenges in evolving the River City environment and curriculum via DBR.

First Implementation

The initial pilot study of our project's 'River City' educational environment extended typical MUVE capabilities in order to study the science learning potential of immersive simulations, interactive virtual museum exhibits, and "participatory" historical situations. Our goal was to promote learning for all students, particularly those unengaged or low performing. Using a guided social constructivist design, students learned to behave as scientists while they collaboratively identified problems through observation and inference, formed and tested hypotheses, and deduced evidence-based conclusions about underlying causes.

The River City virtual "world" consists of a city with a river running through it; different forms of terrain that influence water runoff; and various neighborhoods, industries, and institutions, such as a hospital and a university (muve.gse.harvard.edu/muvees2003). The learners themselves populate the city, along with computer-based agents, digital objects that can include audio or video clips, and the avatars of instructors (Figure 1). Content in the right-hand interface-window shifts based on what the participant encounters or activates in the virtual environment (Figure 2). Dialogue is shown in the text box below these two windows; members of each team can communicate regardless of distance, but intra-team dialogue is displayed only to members of that team.

In our pilot research, students worked in teams to develop hypotheses regarding one of three strands of illness in River City (water-borne, air-borne, and insect-borne). These three disease strands are integrated with historical, social and geographical content to allow students to experience the realities of disentangling multi-causal problems embedded within a complex environment. At the end of the project, students compared their research with other teams of students to discover the many potential hypotheses and avenues of investigation to explore.

Findings

In our pilot implementations of River City, using two public school classrooms in Boston, MA, we examined usability, student motivation, student learning, and classroom implementation issues (Dede & Ketelhut, 2003). One sixth- and one seventh-grade classroom in different schools with high percentages of ESL students implemented the River City curriculum; control classrooms used a similar curriculum delivered via paper-based materials rather than technology.

Using design-based research methods, we collected both qualitative and quantitative data from students and teachers over the three-week implementation period. Both the Patterns for Adaptive Learning Survey (Midgley, 2000) and a content test were administered to students, pre- and post-intervention. In addition, demographic data and teachers' expectations of students' successes were collected. All teachers responded to a pre and post questionnaire regarding their methods and comfort with technology. The experimental intervention classroom teachers also wrote a narrative at the end of the project about their perceptions of the MUVE. We used this data to analyze the learning outcomes for students and to inform our understanding of how the MUVE worked, in order to refine the design of our next iteration. Students found the MUVE interface readily usable and the learning experiences motivating, even after repeated exposures.

Preliminary results indicated the MUVE is motivating for all students, including lower ability students typically uninterested in classroom activities. Six out of seven experimental students scoring less than 35% on the content pre-test improved their content knowledge above that level, while only two of five control students did so. We also found that students discovered multiple intriguing situations in the MUVE to investigate. In the seventh grade classroom, five different hypotheses about these situations emerged, with posited causes ranging from population density to immigration to water pollution. Another finding is that the MUVE seemed to have the most positive effects for students with high perceptions of their own thoughtfulness of inquiry. These students, on average, scored higher on the post content test, controlling for SES, science GPA, ethnicity, and content pre-test score (Dede & Ketelhut, op cit).

Overall, these findings encourage further refinement and experimentation with curricular MUVEs as a learning modality that can help teachers reach students struggling with motivation, self-worth, and lack of content knowledge. These data are promising, but not conclusive about the curriculum's educational value or the effectiveness of MUVEs for learning. By examining student interactions with the pilot curriculum, we saw ways to strengthen our content and pedagogy. Despite the fact that students found the MUVE readily usable and the learning experiences motivating, we found weaknesses in this design, both from a graphical and curricular perspective.

Design Outcomes

Based on these findings, we refined our underlying learning theory. In the pilot, students were immersed in conducting an authentic task, similar to the goals of situated learning. Consequently, the next iteration of our project was to extend our MUVE curriculum to study classroom-based situated learning in comparison to guided social constructivist pedagogy.

MUVEs are a promising medium for creating and studying situated learning because they can support immersive, extended experiences, incorporating modeling and mentoring, about problems and contexts similar to the real world. Our MUVE is based on a problem-solving community in which students gain knowledge and skills through co-interpreting data with other participants at varied levels of skills. Our current research is studying whether the River City MUVE can replicate authentic contexts and multi-leveled communities of practice sufficiently well to provide students and teachers with classroom experiences in situated learning.

We also found that teachers need more support for implementing this curriculum. Although both pilot teachers professed to have a constructivist pedagogical approach, they struggled to find their role in the classroom—one feeling as if she should be more involved, the other less. As a result, we are creating a rigorous 8-hour program of professional development for our teachers.

Research Context

Reports such as the National Research Council's study, *How People Learn: Extended Edition* (2000) delineate alternative theoretical constructs for understanding teaching and learning. The three major schools of thought cited are behaviorist theories of learning (e.g., presentational instruction), cognitivist theories of learning (e.g., intelligent tutoring systems, guided constructivist instruction), and situated theories of learning (e.g., mentoring and apprenticeships in communities of practice).

The least studied of these perspectives is situated learning, in part because conducting controlled experiments about tacit, relatively unstructured learning in complex real-world settings is very difficult. Several studies (Griffin, 1995; Young, 1993; Hendricks, 2001) attempted to validate parts of situated learning theory, but were forced to modify the theoretical intent to match constraints of the classroom. For example, the Jasper Series (Young, 1993; Cognition and Technology Group at Vanderbilt, 1997) situates mathematics problems in an authentic context, but the learner is still a student in a classroom rather than situated in the same context (Tripp, 1993).

Situated Learning Theory

Brown, Collins, & Duguid (1989) and Lave & Wenger (1991) define situated learning theory as embedded within and inseparable from participating in a system of activity (Chaiklin & Lave, 1993; Lave, 1988) deeply determined by a particular physical and cultural setting. The unit of analysis is neither the individual nor the setting, but instead the relationship between the two, as indicated by the student's level of participation in the setting (Barab & Plucker, 2002). Studies of apprenticeship in "communities of practice" (moving from newcomer to expert within a sociocultural structure of practices) are a central construct for situated learning (McLellan, 1996; Kirshner & Whitson, 1997; Wenger, 1998; Wenger, McDermott, & Snyder, 2002). Brown, Collins and Duguid (1989) proffer the graduate student experience as an example of apprenticeship in a community of practice. As part of their academic program, graduate students evolve from pupils to researchers through a series of learning activities embedded within a specific community of practice.

In essence, situated learning requires authentic contexts, activities, and assessments coupled with guidance based on expert modeling, situated mentoring, and legitimate peripheral participation. For example, graduate students may work within the laboratories of expert researchers, who model the practice of scholarship. These students will interact with experts in research, as well as with other members of the research team who understand the complex processes of scholarship to varying degrees. While in these laboratory settings, students gradually move from novice researchers to more advanced roles, with the skills and expectations for them evolving (legitimate peripheral participation). In contrast to courses, students learn the knowledge and skills expected of them in their future research careers through modeling, mentoring, and legitimate peripheral participation.

MUVEs as a Research Method for Assessing Theories of Situated Learning

Our current River City curriculum unit is based on students participating in an elaborate context modeled on the real world, interacting with novices and experts who are part of its culture. Learners are not passively observing this situation, but actively investigating multivariate problems with aid from community members who have various types of expertise. The digital artifacts, computer-based agents, avatars, and virtual contexts in our MUVE can simulate the communities of practice envisioned by Lave & Wenger (1991). Proponents of situated theory approach mentoring in two different ways, either through direct guidance from an expert or through the legitimate peripheral participation model proposed by Lave & Wenger (1991). The first, the method used by Griffin (1995), employs an expert to model and then coach students while they learn the techniques. The expert is not employing the techniques for his/her own benefit, but is using skills to show novices how to apply them. In contrast, within the legitimate peripheral participation model, the entire community of practice is working on the product, and novices learn as much or more from observation of other, somewhat more advanced participants as by direct guidance from an expert. One advantage of a MUVE is that it can be modified to favor either form of guidance, allowing comparison of the two models.

Research Strategy

Based upon what we learned from our prior design-based research, we are developing three variations of the River City curriculum. Variant GST centers on a guided social constructivist (GST) model of learning, in which guided inquiry experiences in the MUVE alternate with whole-class interpretive sessions. This variant is currently undergoing testing and evaluation while the other two are developed. Variant EMC shifts the learning model to center on expert modeling and coaching (EMC), based on expert agents embedded in the MUVE and experts collaborating with teachers in facilitating the whole-class interpretive sessions. Variant LPP shifts the learning model to focus on legitimate peripheral participation (LPP), in which the entire community of practice in the MUVE works on problem-solving, and students learn more from observation of other, somewhat more advanced participants (avatars, computer-based agents) than via direct guidance by experts. Our fourth "control" condition utilizes a curriculum in which the same content and skills are taught in equivalent time to comparable students in a paper-based format without technology, using a guided social constructivist-based pedagogy. Where possible, teachers offer both the variants and the control.

As they are developed, we are implementing these alternative curricula in small group settings to allow us to go through continuous cycles of design, enactment, analysis, and redesign, as suggested by the DBR collective (2003). As advocated by Dede (2004), we will determine the necessary criteria to declare a design finished. At that point, later this spring, we will conduct a controlled implementation on a larger scale in Boston- and Milwaukee-area classrooms with high proportions of ESL and free-and-reduced-lunch students, matching experimental to control classrooms to determine the relative efficacy of our approach. To control for threats to validity, all three variants will be randomly assigned to students within a single classroom, with teachers instructed to minimize cross-contamination of treatments. We are also creating and delivering approximately eight hours of professional development for teachers, focused on content review, alternative pedagogical strategies based on different theories of learning, facilitation strategies while students are using the MUVE, and interpretive strategies for leading class discussions.

Research Questions

1. Can a MUVE simulate authentic contexts and multi-leveled communities of practice within a classroom setting? For variants EMC and LPP, to what extent do constellations of architectural, social, organizational, and material vectors that shape the setting provide a context that alters students' sociocultural interactions and practices of learning to patterns characteristic of situated learning rather than guided social constructivism (version GST)?
2. When compared to the "control" version, what types of contrasting, significant gains in motivation and learning (if any) for various kinds of students do versions GST, EMC, and LPP produce? What are the complementarities (if any) between legitimate peripheral participation and expert mentoring as methods of situated learning?

Both in and out of MUVEs, insights obtained into these questions may enhance educators' and researchers' understanding and application of situated learning theories and may increase students' abilities to transfer knowledge from academic to real world settings.

Results from Two Cycles of DBR

Our first implementation of the current MUVE curriculum was held in a focus group in December, 2003. At that time, we implemented changes based on our analysis of the first River City pilot study. Our graphical interface had been improved to allow students to teleport, to have dialogues with computer agents, and to select their avatar—all suggestions made by pupils in the first pilot study. By observing and conducting interviews with focus group participants, these changes seemed to elicit 'ah-ha' moments for students. We also actively solicited focus group suggestions of changes students would like to see. From observations and these interviews, we learned:

Students needed time to experience the world before beginning the formal curriculum. This experience helps them to become immersed in the context.

The MUVE interface's option of having two communication modes -- a chat and a whisper function -- was confusing to students. As a result, most of them relied on the whisper function, which interfered with group collaborative work.

Students were confused by the connection and relevance of the digitized Smithsonian artifacts in the world.

Some students became easily lost in the world.

Students sought to access the books in the virtual library of River City when they were confused.
Students wondered why, if their avatars were in the world, they too weren't getting sick.

We made design changes based on this feedback (see below) and then conducted a full-scale pilot study in January, 2004. As we implemented this new version, we determined that our alterations made significant improvements to the curriculum likely to result in improvements in student engagement and learning outcomes.

We reorganized our lab book to allow students time to negotiate the interface, to learn how to maneuver, and to constructively explore the world. We found that students were immediately engaged and used their exploration time both to become comfortable with the MUVE interface and to start understanding what problems existed in River City. When we handed out the lab book on the second day, the students used this to guide their investigation more readily than they had previously.
We scaffolded proper use of the chat and whisper functions, renaming them: "Group chat" and "River City Resident Chat". Students now were able to identify their team members more easily using the "group chat" and used that in addition to the whisper mode, allowing for greater team collaboration.
We added a section to our lab book that guides students in understanding the digital images and artifacts embedded in the world. Prior to creating this new section, we found that students were more likely to rely only on computer-based agents to understand the problems in River City. Since much of the curriculum is attached to embedded artifacts, students were limiting their understanding. After creating this section, students interacted with the digital pictures more than previously, thus increasing their involvement with the curriculum.
We created a permanent link on the interface to our map that we had previously made interactive. The ease of finding where they were or where they wanted to go allowed students to access more of the curriculum than before, when it was more difficult to maneuver.
Realizing that allowing students to locate background information on disease and on the scientific method would strengthen their learning outcomes, we created several clickable volumes in the library: a dictionary, encyclopedia, primers on microbes and scientific method, and background in pertinent algebraic concepts. Once the students discovered that they could "find answers" in the library (which to many middle-school students is another ah-ha moment!), it became a popular place to get more information.
Although we would like to let students' avatars get sick, this is not yet technologically possible in our authoring system. Instead, we created a health meter that would rise and fall as students wandered close to polluted waters or stepped on manure. This health meter gives students an additional source for data in their experimentation. As they walked through the world, the movement of the meter intrigued students, and they sometimes explored ways to make it change. As a result, students realized where the pollution and contamination in River City was at an earlier stage.

The outcomes of these changes in our January implementation appear to enhance student learning. We are cognizant of trying not to make changes just to improve the design, but instead to create different avenues for student investigation that fit different learning styles. In a typical middle-school classroom faced with a diverse set of learning styles, the teacher must alternate pedagogical strategies to aid each of these. Even under the best of circumstances, at any single moment some students' learning styles block them from understanding the lesson. In a MUVE, students can individualize their learning based on their own styles.

Conclusion

Because of the infancy of MUVE research, and because we are still evolving curriculum and articulating insights about MUVE-based situated learning in the crucible of practice, we are using a design-based research (DBR) approach in our study. By engaging in DBR-inspired cycles of design, implementation, analysis, and redesign, we have been able to refine both our curriculum and the MUVE environment prior to conducting formal randomized experimental trials. In each iteration of our study, quantitative data are revealing findings about the relative effectiveness of learning theories as instantiated in our environment/curriculum, and qualitative data are providing insights about the reasons underlying those comparative differences.

After several study iterations, substantial parts of the design (e.g., strategies for student motivation) have remained relatively unchanged from the initial implementation because our analysis indicated these were successful in meeting our objectives. Other parts of the design have changed based on feedback from the initial studies. For example, we are increasing the number of tacit clues embedded in the learning environment to ensure that high academic achievement students find continued interactions with the curricular unit challenging and interesting. Our quantitative studies, including the use of a control curriculum, are helping us determine whether the leverage for learning and engagement provided by our work are substantial enough to merit moving beyond DBR to large-scale experimental research on implementation.

An important emphasis in our research is assisting students across the spectrum of academic achievement gain in motivation, self-efficacy, and science knowledge and skills. In particular, educators need help in engaging and teaching subpopulations of learners with special needs, students unmotivated by standard instructional approaches, and pupils with learning styles more visual and kinesthetic than symbolic and auditory. Our environment/curriculum is targeted specifically to narrowing the gaps among students by helping all learners reach their full potential via methods particularly applicable to students who are currently underperforming because of how they are taught in conventional classroom settings. DBR methodologies are providing a way of identifying which elements of our curriculum and MUVE environment are best suited to this goal.

We believe that this type of controlled evolution of DBR is important to its acceptance as a legitimate methodology by the conservative end of the scholarly community in education. We hope to contribute to the field and legitimacy of DBR by sharing our development strategy in the context of our MUVE science curriculum.

Figures

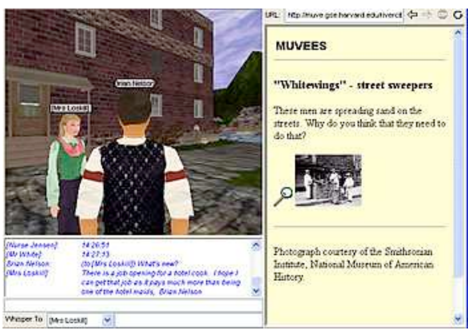

Figure 1: Talking with an Agent

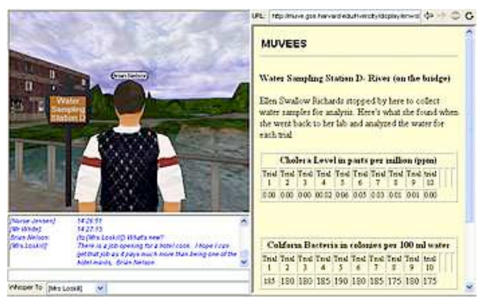

Figure 2: Collecting Water Quality Data

References

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

Barab, S. A., & Plucker, J. A. (2002). Smart People or Smart Contexts? Cognition, Ability, and Talent Development in an Age of Situated Approaches to knowing and Learning. *Educational Psychologist, 37*(3), 165-182.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions. *Journal of the Learning Sciences, 2* (2), 141-178.

Brown, J. S., Collins, A., & Duguid. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher, 18*(1), 32-42.

Chaiklin, S., & Lave, J. (1993). *Understanding practice: Perspectives on activity and context*. New York: Cambridge University Press.

Cognition and Technology Group at Vanderbilt. (1997). The Jasper project: Lessons in curriculum, instruction,assessment, and professional development. Hillsdale, NJ.

Collins, A., Joseph, D., & Bielaczyc, K. (in press). Design research: Theoretical and methodological issues. In J. Campione (Ed.), *Volume in honor of Ann Brown.*

Collins, A. (1992) Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.) *New directions in educational technology*. Berlin: Springer-Verlag, 1992.

Dede, C. (2004). If Design-Based Research is the Answer, What is the Question? *Journal of the Learning Sciences, 13(*1), 105-114.

Dede, C., & Ketelhut, D. (2003). *Motivation, Usability, and Learning Outcomes in a Prototype Museum-based Multi-User Virtual Environment*. Presented at American Educational Research Conference April, 2003.

The Design-Based Research Collective. (2003). Design-Based Research: An Emerging Paradigm for Educational Inquiry. *Educational Researcher, 32*(1), 5-8.

Griffin, M. M. (1995). You Can't Get There from Here: Situated Learning, Transfer, and Map Skills. *Contemporary Educational Psychology, 20*, 65-87.

Hendricks, C. C. (2001). Teaching Causal Reasoning Through Cognitive Apprenticeship: What are Results from Situated Learning? *Journal of Educational Research, 94*(5), 302-311.

Kirshner, D., & Whitson, J. A. (Eds.). (1997). *Situated cognition: Social, semiotic, and psychological perspectives*. Mahway, N.J.: Lawrence Erlbaum.

Lave, J. (1988). *Cognition in practice***.** Cambridge, UK: Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

McLellan, H. (Ed.). (1996). Situated learning perspectives. Englewood Cliffs, NJ: Educational Technology Publications.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., Gheen, M., Kaplan, A., Kumar, R., Middleton, M. J., Nelson, J., Roeser, R., & Urdan, T. (2000*). Manual for the Patterns of Adaptive Learning Scales (PALS)*, Ann Arbor, MI: University of Michigan

National Research Council (2000). *How People Learn: Brain, Mind, Experience and School: Extended Edition*. Washington, DC.

National Research Council. (1996). National science education standards. Washington, DC.

Tripp, S. D. (1993). Theories, Traditions, and Situated Learning. *Educational Technology, 33*(3), 71-77.

Wenger, E. (1998). Communities of practice. New York: Cambridge University Press.

Wenger, E., McDermott, R., & Snyder, W. (2002). Cultivating Communities of Practice. Cambridge, MA: Harvard Business School Press.

Young, M. F. (1993). Instructional Design for Situated Learning. *Educational Technology Research & Development, 41*(1), 43-58.