

Fostering Reflection with Socratic Tutoring Software: Results of Using Inquiry Teaching Strategies with Web-Based HCI Techniques

Baba Kofi A. Weusijana, Christopher K. Riesbeck & Joseph T. Walsh, Jr.
Learning Sciences / Computer Science / BioMedical Engineering, Northwestern University
1890 Maple Ave, Evanston, IL 60201
Tel: (847) 467-1619, Fax: (847) 491-5258
Email: k-weusi@northwestern.edu

Abstract: Often a lesson requires learners to reflect on their learning task, to stop and think about what they are doing, how, and why. One way educators can foster this kind of metacognition is to utilize the Socratic Method; ask learners probing questions to get them to examine their own thinking and assist them in an inquiry process. When there are not enough educators skilled in facilitating Socratic dialogue available to help all the learners in a learning environment in a timely manner, we propose using Web-based Socratic tutoring software to assist with the workload. We designed such a system and evaluated its effectiveness. The system, known as Sask, utilizes the Socratic Method via human computer interaction techniques to make the experience feel like a conversation with one's educator. It also includes an authoring system designed to help educators quickly and easily build the tutors, called Mentors. In a quasi-experimental study the Sask system had a statistically significant effect size of 0.69 on students' post-test scores. This implies that the unique interactive design of Sask Mentors can make such Socratic software effective tools for improving a learning environment where human educators already engage students in inquiry dialog.

Introduction

In this paper we address the question of whether adding a computerized Socratic tutor to a learning environment improves learning environments where Socratic tutoring already occurs yet sub-optimally because there are too few educators per student. We hypothesize that the additional computer tutor helps increase students' conversational access to partially Socratic human and software tutors, and this will in turn foster more reflection and improve student outcomes. We are particularly concerned about students' ability to transfer what they have learned to analogous situations.

We have designed an inexpensive yet innovative system called Sask 0 for creating Web-based Socratic software tutors called Mentors. With the Sask Web-based authoring tool, called the Designer, educators who have personally taught using the Socratic Method before can design Mentors for specific lessons for their learners. Mentors are meant to be built by educators in a short amount of time and without extensive assistance from knowledge and software engineers. We used Sask in a learning environment and investigated whether it helped foster learners' reflections on a well-defined learning task.

Theoretical Framework

A common tool in education is to assign learners complex tasks, the doing of which leads to learning. Unfortunately, learners often do not have enough experience or knowledge to perform the learning task proficiently. They often see their goal as completing the learning task as quickly as possible. They rarely stop to examine ways in which the task may be challenging hidden misconceptions. These factors reduce the learning value of the task. An educator may know of these issues and choose to tutor their learners Socratically; to conversationally engage with learners, often while they work on their learning task, with pertinent and probing questions. By doing this Socratic tutoring, the educator helps the learners see pitfalls in the task and recognize their own misconceptions (Collins & Stevens, 1982; Collins & Stevens, 1983; Collins & Stevens, 1991; Chang, et. al., 2000). Thus we think a Socratic tutor can support and foster learners' task-situated reflection 0 by getting them to question what they are doing, how, and why (see Figure 1).

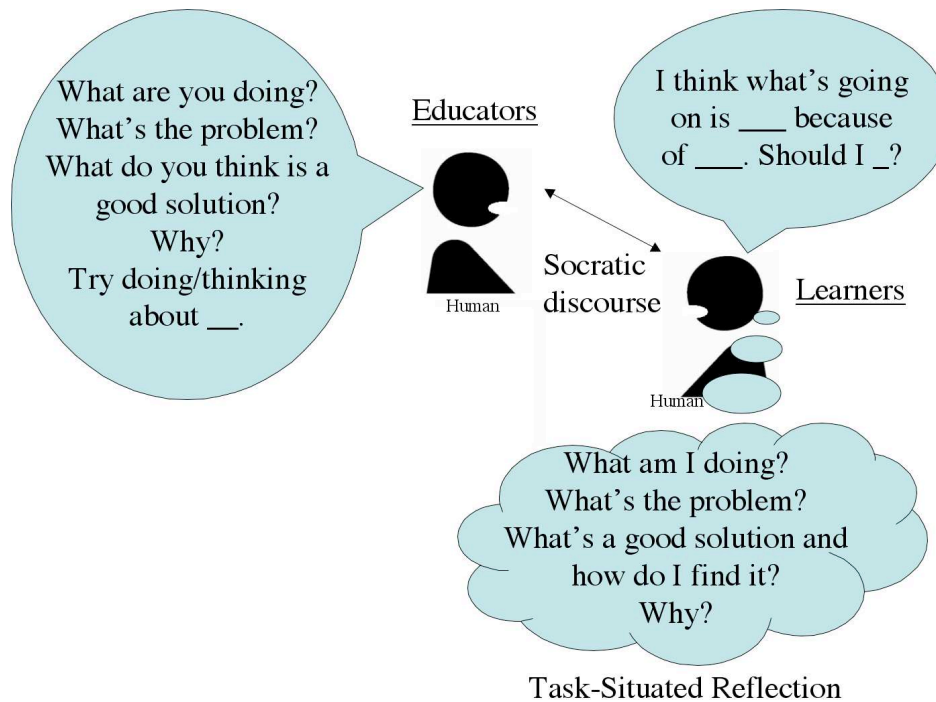


Figure 1. : Human Socratic tutor supporting learner's task-situated reflection on a learning task.

When there are many learners and too few educators, the educators are often not able to properly foster task-situated reflection in learners. One solution we propose is to provide artificial Socratic tutors, such as Sask Mentors, that provide a similar Socratic dialog while learners are working on their learning task. The human educator will likely also interact with the learner, but for the most part the software is supporting their task-situated reflection (see Figure 2).

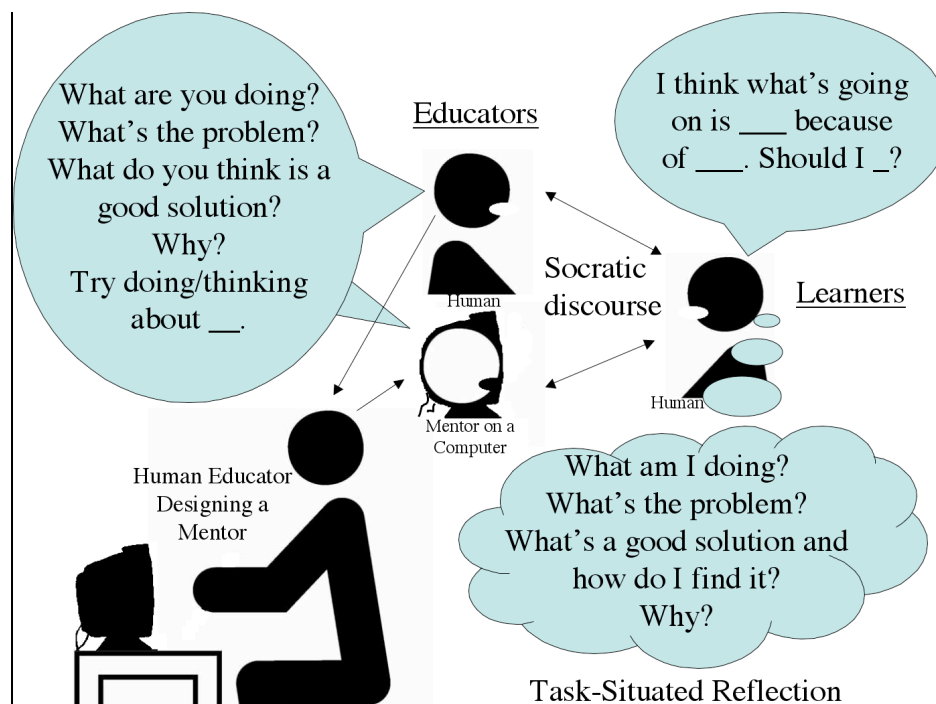


Figure 2: Software Socratic tutor supporting learner's task-situated reflection on a learning task with optional help from busy educator.

By adding a Sask system into a learning environment the amount of interaction between learners and Socratic educators increases, and consequently learning outcomes improve on both a short and long-term basis (compared to a control learning environment). Expected improvements are more accurate long-term memory and understanding, and better transfer. These improvements will be due to learners being situated in dialog that is more Socratic, and thus having more task-situated reflection and hypothesis construction experiences which foster better learning behaviors such as increased short-term motivation, better use of inquiry practices, and more thoughtful group interactions.

Unique Features For Fostering Reflection

When using the Mentor, learners can be engaged in a dialog with the program and thus get assistance in a similar manner as they would from their human educators. Using video or audio clips of the human educators (3) can increase the learners' feeling that they are having a conversation with their human educator and thus help transfer (Moreno, Mayer, Spires, & Lester, 2001; Reeves & Nass, 1996).

To meet our design goals, we knew the Mentor must allow learners to express themselves in a manner as close to human educator and human learner conversational interactions as possible. However, if we allowed learners to type in full sentences in response to the Mentor's utterances, we would have been faced with the unsolved Natural Language Processing problem of getting the system to understand the text without having all the experiential and cultural background a human mentor brings to bear. It would be frustrating for learners to type long texts that are then rejected for being improperly formatted, or worse misinterpreted thus making the tutoring session incongruent to their needs. A similar Socratic tutoring computer system, coincidentally called Mentor, solved this problem by giving students a menu (called the "Vocabulary") of all the utterances they could say that the system would recognize (Feurzeig, Munter, Swets, & Breen, 1964). We chose not to use this technique to avoid the possibility of students reflecting less because they know the best student response must be among the list given to them. We want the conversational interaction to encourage students to be more reflective. It should not be a multiple choice test where the student just needs to eliminate the "wrong" responses. To do that we used a technique used in the Creanimate system. Creanimate was designed to teach elementary school children about animal adaptation using Socratic techniques (Edelson, 1993). Creanimate encouraged students to make and explore hypotheses using a series of thought-provoking questions. When Creanimate asked a student a question, it provided templates for the student to use in their response. Our system does the same. When a Mentor presents its utterance it provides students response choices in the form of links (see Figure 3).

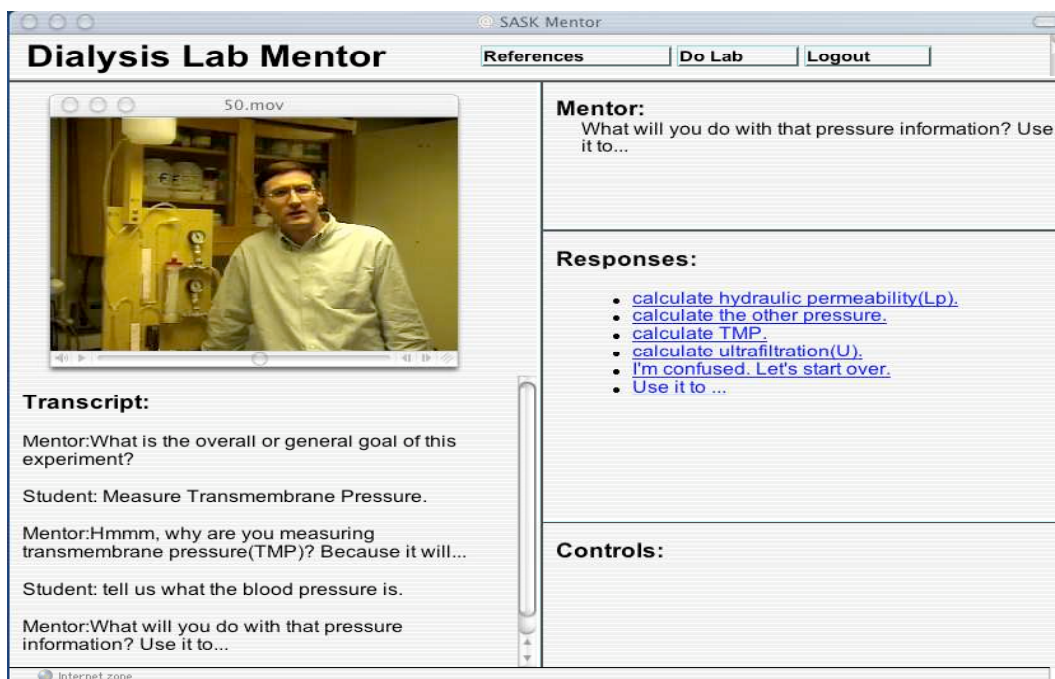


Figure 3: Before student clicks on "Use it to ...".

When students click on a link whose label text does not end with an ellipsis, the text of that link becomes their chosen response utterance. However, if the students click on a link whose label text ends with an ellipsis, such as in the "Use it to ..." link in Figure 3 above, we present them with a form, similar to those commonly used on a Web page, to provide their answer (see Figure 4). This avoids the problem of trying to understand open-ended text. It gives students some, but not too much, information regarding what type of response they can give. Educators then specify what they think their learners may type in the field and what the Mentor should say next in each case.

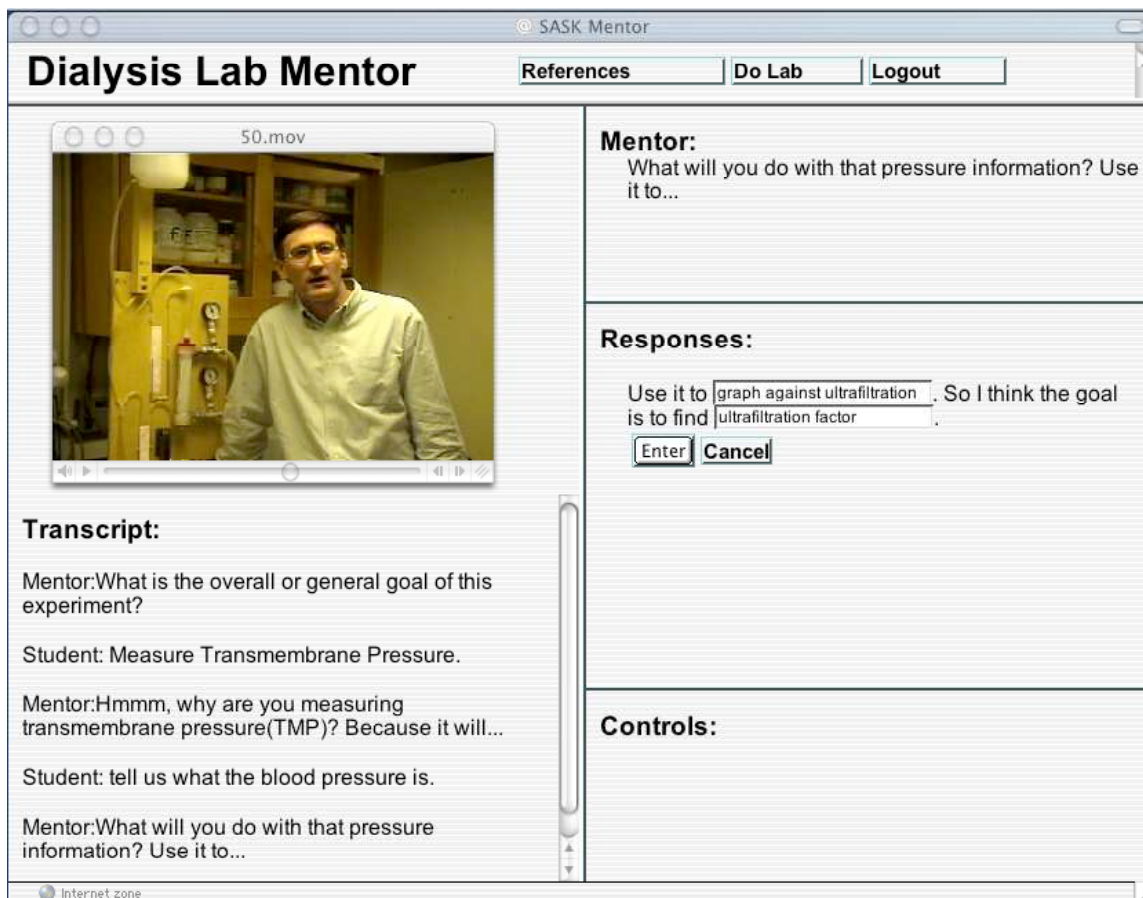


Figure 4: After student clicks on "Use it to ..." and starts filling resulting form.

With this system we expect educators to be able to write Socratic dialog tailored to a specific learning task that assists students in solving a problem or exploring a concept. Students can use such Mentors interactively during an inquiry process, receiving feedback similar to what they would get from their human educators, rather than passively listen to an explanation or a demonstration of logic. It should be noted that every aspect of the inquiry process of the learning task does not need to be handled by the Mentor, nor should it be. Mentors are meant to augment a learning environment where Socratic dialog among human educators and learners already occurs, not to completely replace the human educators. Mentors should be used for the most typical Socratic dialog students tend to need to be engaged in, while the human educator should be available in some manner for atypical dialog and clarification.

Results

Subjects in the current study were 64 seniors in a Fall 2002 biomedical engineering laboratory course at a research university. A Socratic Mentor, we named the Dialysis Lab Mentor, was designed to assist students during certain stages of a lab on determining the quality of a dialyzer (an artificial kidney). The Dialysis Mentor helped them understand the purpose of their lab experiment, as well as what variables need to be held constant and which need to be changed.

In this study we are using a quasi-experimental approach. All students shared the same lecture section, while half were also taking a morning lab session and the other half took a lab session that same afternoon. The afternoon session was made the experimental section by providing them with the Dialysis Mentor during their dialysis lab. The morning section was the control section and did not use the Mentor software. Both sections had the same professor and the same two teaching assistants available to them under usual conditions. Each section of about 30-33 students was broken into 10-11 work groups of about 3 students. Each workgroup worked on a lab apparatus along with a computer that was used to record and graph data and to possibly run the Dialysis Mentor. After initial lab setup and safety instructions, every group in the experimental section was directed by their human educators to use the Dialysis Mentor. The Mentor started the conversation by asking the group “What is the overall or general goal of this experiment?”. The students had been given a pre-test immediately prior to their first lecture related to the dialysis lab. Students were also given a post-test included in their final exam that was essentially the same question as the pre-test. The tests were designed to measure the correct remembrance and utilization of key concepts for the dialysis lab and the lack of common misconceptions for that lab.

The difference between the control section’s mean post-test scores and the experimental section’s mean post-test scores (middle bars in the Figure 5 below) is statistically significant. Furthermore, the mean gains on the post-test scores over the pre-test scores (rightmost bars in Figure 5 below) are also statistically significant. Both sections scored similarly on the pre-test with a mean of about 3.4 out of 15 possible points, implying that the two sections started with students of similar levels of expertise in the domain of the learning task. The experimental section’s mean post-test score (taken as part of courses final, almost 2 months after using the Dialysis Mentor) was about 8.72, about 2.24 points higher than the control section. That post-test mean score in the experimental group was also over 5 points higher than their mean pre-test scores compared to only about a 3 points gain in the control section. This is even more significant when you consider that the overall final exam grade is curved so that median students get a B grade, and thus a consistent score of 10/15 (two-thirds) on each final question would get a student an A grade on the final. In addition we found that 6 of the 25 reporting control section students did worse on their post-test than they did on their pre-test, while only 1 of the 29 reporting experimental section students did worse.

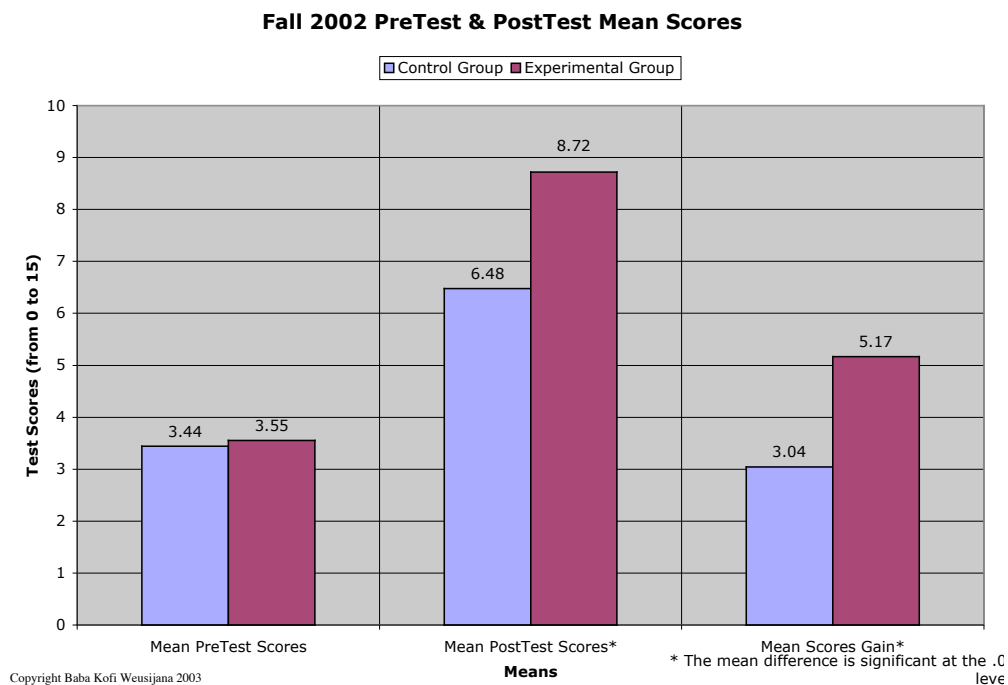


Figure 5: Analysis of outcomes data.

Tables 1 and 2 below display more statistical information. Of particular importance are the effect sizes (corrected for bias from sampling error) in Table 2. The use of the 2002 Sask Dialysis Mentor had an effect size of

.69 on students' scores on the post-test. That means almost 76% of the control group scored below the average person in the experimental group. There is also an effect size of .54 on students' gain in scores on the post-test versus their pre-test scores. That means about 71% of the control group made smaller gains than the average person in the experimental group. The .69 post-test effect size is one of the highest found for computer-assisted instruction (CAI). Reported effect sizes from different meta-analyses of CAI and other computer interventions in education generally fall in the range of .2 to .4. The few similar effect sizes are .69 regarding the use of computer simulations in a study in 2000, .56 regarding the use of Logo in a 1994 study, and .71 regarding computer usage in the classroom as a supplement to elementary math instruction in a 1982 study (Smith, 2001).

Table 1: Descriptive Statistics

Outcome Measurements	Experimental Group			Control Group		
	mean	n	standard deviation	mean	n	standard deviation
Pre Test	3.55	29	2.60	3.44	25	1.92
Post Test	8.72	29	3.12	6.48	25	3.31
Post-Pre Test	5.17	29	3.51	3.04	25	4.27

Table 2: Raw Differences & Standardized Effect Sizes

Outcome Measurements	Pooled Standard Deviation	p-value for difference in SDs	Mean Difference	p-value for mean diff (2-tailed T-test)	Effect size with bias corrected for sampling error (Hedges & Olkin, 1985)
Pre Test	2.31	0.06	0.11	0.860	0.05
Post Test	3.21	0.38	2.24	0.013*	0.69
Post-Pre Test	3.88	0.16	2.13	0.049*	0.54

* Statistically significant at the .05 alpha level.

Future Work

We are currently performing qualitative analyses of video, audio, computer interaction transcripts and survey data collected from the dialysis laboratory. We are attempting to determine what Socratic teaching strategies fostered reflective discourse, helped students the most in terms of achievement outcomes, and which strategies were used by the Sask Dialysis Mentor, the course professor, and graduate assistants. We will also investigate ways to improve the Sask Designer authoring tool to assist educators who do not have much experience teaching Socratically.

Conclusions and Importance

We have shown that increasing the amount of Socratic tutors and access to an educator by using a Socratic, non-menu based, software tutor can improve learning environments. This also implies that the partial use of forms instead of solely relying on multiple-choice interfaces can foster more reflection in learners. Educators can make such Socratic software tutors for their learners and can expect significant effects on exam outcomes.

This work has important implications for the usefulness of simple Socratic software tutors and for design research of teaching strategies. It provides a basic theoretical framework for fostering task-situated reflection through Socratic dialogue as well as software that can be used to refine the framework.

Endnotes

- (1) Sask is short for Socratic ASK. ASK systems are a form of hypermedia based on the metaphor of having a conversation with an expert, or a group of experts (Cleary & Schank, 1995). Currently Sask is a Tomcat webapp using Java Servlet and JSP technologies.
- (2) Task-situated reflection is domain-relevant reflection in the context of the learning task (Radinsky, 2000).

- (3) Any other Web-based content can be presented along with the Mentor's utterances. Also available is a visible transcript of what a Mentor and the learners have said that can be saved by learners to a text file for reference.

References

- Cleary, C., & Schank, R. C. (1995). *Engines for Education*. Hillsdale, NJ: Lawrence Erlbaum.
- Collins, A., & Stevens, A. L. (1982). Goals and Strategies of Inquiry Teachers. In R. Glaser (Ed.), *Advances in Instructional Psychology* (pp. 65-119). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Collins, A., & Stevens, A. L. (1983). A Cognitive Theory of Inquiry Teaching. In C. M. Reigeluth (Ed.), *Instructional-design theories and models* (pp. 203-230). Hillsdale, N.J: Lawrence Erlbaum Associates.
- Collins, A., & Stevens, A. L. (1991). A Cognitive Theory of Inquiry Teaching. In P. Goodyear (Ed.), *Teaching Knowledge and Intelligent Tutoring*. Norwood, NJ: Alex Publishing Corporation.
- Chang, K., Lin, P., Sung, Y., & Chen, S. (2000). Socratic-Dialectic Learning System of Recursion Programming. *Journal of educational computing research*, 23(2), 133.
- Edelson, D. C. (1993). *Learning from stories: indexing and reminding in a Socratic case-based teaching system for elementary school biology*. Doctoral dissertation, Computer Science, Northwestern University.
- Feurzeig, W., Munter, P., Swets, J., & Breen, M. (1964). Computer-Aided Teaching in Medical Diagnosis. *Journal of Medical Education*, 39(8), 645-754.
- Hedges, L. & Olkin, I. (1985) *Statistical Methods for Meta-Analysis*. New York: Academic Press.
- Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The Case for Social Agency in Computer-Based Teaching: Do Students Learn More Deeply When They Interact with Animated Pedagogical Agents? *Cognition and Instruction*, 19(2), 177-213.
- Radinsky, J. (2000). *Making sense of complex data : a framework for studying students' development of reflective inquiry dispositions*. (pp. 38-47). Doctoral dissertation, Education and Social Policy-Learning Sciences, Northwestern University.
- Reeves, B., & Nass, C. (1996). *The Media Equation*. New York, NY: Cambridge University Press.
- Smith, J. E. (2001). *The effect of the Carnegie algebra tutor on student achievement and attitude in introductory high school algebra*. Retrieved 20010523<http://scholar.lib.vt.edu/theses/available/etd-04232001-161143>; <http://scholar.lib.vt.edu/theses/available/etd-04232001-161143>.

Acknowledgments

We thank Dr. Allan Collins, Dr. Daniel Edelson, and Dr. Brian Reiser for their many suggestions and groundwork for this study. We also thank Dr. Ann McKenna and Dr. Jelani Mandara for their assistance with the project. This work was supported in part by the Engineering Research Centers Program of the National Science Foundation under Award Number EEC-9876363. EEC-9876363 funds the VaNTH engineering research center composed of the bioengineering and learning sciences faculties of Vanderbilt University, Northwestern University, the University of Texas at Austin and the Health Science and Technology Program of Harvard and MIT. See <http://www.vanth.org> for more information.