Teaching with Dialectic Arguments vs. Didactic Explanations

Ravi Desai (rpdst6@pitt.edu)
Intelligent Systems Program, University of Pittsburgh,
LRDC, 3939 O'Hara Street
Pittsburgh, PA 15260 USA
Kevin D. Ashley (ashley@pitt.edu)
Learning Research and Development Center, University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260 USA

Abstract

We compared two automated approaches to teaching *distinguishing*, a skill that involves assessing the relevant differences among cases in a context-sensitive way. We implemented the two approaches in two versions of an Intelligent Tutoring System designed to teach law students basic skills of case-based legal argument. The original version of CATO employs didactic explanation. The newer version, CATO-Dial, teaches the same skill with a simulated dialectic argument in a courtroom setting. We hypothesized that students would learn the skill better by engaging in the simulated argument than by receiving interactive explanation. We expected that dialectic argument would help students to construct the target knowledge on their own as they developed responses to arguments, and that this would lead to more robust learning. We showed that students in the dialectic argument simulation group performed significantly better on a section of the post-test aimed at assessing transfer of their skills of distinguishing. We discuss a number of cognitive and motivational benefits that may explain this effect

Introduction

The skill of distinguishing is important in dialectical domains such as law, applied ethics, policy analysis, and business, where arguments by analogy are routinely employed in professional education and practice. Distinguishing is a way to respond to an analogical argument. The argument claims that a target problem should be decided in the same way as a cited source case by virtue of their relevant similarities, that is, factual patterns the cases share that form the basis of legal reasons for deciding them in the same way. A *distinction* is a factual difference underlying a legal reason to decide the target problem differently from the cited case. There may be many differences, but only those that give rise to legal reasons for treating the cases differently are distinctions.

A distinction may be a factual strength of a side (i.e., a party, either plaintiff, the one who commences suit, or the defendant) in the target problem not shared in the source case, or a factual weakness in the source case not shared in the target. For the cases in our experiment, students play the role of the defendant's attorney; a distinction is defined as a pattern of facts that strengthens defendant's legal argument in the problem not found in the cited-case, or a fact pattern that weakens defendant's side in the cited-case not found in the problem.

In order to distinguish effectively, one must be sensitive to the argument context in which a case has been used. The context includes the role a relevant difference plays in an argument, its underlying legal significance both in absolute terms and relative to the other combinations of facts in the target problem and cited precedent.

Professors note that law students often demonstrate only a shallow understanding of the concept of a distinction. Students may be able to find different facts but fail to realize that only some differences are distinctions. Students may also ignore which side a difference favors, or they fail to view the significance of a difference in the context of the other facts in the problem and cited-case. Because of their shallow knowledge, students may make arguments citing differences that hurt, rather than help, their side's argument.

Ideally, students "pick up" the skill of distinguishing through the trial-and-error experience of making arguments. In a law school, students engage in arguments in the classroom. Sometimes, however, students are reluctant to expose themselves in class by making arguments, and, in any event, could benefit from additional instruction outside the classroom. In an effort to meet this need, Aleven and Ashley (1997) developed the CATO program (i.e., Case Argument TutOrial), an Intelligent Tutoring System designed to teach basic case-based argument skills, including how to distinguish cases (see also Aleven, 1997).

In designing an ITS to teach distinguishing, one may try two approaches. *Didactic explanation* involves presenting good and bad examples of distinguishing. The system explains

why the examples are instances of successful or unsuccessful argument. The bad examples illustrate the various kinds of shallow knowledge. This is how CATO teaches distinguishing.

Dialectic argument attempts to teach students distinguishing by engaging them in arguments, affording an opportunity to learn the skill through a process of trial-and-error. Technically, it is harder to design this kind of pedagogical interaction. Before undertaking to develop a large-scale dialectic argument system, we wanted to see whether it was likely to improve students' learning. We therefore developed a variation of CATO, called CATO-Dial, which employs dialectic argument to teach distinguishing.

We hypothesized that students would learn better to distinguish by engaging in dialectical argument than in didactic explanation. We speculated that students engaged in role-playing and arguing would encounter information in a more relevant context and would be motivated to process the information more thoroughly. They would develop deeper knowledge of the role a difference plays in the argument context, its interactions with other facts in the target problem and source case and the underlying reasons why it matters. Our experiment tested this hypothesis.

In aiding students to learn deeper knowledge, we also hoped that CATO-Dial would promote better transfer of knowledge. A major problem in knowledge transfer is that people tend to access prior knowledge that bears superficial rather than structural similarity to the problem at hand (Thompson et al., 2000).

Legal argument is not as determinate a form of problem-solving as, say, physics or geometry. Legal problems rarely have provably correct answers; there may be reasonable arguments on both sides of a dispute based on analogies to competing cases (Ashley, 1990). CATO-Dial attempts to engage students in argument dialogues that focus them on comparing cases.

Law school professors engage students in Socratic dialogues about cases in the casebooks. Some of the earliest ITSs used Socratic dialogue and an inquiry teaching method to teach subject matter such as geography (SCHOLAR (Carbonell, 1970)) or meteorology (WHY (Collins and Stevens, 1982)). A subsequent tutoring system (Wong et.al, 1997) incorporated the inquiry teaching method into an ITS shell and geography tutor. The OLIA ITS (Retalis et al., 1996) used a related argument dialogue strategy, playing devil's advocate. Research suggests that students tutored manually with Socratic dialogues learned targeted physics concepts (i.e., rules) better than those taught with more didactic dialogues (Rose et. al, 2001). In the latter, the human tutor provided more explanation before asking questions but asked fewer open-ended questions.

CATO vs. CATO-Dial

CATO is one of the few case-based Intelligent Tutoring Systems that teaches a process of case-based reasoning. (Aleven, 1997, pp. 197-8). It provides a set of specialized tools, accessible through an X-server connection to CATO running on a Unix workstation, and a web-accessible Casebook and Workbook. The Casebook presents excerpts from important legal case opinions in trade secret law. A small set of argumentation and discussion questions follows each case, much like an ordinary legal casebook. The Workbook helps students use CATO's tools to analyze and respond to the argumentation and discussion questions.

Experiments show that CATO is an effective teacher (Aleven & Ashley, 1997; Aleven, 1997). Students work with CATO's textual case summaries and abstract representations of cases in terms of *factors*. Each factor represents a stereotypical collection of facts, which tends normally to strengthen or weaken a conclusion that a side should win a particular kind of legal claim (Ashley, 1990). A Factor Hierarchy represents legal reasons why a factor makes a difference to the legal claim (Aleven, 1997).

CATO helps students analyze target problems and compare them to past cases. It teaches novices to identify factors in a target problem, test hypotheses about their significance against cases in its database, and make legal arguments about how to decide the target problems citing cases. Students encounter problems based on real litigated cases. Novice users identify conflicting factors in the problem, which give rise to conflicting reasons for decision. CATO teaches them how to make arguments to resolve such conflicts.

CATO's Argument Maker tool provides a tutorial on distinguishing. To enable CATO to employ dialectic argument, we developed CATO-Dial, a modification of the program that engages novice users in courtroom-style arguments about target problems. In using the CATO-Dial version of the tutorial on distinguishing, students encounter examples like the *Lynchburg Lemonade* case.[1] With the Case Analyzer tool, students have

[1] In the *Lynchburg Lemonade* case, Tony Mason, the plaintiff, developed a cocktail he dubbed "Lynchburg

identified conflicting factors in the *Lemonade* problem and have begun to consider the conflicting factor-based legal reasons about how to decide its outcome. For comparing cases, the Case Analyzer presents them in a tabular form as in Figure 1. The problem has six factors, four of which favor the plaintiff (p) and two of which favor the defendant (d). The *Boeing* case, won by plaintiff (p), shares two of these factors, the relevant similarities (marked with "="). The relevant differences (i.e., distinctions) are the four unshared factors marked with "*". These favor deciding the *Lemonade* case for the defendant (i.e., differently from *Boeing*). Note that F16 strengthens the defendant in the *Lemonade* case and is not found in *Boeing*, whereas F4, F12 and F14 strengthen the plaintiff's position in *Boeing* and are not found in the *Lemonade* case. F10, F15 and F18 are also unshared factors, but they are not distinctions because they favor deciding the *Lemonade* case for plaintiff (i.e., the same as in *Boeing*).

With CATO-Dial, students play the role of an advocate, Perry Mason, who has to argue a case in court. As shown in Figure 2, the student may put arguments in the mouth of Perry Mason by selecting argument moves and values from a menu. CATO-Dial responds on behalf of the Judge, who mediates the proceedings, Hamilton Burger, Perry's opposing counsel, and Della Street, Perry's savvy assistant, who offers helpful hints.

In the dialogue, Mr. Burger's responses (such as step 5 in Figure 2), generated by CATO-Dial, take advantage of any weaknesses in Mr. Mason's argument, based on the students' menu selections. The Judge's reaction emphasizes the student's mistake, and Della's hints, also generated by CATO-Dial, provide instruction on how to correct them. CATO-Dial can engage in dialogues like this for any pair of relevant cases in its database.

Lemonade". Since Tony took some measures to protect his recipe's secrecy, and since his was the only tavern producing this drink, we say factors F6, Security-Measures, and F15, Unique-Product, apply; both tend to favor the plaintiff (p). On the other hand, Tony disclosed his recipe in negotiations with a sales agent of the defendant, Jack Daniel's Distillery. Thus, F1, Disclosure-In-Negotiations, applies, a factor that tends to favor the defendant (d). The agent was aware, however, that the recipe was a "secret formula", so F21, Knew-Info-Confidential, also applies, tending to favor the plaintiff. Without compensating Tony, the defendant started marketing exactly the same kind of cocktail; F18, Identical-Products applies and favors plaintiff. Finally, the recipe could have been obtained by reverse engineering the cocktail; F16, Info-Reverse-Engineerable, applies and favors the defendant.

CATO-Dial's dialectic argument contrasts with CATO's didactic explanation. Originally, CATO's tutorial engaged students in an interactive exercise employing didactic explanation. In these explanations, CATO did not engage the student in as direct a manner. Instead, the student was presented with an argument that a hypothetical opponent could make. Below is an excerpt from CATO's original version of the dialogue of Figure 2. First, CATO makes the same argument as in Step 1 of Figure 2 and invites the student to select the factors that he would emphasize in distinguishing the *Boeing* case on behalf of the defendant. Second, the student again selects three differences (i.e., factors F15, 516, and F18) which though not shared between the two cases, actually strengthen the argument that plaintiff should win in the problem situation. CATO then responds as follows:

3. CATO: Suppose defendant made the following argument: The *Boeing* case is distinguishable, because in the *Lynchburg Lemonade* case, defendant's product was identical to plaintiff's [F18] and plaintiff was the only manufacturer making the product [F15]. This was not so in *Boeing*. Would this really help defendant? When you have answered this question for yourself, please click on "Continue".

4. STUDENTS: (Continue)

5. CATO: Plaintiff could respond to defendant's argument, as follows: Your Honor, as defendant's counsel correctly pointed out, in the current problem, defendant's product was identical to plaintiff's [F18] and plaintiff was the only manufacturer making the product [F15]. *Boeing*, which was won by plaintiff, does not have these strengths on plaintiff's part but is otherwise similar to the problem at hand. Where these additional strengths are present, a decision in favor of plaintiff is warranted even more strongly than it was in *Boeing*. In short, even though factors F18 and F15 are differences, defendant cannot use them to distinguish *Boeing*. Defendant needs to focus on other differences, namely, on those differences that make the *Lynchburg Lemonade* case look stronger for defendant than *Boeing*. Please click on "Continue". *See* (Aleven, 1997 at 141f).

The argument that CATO uses to show the inadequacy of the student's initial selection of distinctions is the same in both versions; the difference lies in the manner in which it is presented to the students and the way in which students are engaged in the task

Description of Experiment

We compared the impact of teaching distinguishing to senior undergraduates using the two versions of CATO. The students had all been accepted into law schools and were receiving

preparatory instruction through a summer institute. The students were randomly assigned to two groups. The experimental group used the dialectic argument version of CATO-Dial and initially numbered 22 students. The control group worked with the didactic explanation version of CATO and initially numbered 23 students. Each group worked in a series of eight two-hour sessions over a span of about one month from June 5 through July 11, 2000. For each session a student was paired with a different partner from the same group.

Prior to the series of instructional sessions, all students took a pre-test comprising three questions designed to assess argumentation skills. For Questions 1 and 2, students read a problem situation and three short cases. Students were asked to make and respond to arguments about the problem given the cases. Question 3 asked them to define the concepts of a relevant similarity and relevant difference.

Lynchburg Lemonade Case	*Boeing* (p)
= Fl Disclosure-In-Negotiations (d)	= Fl Disclosure-In-Negotiations (d)
= F6 Security-Measures (p)	* F4 Agreed-Not-To-Disclose (p)
F15 Unique-Product (p)	= F6 Security-Measures (p)
* F16 Info-Reverse-Engineerable (d)	F10 Secrets-Disclosed-Outsiders (d)
F18 Identical-Products (p)	* F12 Outsider-Disclosures-Restricted (p)
= F21 Knew-Info-Confidential (p)	* F14 Restricted-Materials-Used (p)
	= F21 Knew-Info-Confidential (p)
= shared factor * distinction	

Figure 1: Case Comparison In Terms of Factors

Court is in session…

1. Mr. Burger for Plaintiff (CATO): Your Honor, where plaintiff adopted security measures [F6] and defendant knew that plaintiff's information was confidential [F21], even though plaintiff disclosed its product information in negotiations with defendant [Fl], plaintiff should win a claim of trade secrets misappropriation, as in *The Boeing Company v. Sierracin Corporation*, 108 Wash.2d 38, 738 P.2d 665 (1987).

2. Judge (to Defendant's counsel): Your response, Mr. Mason?

3. <STUDENT (selecting from menu): Distinguish *Boeing* [F15] [F16] [F18] (Continue)>

Mr. Mason for Defendant (Student): *The Boeing Company v. Sierracin Corporation* is distinguishable, Your Honor, because in the *Lynchburg Lemonade* case, plaintiff's product information could be learned by reverse-engineering [F16]. This was not so in *Boeing*. *Boeing* is also distinguishable, because in the *Lynchburg Lemonade* case, unlike *Boeing,* defendant's product was identical to plaintiff's [F18] and plaintiff was the only manufacturer making the product [F15].

4. Judge (to Plaintiff's counsel): Mr. Burger?

5. Mr. Burger for Plaintiff (CATO): Your Honor, as Mr. Mason concedes, in the current problem, defendant's product was identical to plaintiff's [F18] and plaintiff was the only manufacturer making the product [F15]. *Boeing*, which was won by plaintiff, does not have these strengths on plaintiff's part but is otherwise similar to the problem at hand. Where these additional strengths are present, a decision in favor of plaintiff is warranted even more strongly than it was in *Boeing***.**

6. Judge (to Defendant's counsel): I am sorry, Mr. Mason, but I am afraid I have to agree with Mr. Burger.

7. STUDENT: Help! (Continue)>

8. Della Street (whispering into Perry's ear): Here is a hint. Even though factors F18 and F15 are differences, defendant cannot use them to distinguish *Boeing*. Defendant needs to focus on other differences, namely, on those differences that make the *Lynchburg Lemonade* case look stronger for defendant than *Boeing*.

9. <STUDENT (selecting from menu): Distinguish *Boeing* [F4][F12] [F14] [F16] (Continue)>

Mr. Mason for Defendant (Student): *Boeing* is distinguishable, because in *Boeing*, defendant used materials that were subject to confidentiality restrictions [F14], plaintiff's disclosures to outsiders were subject to confidentiality restrictions [F12], and defendant entered into a nondisclosure agreement with plaintiff [F4]. This was not so in the *Lynchburg Lemonade* case. Also, in the *Lynchburg Lemonade* case, plaintiff's product information could be learned by reverse-engineering [F16]. This was not so in *Boeing*.

10. Judge (to Plaintiff's counsel): Mr. Burger?

11. Mr. Burger for Plaintiff (CATO): Uhh. Well, as you know, Your Honor, *Boeing,* is not the only case we cite in our brief.

12. Della Street (whispering into Perry's ear): Brilliant job, Perry!

Figure 2: Sample Dialectic Argument in CATO-Dial

During the first six sessions, the experimental and control groups were treated identically. In Sessions 1 to 4, both groups of students familiarized themselves with the factor based representation system and the CATO environment. In Sessions 5 and 6 all students learned how to use the Argument Maker tools.

The control group and experimental group were treated differently in Sessions 7 and 8. These two sessions focused on teaching students how to distinguish cases. During both Sessions 7 and 8, students were expected to complete eight pair-wise case comparisons involving distinguishing. The eight pairs of cases were the same in both groups. The only difference was the manner in which CATO and CATO-Dial taught the lesson. The experimental group worked with CATO-Dial's simulated courtroom dialogues like that in Figure 2. The control group worked with the original CATO didactic explanations.

After Session 8, all students took a post-test comprising six questions. The first three questions were worded identically to the pre-test questions, but Questions 1 and 2 involved a different problem and cases. The other three questions tested the following transfer skills:

- Question 4 put students in a new role – instead of making arguments they critiqued an argument.
- Question 5 tested students' recall of a particular problem situation they had encountered in the instruction. This problem had been used extensively in the teaching sessions as a basis of the argumentation lessons. Students were asked to make and respond to an argument about the problem, which they had to recall from memory, by analogizing it to and distinguishing it from a new case presented with the question.
- Question 6 tested students' ability to apply skills they had learned to an unfamiliar new domain: the copyright law doctrine of Fair Use. As such, this question is a telling measure of their ability to distinguish cases.

The director of the University of Pittsburgh School of Law legal writing program graded all but one of the pre-test and post-test questions. The grader was provided a one-page summary of grading criteria and instructed to assign a gestalt grade (between 1 and 10) to each question. He was blind as to the identity of the test writers, but did know which were pre-tests and which were post-tests. The exception was Question 5, the recall question, for which we developed an objective grading scheme. Students were awarded a maximum of ten possible points on the basis of how many of the factors in the problem they referred to in their argument.

Analysis

Post-test data were available for only 22 of the 45 students, 15 in the experimental group and 7 in the control group. The rest of the subjects either dropped out before the end of the experiment or did not complete enough work in Sessions 7 and 8, the only sessions involving differential treatment of the two groups.

Pre-test scores were analyzed for the 22 students, who provided both pre-test and post-test data. For each student, responses to the three pre-test questions were summed, and the mean response of students in the experimental group was compared to that of students in the control group, using a two-tailed t-test. Results showed no statistically significant difference between the two groups. Since the students were paired with different partners across sessions, we used the individual student rather than the pair as the unit of analysis for both pre-test and post-test analyses.

Post-test scores were also analyzed for the 22 students who provided both pre-test and post-test data. A two-tailed t-test indicated no significant difference in the mean post-test scores of the experimental and control groups with respect to the first five questions. For question 6, however, the mean post-test score of the experimental group was significantly higher than that of the control group ($t(7.1) = 3.87$, $p < .05$, effect size of 1.57).[2] On question 5, the mean post-test score of the experimental group was nearly statistically significantly higher than that of the control group ($t(20) = 1.39$, $p = 0.052$, effect size of 0.62). While the difference was not significant the experimental group scored substantially higher on question 4 with an effect size of 0.88.

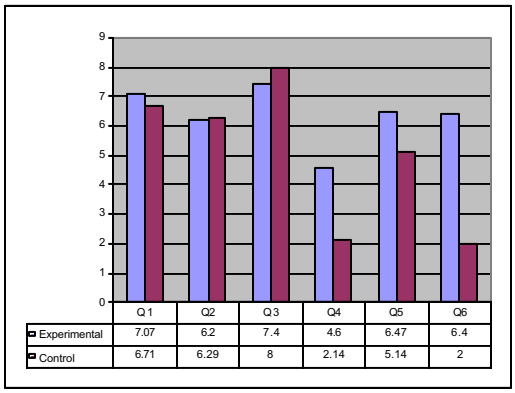

Figure 3: Comparison of Mean Post-Test Scores

Discussion and Conclusion

Our hypothesis was that students would learn the skill of distinguishing better by engaging in simulated dialectical argument than in interactive didactic explanation. The data confirmed our hypothesis in an interesting way. Whereas dialectic argument was not more effective than didactic explanation in teaching argumentation skills, it was more effective in helping students transfer the skills they learned to new tasks and significantly better in helping students apply the skills to an unfamiliar legal domain. Dialectical argument may have induced students to construct a schema for making and responding to arguments, resulting in deeper knowledge and thus better performance on transfer skills. Experimental students also appeared to have a better understanding of the importance that context plays in the task of finding distinctions. The grader also evaluated the

[2] Degrees of freedom for this test were reduced from 20 because Levene's test for equality of variances indicated that the variances of scores in the experimental and control groups were not equal.

answers in terms of four grading criteria, each involving a simple binary positive-or-negative scale. Two differences emerged. Students in the experimental group were more often rated positively on the criteria "Avoids making opponent's argument" and "Avoids errors regarding which side strengths favor". These results support the conclusion that the experimental manipulation helped students to learn better when a difference is a distinction.

It is intriguing that the rather superficial transformation from CATO to CATO-Dial in the presentation of the lesson on distinguishing had such a dramatic benefit. After all, both programs presented the same basic information. The critical difference, we believe, is that CATO-Dial's dialectical argument simulation provided that information in a more useful way. The dialectical argument offers several potential benefits, any or all of which may explain the observed difference in transfer scores.

Students may have found the increased level of involvement and the competitive element in the courtroom simulation motivating and thus conducive to paying attention and learning.

Furthermore students may have found it easier to understand the program's responses in CATO-Dial than in CATO. It is awkward for CATO to explain the quality of a student's response by example. The dialectical argument simulation, by contrast, provides a more natural context for illustrating the effect in an ongoing dialogue regarding a student's choices. Students in the experimental group did report finding the dialogues somewhat (though not significantly) more helpful than did those in the control group. When asked, "When CATO did provide instructional feedback, how helpful was it?", the CATO-Dial students rated it as more helpful than did the CATO students (Ms = 6.76 and 5.56 out of 10, respectively).

Role-playing in a courtroom argument, with its cognitive and emotional expectations, may also be important. Courtroom simulation explicitly prompts the student. An interactive style of human tutoring, in which tutors prompted students, supported learning even when tutors did not provide explanations and feedback (Chi et al., 2001). Having been prompted to participate, students may have been more likely to argue the merits to themselves, a task cognitively similar to self-explanation (Chi et al., 1989). Dialectical argument may also induce a student to feel worse about making a mistake than does didactic explanation. If so, students are more likely to pay attention and to care about learning in the former context. Role-playing may also induce students to compare the cases more carefully, helping the transition from shallow to deeper knowledge. In a recent investigation, business school students who compared cases in a study phase were three times more likely to transfer the cases' implicit principle to a new application than were those who simply read the cases for the purposes of advice-giving (Thompson et al. 2000).

The results suggest that the CATO-Dial approach is potentially quite valuable. Since laws change, law professors value students' ability to transfer skills. Our subsequent work will focus on converting as much of the CATO curriculum as possible to a dialectical format.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 9720359. We thank Professor Kevin Deasy, University of Pittsburgh School of Law, for his many contributions to this work.

References

Aleven, V. (1997) *Teaching Case-Based Argumentation Through a Model and Examples*, Ph.D. Diss., U. Pgh, LRDC.

Aleven, V. & Ashley, K.D. (1997) "Teaching Case-Based Argumentation Through a Model and Examples". In *Proc. 8th World Conf. AI in Education Society*. pp. 87-94. IOS Press: Amsterdam. Kobe, Japan. August.

Ashley, K.D., (1990). *Modeling Legal Argument: Reasoning with Cases and Hypotheticals*. The MIT Press: Cambridge.

Carbonell, J.R. (1970). "AI in CAI: An Artificial Intelligence approach to Computer Aided Instruction". *IEEE Transactions on Man Machine Systems* 11(4) 190-202.

Chi, Michelene T.H., S. Silver, H. Jeong, T. Yamauchi, & R. Hausmann (2001) "Learning From Human Tutoring" in *Cognitive Science*, Vol. 25, pp. 471-533.

Chi, Bassock, Lewis, Reimann & Glaser (1989), "Self Explanations: How students study and use examples in learning to solve problems", *Cognitive Science, 5,145-182*

Collins, A. & Stevens, A. L. (1982). "Goals and Strategies of inquiry Teachers". In *Advances In Instructional Psychology*, R Glaser (ed.) pp. 65-119. Hillsdale, NJ: Erlbaum.

Retalis, S., H. Pain & M. Haggith. (1996) "Arguing with the Devil; Teaching in Controversial Domains". In *Intelligent Tutoring Systems, 3d Int'l Conf., ITS '96*, 659-667. Berlin: Springer.

Rose C. P., J. D. Moore, K. VanLehn, D. Allbritton. (2001) "A Comparative Evaluation of Socratic versus Didactic Tutoring", 2001 LRDC Tech Report LRDC-BEE-1.

Thompson, L., Gentner, D. and Loewenstein, J. (2000) "Avoiding Missed Opportunities in Managerial Life." in *Org. Behavior and Human Decision Proc.*, 82, No. 1. May. pp. 60-75.