Teaching Case-Based Argumentation Concepts using Dialectic Arguments vs. Didactic Explanations*

Kevin D. Ashley, Ravi Desai, and John M. Levine

Learning Research and Development Center

University of Pittsburgh
Pittsburgh, Pennsylvania, USA 15260
ashley@pitt.edu

Abstract. We compared two automated approaches to teaching *distinguishing*, a fundamental skill of case-based reasoning that involves assessing the relevant differences among cases in a context-sensitive way. The approaches are implemented in two versions of CATO, an ITS designed to teach law students basic skills of case-based legal argument. The original version of CATO employed a didactic explanatory dialogue. The newer version, CATO-Dial, teaches the same skill with a simulated dialectic argument in a courtroom setting. Our hypothesis was that students would learn better by engaging in the simulated argument than by receiving interactive explanation. We showed that students in the dialectic argument simulation group performed significantly better on certain sections of the post-test aimed at assessing transfer of their skills of distinguishing.

1. Introduction

In dialectical domains such as law, applied ethics, policy analysis, and business, where arguments by analogy are routinely employed in professional education and practice, the skill of distinguishing cases is fundamental. In law, practitioners regularly make arguments by drawing analogies between a target problem and past precedents. The argument asserts that the target should be decided in the same way as a cited source case by virtue of their relevant similarities. These are factual patterns the cases share that form the basis of legal reasons for deciding them in the same way.

Distinguishing is a way of responding to such arguments. A *distinction* is a factual difference underlying a legal reason to decide the target problem differently from the cited case. Not all differences are distinctions, only those that give rise to legal reasons for treating the cases differently. Such a difference 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 used in our experiment, in which a student plays 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

* 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.

problem. The legal reasons triggered by these factual strengths and weaknesses are based in legal policies and values in statutes, constitutions, and precedents [5].

One who distinguishes successfully must be sensitive to the argument context in which a source case has been used. The role a difference plays in an argument -- its underlying (here legal) significance depends on the other combinations of facts in the particular target problem and cited source case under comparison.

In observing student arguments, instructors note that students often evidence only shallow knowledge of the concept of a distinction. They may be able to find differences between cases but do not necessarily understand that only some differences are distinctions. Students may tend to ignore which side a difference favors, or they may view the significance of a difference independently of the other facts in the problem and cited-case. Because of their shallow knowledge, students may even make arguments which hurt, rather than help, their side's position.

Like other argument skills, the skill of distinguishing is ideally "picked up" by students through trial-and-error practice. In law school, students learn argumentation skills by engaging in classroom dialectical exchanges with a professor and other students and by participating in mock courtroom arguments in moot court competitions and legal writing classes. Students, however, are often 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 [1, 2] developed the CATO program (i.e., Case Argument TutOrial), an Intelligent Tutoring System designed to teach beginning law students basic skills of making arguments with cases, including how to distinguish cases.

How can a tutoring system best teach students to learn to distinguish well? One approach, *didactic explanation*, involves presenting good and bad examples of distinguishing. The bad examples illustrate the various kinds of shallow knowledge, and the system explains why the examples are instances of unsuccessful argument. This is how CATO teaches distinguishing.

Alternatively, a *dialectic argument* approach might attempt to teach students distinguishing by engaging them in arguments and giving them an opportunity to learn the skill through a process of trial-and-error. From a technological viewpoint, this is a more difficult kind of pedagogical interaction to engineer. Before we undertook the effort to develop a large-scale didactic argument system, we wanted to see whether it was likely to make a difference in how well students learned. We therefore developed a variation of CATO, called CATO-Dial, which employs dialectic argument to teach distinguishing.

Our hypothesis was that students would learn the skill of distinguishing better by engaging in simulated argument than in interactive explanation. We speculated that role-playing and arguing would present information to students in a more relevant context and motivate them to process information more thoroughly. This would expose students' superficial knowledge and help them develop deeper knowledge concerning how a difference relates to underlying reasons, to the role it plays in the argument context, and to interactions with other facts in the problem and cited-case. We conducted an experiment to test our hypothesis that CATO-Dial can give students this kind of experience and benefit.

CATO-Dial's tutoring methodology is to engage students in argument dialogues that focus on case comparison. Traditionally, law school professors engage students in Socratic dialogues about cases that students read in their casebooks. While a number of Intelligent Tutoring Systems have been developed to teach legal subject matter [7, 13], few have focused on teaching legal argumentation skills or reasoning with cases. Legal argument is not as determinate a form of problem-solving as, say, physics or

geometry. Legal problems rarely have provably correct answers; instead, there may be reasonable arguments on both sides of the dispute based on analogies to competing cases [3, 11]. The legal domain is also highly focused on texts, requiring any tutoring system to finesse the need for complex natural language processing.

Some early tutoring systems used argumentation to teach subject matter. For example, SCHOLAR [6] taught geography with a kind of Socratic dialogue. Collins and Stevens [9] concluded after systematic study that the best human tutors used a Socratic style. Their WHY system taught students about the causes of rainfall using an inquiry teaching method and Socratic dialogue. A subsequent tutoring system [15] incorporated the inquiry teaching method into an ITS shell and geography tutor. Using another argumentation dialogue strategy, playing devil's advocate, the OLIA ITS [10] defends the weaker position as it engages students in an argument / counter-argument dialogue. More recently, research suggests that students tutored manually with Socratic dialogues learned targeted physics concepts (i.e., rules) better than those taught with more didactic dialogues [12]. In the latter, the human tutor provided more explanation before asking questions but asked fewer open-ended questions.

2. Didactic Explanation vs. Dialectic Argument

A variety of tutoring systems have incorporated case-based reasoning to tell relevant stories, provide worked-out examples, adapt as students solve new problems, generate new problems, model students' knowledge, and provide advice to teachers. CATO, however, is one of the few that teaches a process of case-based reasoning, literally comparing and contrasting problems to past cases in order to draw and justify inferences about the problems [1, pp. 197-8].

CATO's instructional environment comprises a web-accessible Casebook and Workbook and a set of specialized computerized tools, accessible through an X-server connection to CATO running on a Unix workstation. The Casebook presents excerpts from important legal case opinions in trade secret law. Each is followed by a small set of argumentation and discussion questions, much as in an ordinary legal casebook. The Workbook guides students in using the CATO tools to analyze and respond to the argumentation and discussion questions.

In using CATO, students work with 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 [3]. A Factor Hierarchy represents reasons why a factor makes a difference to the legal claim. Experiments show that CATO is an effective teacher [1, 2].

CATO helps students analyze target problems and compare them to past cases. It teaches novices to identify factors in a target, test hypotheses about their significance against cases in its database, and make legal arguments about how to decide the target problems citing cases. Novice users encounter target cases based on real litigated cases, such as the *Space Aero* case.[1] Users identify conflicting factors in the problem,

[1] In the *Space Aero* case, plaintiff Darling manufactured oxygen hoses for Navy aircraft and was the only company capable of their manufacture. Darling had committed over $250,000 in research for hose manufacture. Four long-time employees on Darling's hose-building team left to form Space Aero. Within one month Space Aero was able to manufacture hoses identical to Darling's. None of the defendants had any formal employment contract with Darling: each could be fired or leave at any time. Plaintiff's former supplier provided financial

which give rise to conflicting reasons about how to decide the problem. CATO then teaches them how to make legal arguments to resolve such conflicts.

CATO's Argument Maker tool provides a tutorial on distinguishing. In its original form, the tutorial engages students in an interactive exercise employing didactic explanation. Given a problem situation and case cited by a side in the dispute, students are invited to select from a menu the differences on which to focus in responding to the case by distinguishing. The program then explains whether that selection is a good or bad choice. If a bad choice, the program explains why and illustrates the rebuttal argument one might expect from the other side.

To enable CATO to employ dialectic argument in teaching distinguishing, we developed CATO-Dial, a modification of the program that engages novice users in courtroom-style arguments about target problems like the *Space Aero* case (or any other pair of relevant cases in its database) [4].

Here is an example students encountered with CATO-Dial's version of the tutorial on distinguishing, based on the *Space Aero* case. Students have identified conflicting factors in the *Space Aero* problem and have begun to consider the conflicting legal reasons about how to decide the problem associated with these factors. The Case Analyzer presents the two cases in a tabular form for comparing their factors. As shown in Figure 1, *Space Aero* has five factors, three of which favor the plaintiff (p) and two of which favor the defendant (d). The *Kubik* 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 *Space Aero* for the defendant (i.e., differently from *Kubik*). Note that F1 and F19 strengthen the defendant in *Space Aero* and are not found in *Kubik*, whereas F7 and F21 strengthen the plaintiff's position in *Kubik* and are not found in *Space Aero*. F8 and F16 are also unshared factors, but they are *not* distinctions because they favor deciding *Space Aero* for plaintiff (i.e., the same as in *Kubik*). F8 favors plaintiff in *Space Aero*. Plaintiff won in *Kubik* despite F16.

CATO-Dial places the student in the role of an advocate, Perry Mason, Esq.,[2] 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 7 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 is meant to underscore the student's mistake, and Della's hints, also generated by CATO-Dial, provide instruction on how to rectify them. CATO-Dial can engage in dialogues like this for any pair of relevant cases in its database.

CATO-Dial's dialectic argument contrasts with the didactic explanation of the original CATO. In the latter, the student is not engaged in as direct a manner. Instead, the student is presented with an argument that a hypothetical opponent could make. The footnote shows an excerpt from CATO's didactic explanation of why a student's distinction is incorrect.[3] Just as in step 5 of Figure 2, the student has chosen to

---

backing to Space Aero and conspired with the individual defendants to terminate their employment with Darling. Darling had disclosed the secret information to this supplier in prior dealings. Darling appeared to have taken minimal security measures.

[2] Based on E. S. Gardner's famous Perry Mason. See http://www.ozemail.com.au/~jsimko/.

[3] Here is CATO's didactic explanation of why a student's selection of factor F8 as a distinction is incorrect. It makes a similar point re F16:

distinguish the *Space Aero* case from the *Kubik* case using factors F8 and F16. 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.

Space Aero		**Kubik (p)**
* F1 Disclosure-In-Negotiations (d)		* F7 Brought-Tools (p)
F8 Competitive-Advantage (p)		= F15 Unique-Product (p)
= F15 Unique-Product (p)		F16 Info-Reverse-Engineerable (d)
= F18 Identical-Products (p)		= F18 Identical-Products (p)
* F19 No-Security-Measures (d)		* F21 Knew-Info-Confidential (p)

Fig. 1. Case Comparison in Terms of Factors

COURT IS NOW IN SESSION

1. Judge: You may proceed, Mr. Burger.

2. Mr. Burger for Plaintiff: Your Honor, where plaintiff was the only manufacturer making the product [F15] and defendant's product was identical to plaintiff's [F18], plaintiff should win a claim of trade secrets misappropriation, as in Kubik, Inc. v. Hull, 56 Mich.App. 335, 224 N.W.2d 80 (1974).

3. Judge (to Defendant's counsel): Any response Mr. Mason?

4. Della Street (Whispering in Perry's ear): How are you going to respond on behalf of the defendant? … Burger's analogy to the *Kubik* case is not very strong. Perry, you should distinguish the *Kubik* case, from the *Space Aero* problem. In the top part of this window, click on the factors that are the relevant distinguishing factors of the *Space Aero* problem and the *Kubik* case. If by mistake you clicked on the wrong factor, you can "turn off" that factor by clicking on it again. Finally, click on "Continue".

5. <STUDENT: Distinguish *Kubik* [F8] [F16] (Continue)**> Mr. Mason for Defendant:** Kubik, Inc. v. Hull is distinguishable, because in *Space Aero*, defendant's access to plaintiff's product information saved it time or expense [F8] and plaintiff's product information could be learned by reverse-engineering [F16]. This was not so in *Kubik*.

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

7. Mr. Burger for Plaintiff: Your Honor, as Mr. Mason concedes, in the current problem, defendant's access to plaintiff's product information saved it time or expense [F8]. *Kubik*, which was won by plaintiff, does not have this strength on plaintiff's part but is otherwise similar to the problem at hand. Where this additional strength is present, a decision in favor of plaintiff is warranted even more strongly than it was in *Kubik*. Your Honor, as defendant's counsel pointed out, in Kubik, plaintiff won in spite of the fact that plaintiff's product information could be learned by reverse-engineering [F16]. The current problem does not have this weakness on plaintiff's part and is otherwise similar to Kubik. If plaintiff won in Kubik, then a fortiori plaintiff should prevail in the current problem.

8. Judge (to Defendant's counsel): I am sorry Mr. Mason but at this time I am afraid I have to agree with Mr. Burger. Can you strengthen your argument?

CATO: Suppose defendant made the following argument: "*Kubik, Inc. v. Hull.* is distinguishable, because in *Space Aero*, defendant's access to plaintiff's product information saved it time or expense [F8]. This was not so in *Kubik*." Would this really help defendant? When you have answered this question for yourself, please click on "Continue".

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 access to plaintiff's product information saved it time or expense [F8]. *Kubik*, which was won by plaintiff, does not have this strength on plaintiff's part but is otherwise similar to the problem at hand. Where this additional strength is present, a decision in favor of plaintiff is warranted even more strongly than it was in *Kubik*." In short, even though factor F8 is a difference, defendant cannot use it to distinguish *Kubik*. Defendant needs to focus on other differences, namely, on those differences that make *Space Aero* look stronger for defendant than *Kubik*.

9. <STUDENT: Help! (Continue)**>**

10. Della Street: Perry, here is a hint that may help you. Continue to select factors that can be used to distinguish the case. Even though factor F8 is a difference, defendant cannot use it to distinguish *Kubik*. …. Even though factor F16 is a difference, defendant cannot use it to distinguish *Kubik*. Defendant needs to focus on other differences, namely, on those differences that make *Space Aero* look stronger for defendant than *Kubik*. Continue to select factors that can be used to distinguish the case…. Then click on "Continue".

11. <STUDENT: Distinguish *Kubik* [F1][F19] [F7] [F21] (Continue)**>**

Mr. Mason for Defendant: Kubik, Inc. v. Hull is distinguishable, because in *Kubik*, defendant knew that plaintiff's information was confidential [F21] and plaintiff's former employee brought product development information to defendant [F7]. This was not so in *Space Aero*. Also, in *Space Aero*, plaintiff did not adopt any security measures [F19] and plaintiff disclosed its product information in negotiations with defendant [F1]. This was not so in *Kubik*. These factors are reasons that *Space Aero* is stronger for defendant than is *Kubik*, and therefore, that *Kubik* should not be followed.

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

13. Mr. Burger for Plaintiff: Your Honor, We concede that the defendant's argument can be used to distinguish the case.

14. Della Street: You're so brilliant Perry!

Fig. 2. Sample Dialectic Argument in CATO-Dial

3. Description of Experiment

We compared the impact of teaching distinguishing to senior undergraduates using the two versions of CATO. The students (20 males and 25 females) had all been accepted into law schools and were receiving preparatory instruction through a Summer Institute of the Council on Legal Education Opportunities (CLEO). 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. After taking a pretest, 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.

The pre-test comprised three questions which tested student's case-based 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.

During the first six instructional sessions, students in the experimental and control groups were treated identically. These sessions were designed to introduce students to making arguments with cases and to CATO or CATO-Dial, both of which behaved identically for these initial sessions. 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 programs' Argument Maker tools to select cases to cite.

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. 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 courtroom dialogues like that in Figure 2. The control group worked with the original CATO didactic explanations.

After the students completed all eight sessions of classroom instruction with CATO or CATO-Dial, they took a post-test comprised of three argument questions (Questions 1, 2, and 3) and three transfer questions (Questions 4, 5, and 6). Questions 1, 2, and 3 were worded identically to the pre-test questions but Questions 1 and 2 involved a different problem and cases. The three transfer questions tested the following transfer skills: Question 4 required students to take on 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 required students to apply the skills they had learned with CATO or CATO-Dial to a new domain that they were unlikely to have encountered before: the copyright law doctrine of Fair Use.

A law professor and director of the summer CLEO 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. They received 1 point for citing a factor shared by the problem and case, 2 points for citing a factor in the problem not shared with the case, and 3 points for citing a factor in the problem that could be used to distinguish the case. In grading the recall question, the grader, an undergraduate research assistant, was blind as to which group the students were in.

4. Analysis

Post-test data were available for only 22 of the 45 students, 15 in the experimental group and seven in the control group. Of the 23 students whose data were not available, seven were from the experimental group and 16 were from the control group. Three students dropped out of the CLEO program and our study before Session 7. Five students did not show up for the post-test. Two students completed only the first question of the post-test. Thirteen students were dropped because they had not completed enough work in Sessions 7 and 8, the only sessions involving differential treatment of the experimental and control groups. In determining whether a student had or had not completed enough work in Sessions 7 and 8, the following criterion was adopted. If a student completed one or more of the eight pair-wise case comparisons presented in Sessions 7 and 8 using the relevant dialogue feature of CATO, he or she was retained. By relevant dialogue feature of CATO, we mean the didactic explanation mini-dialogue in the case of the control group and the dialectical argument in the case of the experimental group. In order to complete a pair-wise comparison, a student had to find all the relevant differences between the two cases. Whether a student completed a pair-wise comparison was assessed by examining the command log files of CATO or CATO-Dial for Sessions 7 and 8. Since students worked in pairs, if the log file failed to show completion of the pair-wise comparison,

neither student in the pair received credit for that comparison. Conversely, if the log file showed completion, both students received credit.

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. For each student, responses to the three argument questions (i.e., Questions 1, 2, and 3) were summed, as were responses to the three transfer questions (i.e., Questions 4, 5, and 6) (see Figure 3.) 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 argument questions. For the transfer questions, however, the mean post-test score of the experimental group was significantly higher than that of the control group (t (6.8) =2.63, $p < .05$, effect size of 1.30).[4]

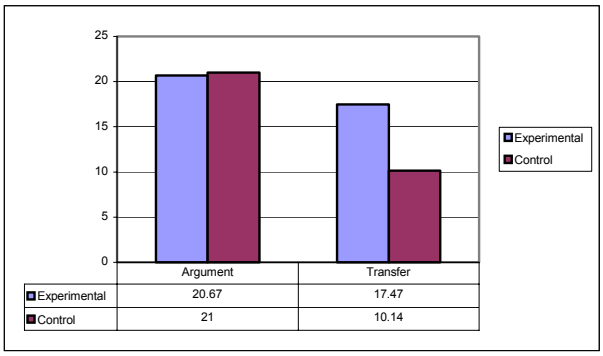

Fig. 3. Comparison of Mean Post-Test Scores on Argument and Transfer Questions

5. Discussion and Conclusion

The data confirmed in an interesting way our hypothesis that students would learn the skill of distinguishing better by engaging in simulated dialectical argument than in interactive didactic explanation. Whereas dialectic argument was not more effective than didactic explanation in teaching distinguishing skills, it was significantly more effective in helping students transfer the skills they learned to new tasks and 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.

Apart from the transfer difference, there was some other evidence that the experimental students had a deeper understanding of the importance that context plays in the task of finding distinctions. In addition to assigning numerical scores to answers, the grader evaluated the answers in terms of four grading criteria, each

[4] 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 two groups were not equal.

involving a simple binary positive-or-negative scale. Two suggestive, though nonsignificant, differences emerged, both of which favored the experimental group. 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 a fairly superficial transformation from CATO to CATO-Dial in the presentation of the lesson on distinguishing had such dramatic benefits. 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 in the courtroom simulation motivating and conducive to paying attention and learning. The courtroom role-playing context adds an element of competition to the task: the students may feel they are trying to defeat CATO-Dial. This context also adds a sense of fun to the task.

In addition, students may have found it is easier to understand the program's responses in CATO-Dial than in CATO. In the latter, the program is in the somewhat awkward position of explaining by example why a student's response is good or bad. The dialectical argument simulation, by contrast, provides a more natural context for illustrating the effect in an ongoing dialogue regarding a student's choices.

Role-playing in a courtroom argument, with that role's cognitive and emotional expectations, may also be important. For one thing, 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 [8]. For another thing, dialectical argument may 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 and thus help transfer. In a recent investigation involving business school education, students who compared cases in a study phase were three times more likely to transfer the implicit principle of the cases to a new application than were those who simply read the cases for the purposes of advice-giving [14].

It is worth noting that students in the experimental group reported 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?", students who used CATO-Dial rated it as more helpful than did students who used the original CATO (Ms = 6.76 and 5.56 out of 10, respectively.) This finding is consistent with the fact that many more students provided complete data in the CATO-Dial group than in the CATO group.

Law professors aim to teach research and argumentation skills that students can transfer to new legal domains and tasks. They know that law students are likely to practice in different areas of law than the ones students studied in law school, and that over the course of a student's legal practice, the laws are likely to change.

Consequently, the results of this study suggest that the CATO-Dial approach is potentially quite valuable. Our subsequent work will focus on converting as much of the CATO curriculum as possible to a dialectical format.

References

1. Aleven, V. (1997) *Teaching Case-Based Argumentation Through a Model and Examples*, Ph.D. Dis., U. Pittsburgh, unnumbered Tech. Rep. LRDC/ISP.
2. Aleven, V. and Ashley, K.D. (1997) "Teaching Case-Based Argumentation Through a Model and Examples". *Proc. 8th World Conf. AI in Ed. Soc.* 87-94. IOS Press: Amsterdam.
3. Ashley, K.D., (1990). *Modeling Legal Argument: Reasoning with Cases and Hypotheticals*. The MIT Press / Bradford Books, Cambridge, MA.
4. Ashley, K. D. (2000) "Designing Electronic Casebooks That Talk Back: The CATO Program". In *Jurimetrics* Vol. 40, No. 3, pp. 275-319.
5. Ashley, K.D. (2002) "An AI Model of Case-Based Legal Argument from a Jurisprudential Viewpoint." In *Journal of Artificial Intelligence and Law*. Kluwer: Dordrecht, Neth.
6. 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.
7. Centinia, F., T. Routen, A. Hartmann, and C. Hegarty (1995) "STATUTOR: Too Intelligent By Half?" In *Legal Knowledge Based Systems JURIX '95*. 121-132. Lelystad: Koninklijke.
8. Chi, Michelene T.H., S. Silver, H. Jeong, T. Yamauchi, and R. Hausmann (2001) "Learning From Human Tutoring" in *Cognitive Science*, Vol. 25, pp. 471-533.
9. Collins, A. and Stevens, A. L. (1982). "Goals and Strategies of inquiry Teachers". In *Advances In Instructional Psychology*, R Glaser (ed.) pp. 65-119. Hillsdale, NJ: Erlbaum.
10. Retalis, S., H. Pain and M. Haggith. (1996) "Arguing with the Devil; Teaching in Controversial Domains". *Int. Tutoring Sys., 3d Intl Conf., ITS-96*. 659-667. Berlin: Springer.
11. Rissland, E.L. (1990) "Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning". *Yale Law Journal* 99. 1957-1981. June 1990. Number 8.
12. 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.
13. Span, G. (1993) "LITES, an Intelligent Tutoring System for Legal Problem-Solving in the Domain of Dutch Civil Law". In *Proc. 4th Intl Conf. AI and Law*, 76-81. New York: ACM.
14. 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.
15. Wong, L., C. Quek, and C. Looi. (1997) "PADI-2: An Inquiry-based Geography Tutor". In *AI in Education*, Proc. AI-ED 97 World Conf. 47-54. Amsterdam: IOS Press.