1-1-2008

No Computer Program Required: Even Pencil-and-Paper Argument Mapping Improves Critical Thinking Skills

Maralee Harrell
Carnegie Mellon University

No Computer Program Required: Even Pencil-and-Paper Argument Mapping Improves Critical Thinking Skills

Maralee Harrell

Argument mapping software abounds,[1] and one of the reasons is that using the software has been shown to teach/promote/improve critical thinking skills. These positive results are very encouraging, but they also raise the question of whether the computer tutorial environment is producing these results, or whether learning argument mapping, even with just paper and pencil, is sufficient. Based on the results of two empirical studies, I argue that the basic skill of being able to represent an argument diagrammatically plays an important role in the improvement of critical thinking skills. While these studies do not offer a direct comparison between the two methods, it is important for anyone wishing to employ argument mapping in the classroom to know that significant results can be obtained even with the most rudimentary of tools.

Introduction

Although there is no universally accepted definition of "critical thinking" skills, most philosophers and other educators agree that one aspect of critical thinking involves the ability to reconstruct, understand, and evaluate an argument—tasks we may call, for the sake of brevity, "argument analysis." For example, Kuhn (1991) says that "argumentative reasoning skills are in fact fundamental to what educators call 'critical' thinking." (p. 5), and Ennis (1987) says that "analyzing arguments" is one of the critical thinking abilities. This covers identifying the stated and unstated premises and the conclusion, and "seeing the structure of an argument." (p. 12). In addition, according to Moore and Parker (2007),

> critical thinking is more than just thinking or making decisions or acting selectively. Above all else, thinking critically means screening your ideas to see if they make sense….Critical thinking requires evaluating arguments that support the claims we are considering, and weighing them against those that support alternative or contrary views. (p. 2-3)

Similarly, Lee (2002) says:

> Critical thinking is reasoning. It involves seeking to establish whether claims are true by considering reasons that may show those claims to be true, or show how they are true. Reasons are themselves claims. Therefore, critical thinking involves considering various claims and determining how some of them may show others to be true…. An argument is an instance of critical thinking or reasoning. It is an effort to show that some claim is true by giving reasons that support the claim. The claim that is supported by the reasons is the conclusion of the argument…. There are three main skills involved in critical thinking: (1) identifying the reasoning or arguments of others, (2) evaluating the reasoning or arguments of others, and (3) creating reasoning or arguments of your own. (p. 2-3)

[1] I have recently reviewed several of these software packages in this journal (Harrell, 2005): Araucaria v. 2.0, © Chris Reed and Glen Rowe, http://www.computing.dundee.ac.uk/staff/creed/araucaria, Argutect v. 4.0, © Knosis, http://www.knosis.com/argutect.html, Athena Standard v. 2.2, © Bertil Rolf and Charlotte Magnusson, http://www.athenasoft.org, Inspiration v. 7.5, © Donald Helfgott and Mona Westhaver, http://www.inspiration.com, Reason!Able v. 1.1, © Tim van Gelder and Andrew Bulka, http://www.goreason.com. Since that review, Austhink Software, headed by Tim van Gelder has developed a new argument mapping software package called Rationale, available here: http://www.austhink.com.

In the introductory philosophy class at Carnegie Mellon University (*80-100: What Philosophy Is*) one important learning goal is the development of argument analysis skills. But typically, even in philosophy classes at Carnegie Mellon (other than logic courses), students are not taught these skills explicitly. Instead, if they are taught these skills at all, they are taught implicitly by demonstration from an instructor—e.g., when an instructor writes out the premises of an argument on the chalkboard and leads students through a discussion of the truth of the premises and how well they support the conclusion. Students are also often asked to reconstruct an author's argument—in, e.g., a critical essay—but are not given any general guidelines or methods for completing this kind of task.

We do, however, believe that our students improve their argument analysis skills after taking the introductory philosophy course, and we predict that students in the introductory philosophy course will exhibit significant improvement on argument analysis tasks over the course of the semester. In addition to determining whether they are improving, though, we are particularly interested in the efficacy of alternative teaching methods to increase this type of critical thinking performance.

Argument Maps

I first became fascinated with argument mapping after stumbling across Tim van Gelder's Reason!Able argument mapping software. I began to explore the software by mapping the arguments I was teaching in my classes. I found, to my surprise, that I did not understand these arguments as well as I thought I had, and that the mapping was forcing me to analyze and synthesize in a way that I had never done before.

This epiphany lead me to wonder whether learning how to map arguments could be helpful to my students, especially in introductory philosophy, who have trouble reading philosophy. It's not, of course, that they can't *read*, rather that they don't look for, and therefore don't see, the different parts of an argument—premises, conclusion, inferential connections, objections, etc.

Even for those experienced with philosophy, it is often difficult to determine just what the argument is in a philosophical text. Not only does the text contain many more sentences than the propositions that comprise the argument, but also, proceeding necessarily linearly, the prose obscures the inferential structure of the argument. Thus anyone who wishes to understand and evaluate the argument may reasonably be confused. This is where argument mapping is useful. This activity helps us to discover the various elements of an argument, and ultimately demonstrates where the argument may be criticized.

I begin with the logician's definition of an argument: a series of statements in which one is the conclusion and the others are premises, purportedly providing support for the conclusion. An argument map uses boxes and arrows to represent two particular aspects of an argument: the statements—contained in the boxes—and the inferences—indicated by the arrows. Consider the short, albeit profound, argument given by Descartes. "I think, therefore I exist." There are two statements here—"I think" and "I exist." As indicated by the "therefore" the first is the premise and the second is the conclusion. So we say that the premise, "I think," is intended to support the conclusion; it is intended to be a reason to believe that the conclusion, "I exist," is true. One way to represent this argument is fairly standard in logic textbooks and else where. It is what I call the "list:"

P1: I think.
C: I exist.

Alternatively, this argument can be represented in a simple map:

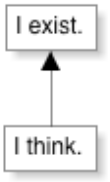

Figure 1 Argument map representing a version of Descartes' *cogito*.

Here, the arrow represents the inference from premise to conclusion, and so is pointing in the direction of support.

Of course, most arguments are more complex than this. Often there are sub-arguments supporting some of the premises, some premises need to be combined with other premises and/or objections and replies may be included. For example, consider the following argument from Descartes' *Third Meditation* concerning the origin of his idea of God:

> *It only remains to me to examine into the manner in which I have acquired this idea from God; for I have not received it through the senses, [since] it is never presented to me unexpectedly, as is usual with the ideas of sensible things when these things present themselves, or seem to present themselves, to the external organs of my senses; nor is it likewise a fiction of my mind, for it is not in my power to take from or add anything to it; and consequently the only alternative is that it is innate in me, just as the idea of myself is innate in me.* (Descartes, 1641)

Here, Descartes is arguing that the idea of God is innate, i.e. this is the conclusion of his argument. The reasons he gives for believing this conclusion are the two premises: the idea did not come from to him through his senses, nor did he invent the idea. And he gives further reasons for each of these premises. The way that all of these premises are inferentially connected is what I call the "structure" of the argument. Representing the argument as a list can obscure this structure because there is no indication of how the premises work together to support the conclusion:

P1: The idea of God is never presented to me unexpectedly.
P2: It is usual that the ideas of sensible things present themselves, or seem to present themselves, to the external organs of my senses unexpectedly.

P3: I have not received the idea of God through the senses.
P4: It is not in my power to take from or add anything to the idea of God.

P5: The idea of God is not a fiction of my mind.

C: The idea of God is innate in me.

The argument can be represented more fruitfully, I believe, by the following map:

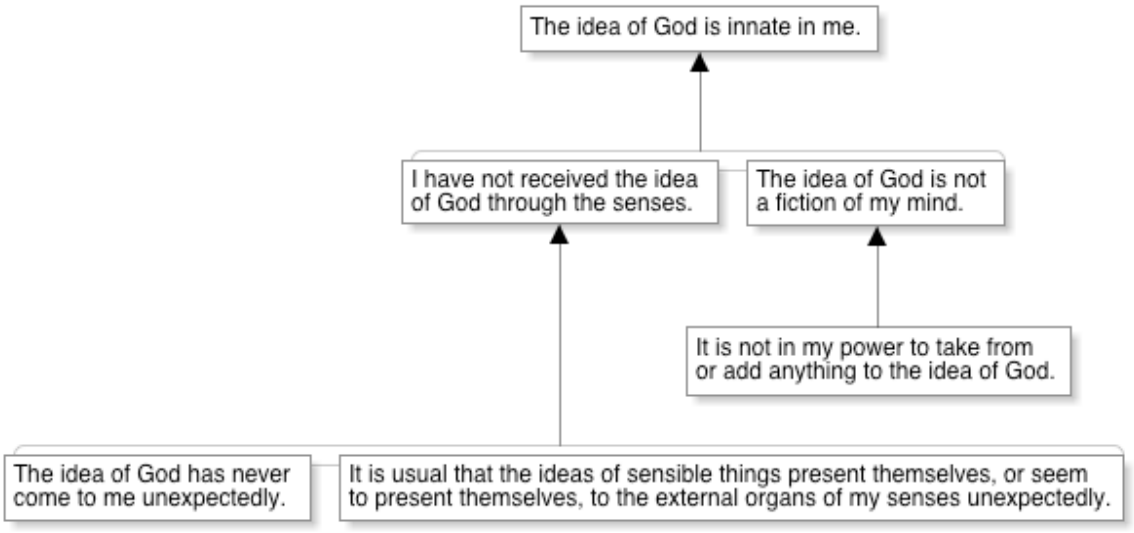

Figure 2 Argument map representing an argument from Descartes' *Third Meditation*.

This map makes both the content and the structure of the argument clear by eliminating unnecessary language, and visually representing the way the premises work to support the main conclusion. Because of this, it is clear how to proceed in evaluating the argument. We can consider validity/strength by assuming that the statements in the boxes are true and focusing on the quality of the inferences (the arrows); and we can consider soundness/cogency by focusing on the status of the premises (the contents of the boxes).

There are a few rules of thumb to follow when mapping an argument. First, look for premise and conclusion indicators. These are words that serve to signal that a particular statement is supporting (a premise for) or follows from (a conclusion based on) another statement. Some common premise indicators are: since, because, for, given that; and some common conclusion indicators are: therefore, thus, so, hence. Second, rewrite statements as individual sentences that can be understood on their own, eliminating the connecting words and any other rhetorical content. For example, rhetorical questions are often disguised claims that are an important part of the argument and should be rewritten as statements. Third, supply missing or implied premises and/or conclusions where appropriate.[2] For example, often an author, instead of explicitly stating a conclusion, just assumes her reader will draw it from one argument and then use that conclusion as a premise in a further argument.

Finally, clearly indicate the difference between premises that need to be combined in order to support a conclusion, and premises that are each separate reasons to believe a conclusion. In some arguments, like *modus ponens* or disjunctive syllogism, the premises work together to support the conclusion; neither premise itself provides any support, but together they provide a good reason to believe the conclusion is true. These are called "linked" arguments, and

[2] I say "where appropriate" here because, with regard to unstated premises, different kinds of argument maps can be constructed depending on the intent of the map constructor. If the map constructor wants to map only the argument actually given in the text, then she may opt only to insert missing sub-conclusions, for otherwise the premises could not be fully connected. If, on the other hand, she wants to map the argument given in the text along with premises that the author seems take for granted, or that the author doesn't seem to realize the argument needs, etc., then she may opt to insert everything implied, assumed, or otherwise needed to create possibly a better argument than the author originally put on paper.

are represented in an argument map by linking the premises together and having only one arrow point to the conclusion. For example, consider the following argument:

> *If you earned an A on the final paper, then you earn an A in the course. You earned an A on the final paper, so you earn an A in the course.*

This argument can be mapped as:

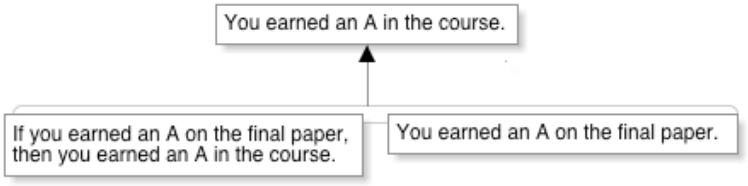

Figure 3 Example of a linked argument.

On the other hand, in some arguments, the premises each would provide support for the conclusion, even if the other premises were ignored; these are called "convergent" arguments, and are represented in a map by having an arrow pointing to the conclusion for each independent premise. Consider this argument:

> *David is a good student. He studies hard, he always participates in class discussions, and he comes in for help on all of his papers.*

Each of these reasons by themselves would support the contention that David is a good student, but the argument may be stronger if the are all included. This can be mapped as:

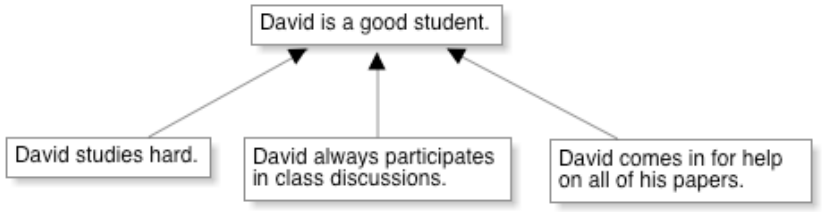

Figure 4 Example of a convergent argument.

Teaching Argument Mapping

There is a substantial amount of evidence that visualization, mapping, and the use of graphic organizers in general are very effective learning aids. Both Larking and Simon (1987) and Winn (1991) argue that diagrams, compared to plain text, make information search and recognition faster and more efficient, and research on student learning has consistently shown the efficacy of using diagrams to aid text comprehension (Armbruster & Anderson, 1984; Dansereau, et al.; Novak & Gowin, 1984; Schwartz & Rafael, 1985), as well as vocabulary development, postreading activities and writing preparation (Johnson, et al., 1986).

Thus, I began, to teach argument mapping in the introductory philosophy course I teach every semester at Carnegie Mellon. I use the mapping mainly for assistance in the analysis and critique of arguments presented in primary source texts (as with the example from Descartes above). At the start of the semester, I teach the students the basics of constructing argument maps to accurately reflect the arguments presented in the text. Then, the students are given short passages to map as homework throughout the semester.

Research on computer-supported argument visualization by Tim van Gelder and others has shown that the use of software programs specifically designed to help students construct argument maps can significantly improve students' critical thinking abilities over the course of a

semester-long college-level course (Kirschner, Shum & Carr, 2003; Twardy, 2004; van Gelder, 2001, 2003). But, of course, one need not have computer software to construct an argument map; one needs only a pencil and paper, and my students were allowed to construct their maps using any method they desired. To my knowledge there has been no research to determine whether the crucial factor is the mere ability to construct argument maps, or the aid of a computer platform and tutor, or possibly both.[3]

My students are not required to use any sort of computer program to draw their maps, and while many do, many others simply draw them with pencils on paper. I saw marked improvement in my students' argument analysis skills when I taught the use of argument maps, and hypothesized that the ability to construct argument maps that accurately reflect the text they are analyzing is a considerable aid for improving of students' argument analysis skills, even if they are pencil-and-paper maps. This implies, of course, that students who are able to construct pencil-and-paper argument maps and use them during argument analysis tasks should perform better on these tasks than students who do not have this ability.

Testing the Efficacy of Argument Mapping

Carnegie Mellon University's introduction to philosophy course (*80-100 What Philosophy Is*), was a natural place to study the skills acquisition of our students. We typically teach 4 or 5 lectures of this course each semester, with a different instructor for each lecture. While the general curriculum of the course is set, each instructor is given a great deal of freedom in executing this curriculum. For example, it is always a topics based course in which epistemology, metaphysics, and ethics are introduced with both historical and contemporary primary-source readings. It is up to the instructor however, to choose a text, the order of the topics, and the assignments. The students who take this course are a mix of all classes and all majors from each of the seven colleges across the University.

In this study, we tested two hypotheses. The first is that all of our students, no matter how they are taught, are gaining argument analysis skills by taking our introductory course. This is important to know if we are then going to inquire which students gained more. This hypothesis implies that, on average, all students in introductory philosophy will exhibit significant improvement on argument analysis tasks over the course of the semester. The second hypothesis is that the ability to construct argument maps that accurately reflect the text they are analyzing is a considerable aid for improving students' argument analysis skills (more of an aid that being able to represent an argument some other way). This second hypothesis implies that students who are able to construct argument diagrams and use them during argument analysis tasks should perform better on these tasks than students who do not have this ability.

In the Spring of 2004 (Semester 1), four different sections of 80-100 were taught, and in the Fall of 2004 (Semester 2), five different sections were taught. The students in Sections 1 and 3 in Semester 1 and the students in Sections 1, 4 and 5 in Semester 2 were explicitly taught how to construct argument maps to represent a selection of text. In contrast, students in Sections 2 and 4 in Semester 1 and students in Sections 2 and 3 in Semester 2 were not explicitly taught the use of argument maps, but rather—if they were taught any specific methods of analyzing arguments—were taught (only implicitly) to use more traditional kinds of representations (e.g. lists as described above).

[3] This paper is not intended to be this sort of direct comparison; rather, it is an argument that one can obtain substantial gains in critical thinking with argument mapping, even without a computer program.

Participants

One hundred and thirty-nine students (46 women, 93 men) in each of four lectures (each with a different instructor) of introductory philosophy in the Spring of 2004 (Semester 1), and 130 students (36 women, 94 men) in each of five lectures (each with a different instructor) of introductory philosophy in the Fall of 2004 (Semester 2) of Carnegie Mellon University's introductory philosophy course (80-100) were studied. Each section of the course had a different instructor and teaching assistant, and the students chose their section. Sixty-seven students (23 women and 44 men) in Semester 1, and 68 students (21 women and 47 men) in Semester 2 were explicitly taught how to construct argument maps, while 72 students (23 women and 49 men) in Semester 1 and 62 students (15 women and 45 men) in Semester 2 were not taught how to construct argument maps.

Materials

To evaluate the development of our students' skills over the course of a semester, the four instructors of the course in Spring 2004 developed a "pretest" and a companion "posttest". The tests in Semester 1 each consisted of 6 questions, each of which asked the student to analyze a short argument. In questions 1 and 2, the student was only asked to state the conclusion (thesis) of the argument. This proved to be an easy task for all students, and so questions like these were not included in the Semester 2 tests. Questions 3-6 in Semester 1 and Questions 1-5 in Semester 2 each had five parts: (a) state the conclusion (thesis) of the argument; (b) state the premises (reasons) of the argument; (c) indicate (via multiple choice) how the premises are related; (d) provide a visual, graphical, schematic, or outlined representation of the argument;[4] and (e) decide whether the argument is good or bad, and explain this decision. For each argument on the pretest, there was a structurally (nearly) identical argument with different content on the posttest (see Appendices A & B for examples of the tests).[5]

In this study, we tested our first hypothesis—that all our students are improving their argument analysis skills—by comparing pretest and posttest scores of all the students in 80-100 in both semesters. We tested the second hypothesis—that students who learn argument mapping will improve their argument analysis skills more than students who do not by comparing pretest and posttest scores in both semesters. Specifically, we compared (a) the pretest and posttest scores of students who were explicitly taught argument mapping to students who were not taught this skill, (b) the pretest and posttest scores of students who actually constructed argument maps on the posttest to those students who did not, and (c) the pretest and posttest scores of students who constructed *correct* argument maps on the posttest to those students who did not in both semesters.

[4] No student was *required* to provide any *particular* representation on either the pretest or posttest. In fact, several students who were taught argument mapping did not, in fact, provide argument maps on every question.

[5] The scores on the pretests and posttests were calculated as follows. Each question part, except for part (d), were coded 1 for a correct answer, and 0 for an incorrect answer. Part (e) of each question was coded 1 if the student gave as reasons claims about support of premises for the conclusion and/or truth of the premises and conclusion. The score on the test was then the sum of the scores on parts (a), (b), (c) and (e) of each question divided by the total number of question parts (20, when (d) is excluded). After reviewing the tests, we determined that there were some standard representation schemes that students used, so for part (d) of each question, answers were coded according to the type of representation used: Correct argument diagram, Incorrect or incomplete argument diagram, List, Translated into logical symbols like a proof, Venn diagram, Concept map, Schematic like: P1 + P2/C, Other or blank.

The primary variables of interest were the pretest and posttest scores (expressed as a fraction correct of the equally weighted question-parts), whether the student provided an argument map for each multi-part question on the posttest for each semester, and whether the student provided the *correct* argument map for each multi-part question on the posttest for each semester. In addition, the following data was recorded for each student in each semester: which section the student was enrolled in, the student's final grade in the course, the student's year in school, the student's home college,[6] the student's sex,[7] and whether the student had taken the concurrent honors course associated with the introductory course.

Gain from Pretest to Posttest for All Students

The first hypothesis was that the students' argument analysis skills improved over the course of the semester. This hypothesis was tested by determining whether the average gain of the students from pretest to posttest was significantly positive for each semester.[8]

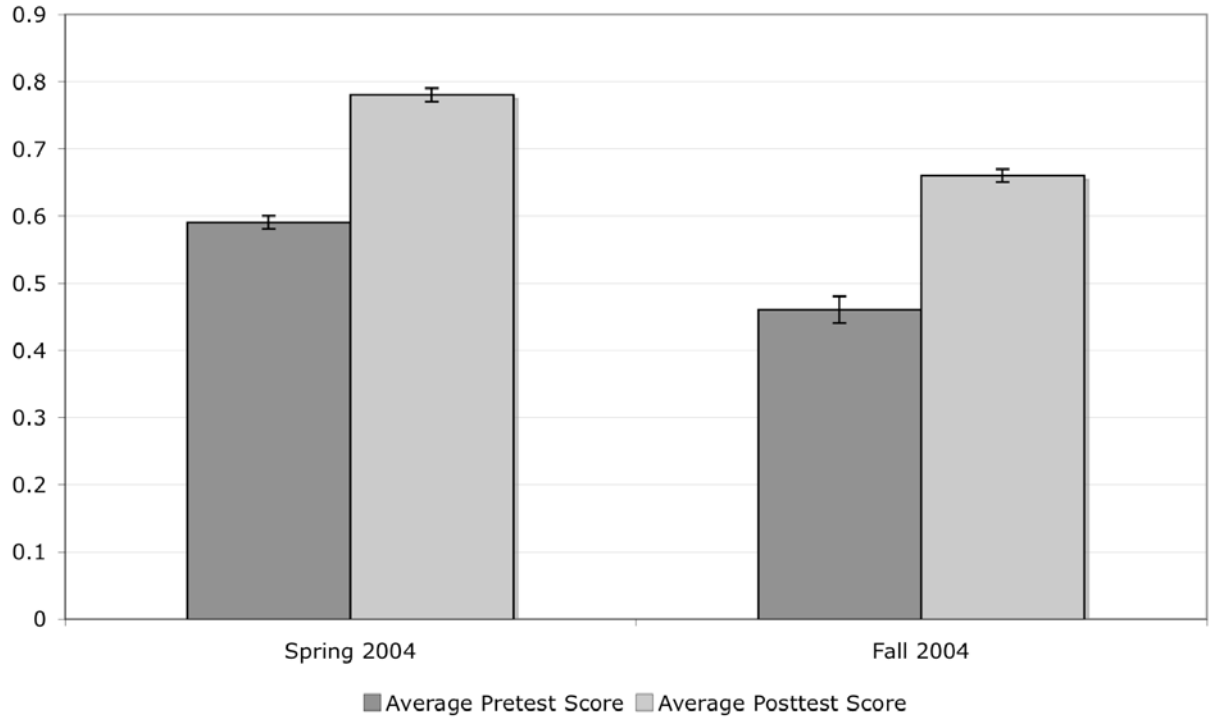

Figure 5 Comparison of pretest and posttest scores for each semester.[9]

[6] There are seven colleges at Carnegie Mellon in which undergraduate students may be enrolled: the College of Fine Arts (CFA), the Carnegie Institute of Technology (CIT), Carnegie Mellon University Honors College (CMU), the College of Humanities and Social Sciences (HSS), the Mellon College of Science (MCS), the School of Computer Science (SCS), the Tepper School of Business (TSB). The distribution of students in 80-100 from each college is given in Table A in Appendix C.

[7] The distribution of students in 80-100 by sex is given in Table B in Appendix C.

[8] For the benefit of those who are interested, I give the standards of statistical significance used in footnotes. Here, in both Semesters, the difference in the means of the pretest and posttest scores was significant (paired *t*-test; $p < .001$), and the mean gain was significantly different from zero (1-sample *t*-test; $p < .001$).

[9] For each bar graph, the height of the bar indicates the value of the mean of the variable, and the I at the top of each bar indicates the standard error of the mean.

From these results we can see that our first hypothesis is confirmed: in each semester, overall the students did have significant gains from pretest to posttest.

Gain from Pretest to Posttest by Map Use

The second hypothesis was that students who learn argument mapping will improve their argument analysis skills more than students who do not learn this skill. This hypothesis was first tested by comparing the gain from pretest to posttest as well as the relative improvement (relative gain) from pretest to posttest (gain as a percentage of what could have been gained) for students who were taught argument mapping vs. students who were not taught argument mapping. Relative improvement is as important to measure as the straight gain, since the gains of many students can be hampered by ceiling effects—i.e., they did not gain as much because they had high pretest scores, and so did not have as much to gain as students with lower pretest scores.

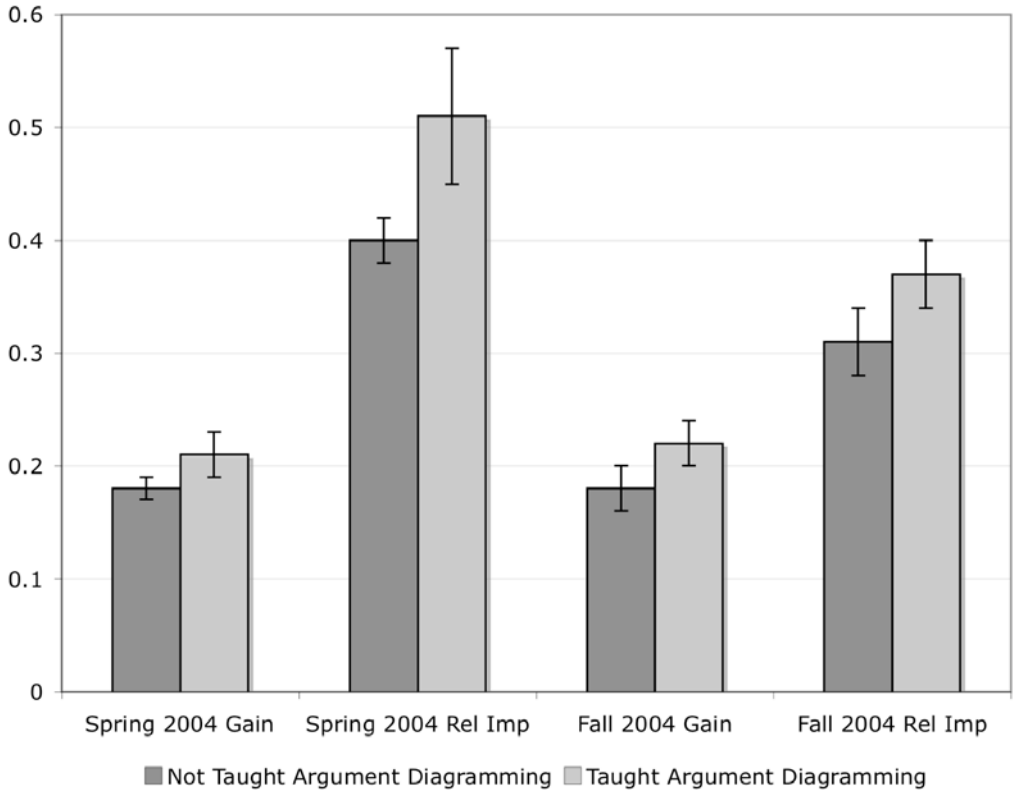

Figure 6 Comparison of gain and relative improvement for students who were taught argument mapping vs. students who were not taught argument mapping for each semester.

These results are striking—the differences between both the gain and the relative improvement among the two groups of students were significant in both semesters.[10]

[10] Here we tested a generalized liner model (GLM) of the variable Gain using the variable Taught as the predictor, and the variable Pretest as a covariate (Minitab 15, Minitab Inc., State College, PA). This analysis indicates that for both semesters, the differences in the pretest scores was significant for predicting the gain (Semester 1: $F = 106.98$, $p < .001$; Semester 2: $F = 83.62$, $p < .001$), and the relative gain (Semester 1: $F = 29.14$, $p < .001$; Semester 2: $F = 18.06$, $p < .001$), and that, even controlling for differences in pretest score, the differences in whether a student was

These results may not be fully informative, though, if we are interested in whether acquiring the skill of argument mapping contributes significantly to improvement on argument analysis tasks. For, as much we try, students do not always absorb and/or apply what we teach, and many students already know what we teach them. As it turned out, many students who were taught argument mapping did not construct argument maps at all on the posttest while many students who were not taught argument mapping in this course did in fact construct argument maps on the posttest. Thus, we tested our second hypothesis a second time, by comparing the gain and relative improvement from pretest to posttest of students who did construct argument maps on the posttest vs. students who did not.

This comparison was not straightforward, though, since there were four opportunities to construct a map on the posttest in Spring 2004 and five in Fall 2004, and it wasn't all or nothing for most of the students. So, to do this comparison, we grouped the students in each semester into two categories: those who constructed few (0-2) maps on the posttest and those who constructed many (3 or more) maps on the posttest. We then compared the gain and relative improvement from pretest to posttest among these groups.

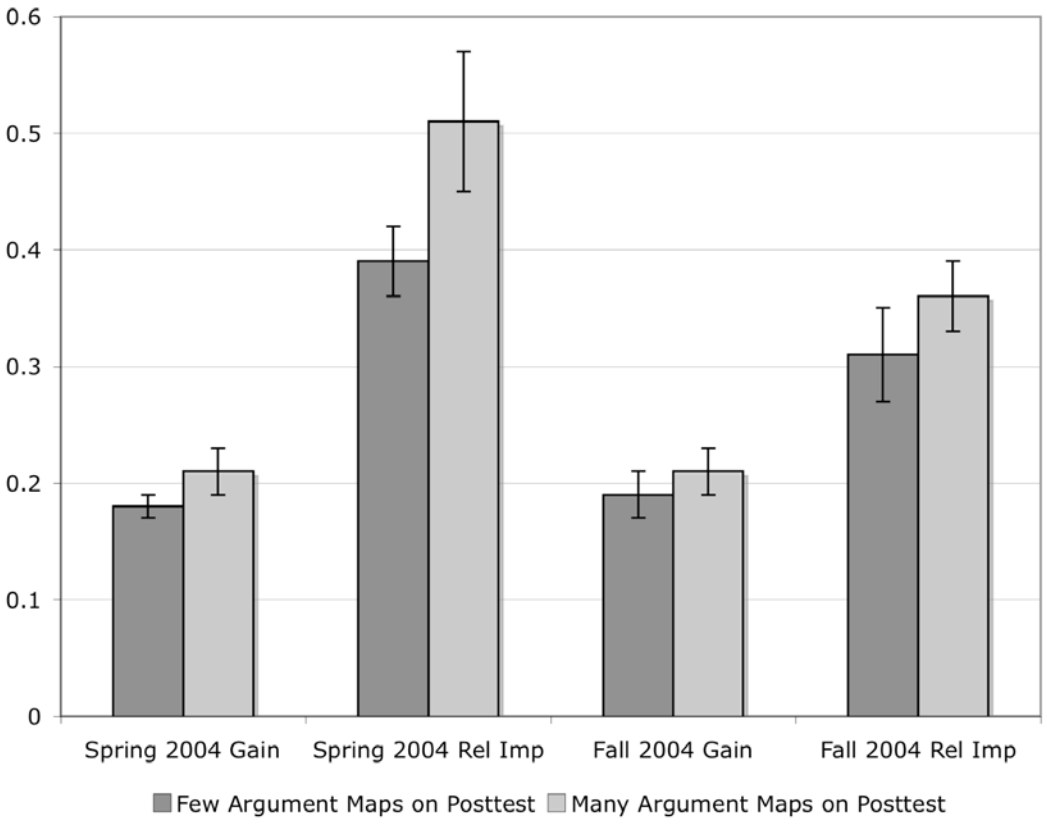

Figure 7 Comparison of gain and relative improvement for students who constructed few vs. many argument maps for each semester.

These results show that while the straight gains of the students who constructed many argument maps are only marginally significantly better than the gains of students who constructed no or few argument maps, the relative improvement of the students who constructed many is

---

taught argument mapping were significant for predicting gain (Semester 1: $F = 13.92$, $p = .001$; Semester 2: $F = 6.07$, $p = .01$), and relative gain (Semester 1: $F = 6.84$, $p < .001$; Semester 2: $F = 4.34$, $p < .001$).

significantly better.[11] The explanation may be that that students who constructed many had higher pretest scores and so are more affected by the ceiling effects that the relative improvement measure dispels.

An even more informative test of our second hypothesis, though, would be to compare the gains made by students who applied their argument-mapping skill *correctly*, and not just attempting some kind of map. To do this, we divided the students into similar groups as above: few (0-2) correct argument maps, and many (3 or more) correct argument maps. We then compared the gain and relative improvement of the students in each group.

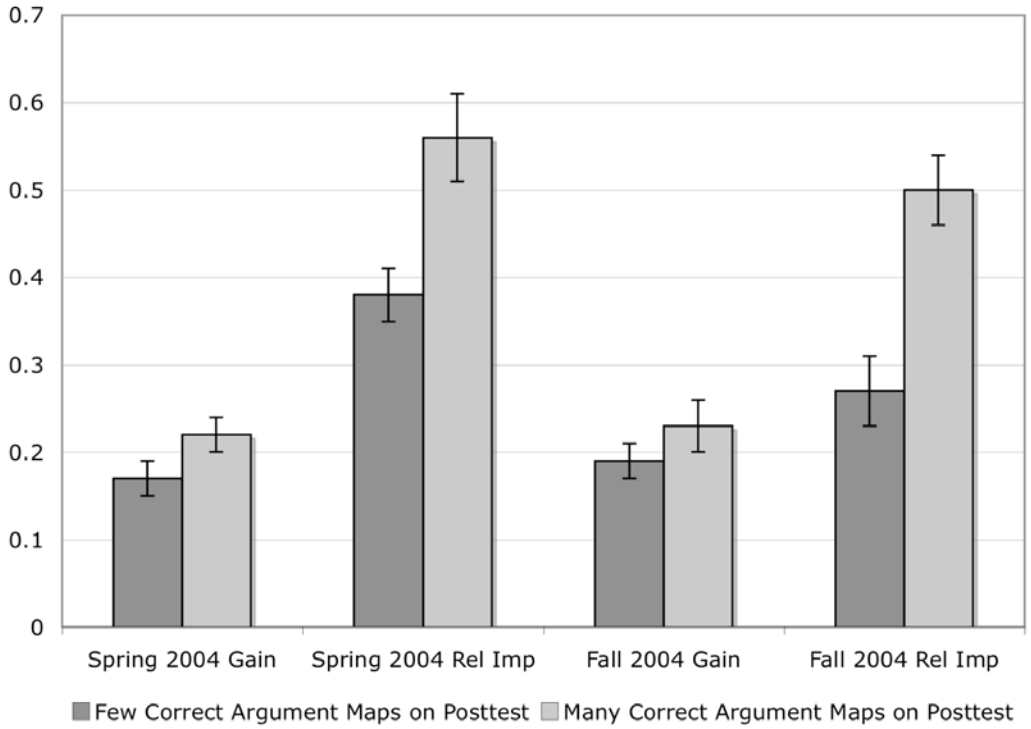

Figure 8 Comparison of gain and relative improvement for students who constructed few vs. many correct argument maps for each semester.

These results show that the differences between both gains and relative improvements for the students who constructed many correct argument maps on the posttest are significantly better than for the students who constructed few.[12]

[11] Again, we tested a GLM on the variable Gain using the variable Maps as the predictor and the variable Pretest as a covariate. This analysis again indicates that the differences in the pretest scores was significant for predicting the gain (Semester 1: $F = 132.00$, $p < .001$; Semester 2: $F = 134.58$, $p < .001$). In addition, this analysis indicates that, even accounting for differences in pretest score, the differences in the gain between students who attempted to construct many argument maps on the posttest and the students who attempted to construct few argument diagrams on the posttest were significant (Semester 1: $F = 28.13$, $p < .001$; Semester 2: $F = 34.67$, $p < .001$).

[12] Again, we tested a GLM on the variable Gain using the variable CorrectMaps as the predictor and the variable Pretest as a covariate. This analysis again indicates that the differences in the pretest scores was significant for predicting the gain (Semester 1: $F = 132.19$, $p < .001$; Semester 2: $F = 133.00$, $p < .001$). In addition, this analysis indicates that in each semester, even accounting for differences in pretest score, the differences in the gain from pretest to posttest between students who constructed many correct argument maps and students who constructed few correct argument maps were significant (Semester 1: $F = 28.13$, $p < .001$; Semester 2: $F = 37.78$, $p < .001$).

Limitations of the Study

There are at least two points about this study that are worth scrutiny. First, we did not use any standardized tests, like the California Critical Thinking Skills Test (CCTST) for our pre- and posttests. The reason we did is that we are specifically concerned with the subset of critical thinking skills that we call "argument analysis skills," and the CCTST does not test specifically for these skills. It tests, rather, some more general problem-solving skills that, while clearly important, are not the focus of our introductory philosophy course. One remedy for this, then, may be the use of another standardized test that does specifically target argument analysis skills. For example, the Ennis-Weir Critical Thinking Essay Test, developed by Robert Ennis and Eric Weir in 1985, focuses on argument analysis and construction. At the time of this study we did not have the resources to implement this test, but we would like to pursue this option in the future.

The second is that, while we did employ control groups in the form of students who were not taught argument mapping, this difference was not the only one, and others may be relevant. For instance, the students who were taught argument mapping had different instructors, different TAs and different assignments than the students who were not taught mapping. To alleviate possible errors introduced by these uncontrolled variables, we performed several analyses on our data that indicated that the specific instructor a student had was not a factor in determining that student's gain or relative improvement. In addition, our analysis showed that none of the following variables was a factor either: the student's final grade in the course, the student's year in school, the student's home college, the student's sex, whether the student had taken the concurrent honors course associated with the introductory course. Rather, the only factors that were factors were the student's pretest score and the type of representation used on the posttest.[13]

Lastly, although not a critique of this particular study, it is the case that there may be many different ways of teaching argument mapping, e.g. focusing on computer-based tools, or focusing more on argument construction rather than analysis. We hope that the results of our study will encourage others to experiment with argument mapping in their classrooms, and share experiences and ideas.

Conclusion

One set of skills we would like our students to acquire by the end of our introductory philosophy course can be loosely labeled "the ability to analyze an argument." This set of skills includes the ability to read a selection of prose, determine which statement is the conclusion and which statements are the premises, determine how the premises are supposed to support the conclusion, and evaluate the argument based on the truth of the premises and the quality of their support.

One purpose of argument maps is to aid students in each of these tasks. An argument map is a visualization of an argument that makes explicit which statement is the conclusion and which statements are the premises, as well as the inferential connections between the premises and the conclusion. Since an argument map contains only statements and inferential connections, it is

[13] Regression analyses indicated that of all the variables tested in both semesters (Pretest, Lecturer, Sex, Year, College, Honors, Grade, and RepresentationType), only Pretest and RepresentationType were significant predictors of gain and relative gain ($p < .01$). Note that it is not expected, of course, that the grade a student received in the course would *cause* the gain from pretest to posttest, but one might think it would be correlated (and thus register as a predictor) since the score on the posttest was a part of the student's final grade. The fact that it is not a predictor suggests that students' abilities in other aspects of the course—paper writing, quizzes on readings, etc.—were not correlated with their argument analysis ability. This suggests an open area to investigate how argument mapping skills can be made relevant to these other sorts of skills.

clear which are the premises and which is the conclusion and how they are connected, and there is little ambiguity in deciding on what bases to evaluate the argument.

There are several software programs with can facilitate any kind of mapping, and a few that are designed for argument mapping in particular. To date, the only studies published on the efficacy of argument mapping have been in the context of a specific computer program used by the instructor and the students throughout the semester. As cognitive tutors[14] in general have been shown to greatly improve students' problem-solving abilities, one may wonder whether it is the computer environment which is contributing the most to the successes of these students.

In our course, instructors and students are free to use any kind of medium they prefer to build argument maps. Some instructors use computer software, while others use the chalk board or overhead slides; similarly some students use one of a variety of drawing programs, while others use just pencils and paper. Our results show that the argument mapping skill, no matter how the maps are produced, is an important part of the gains in argument analysis abilities our students achieve. While on average all of the students in each of the lectures improved their abilities on these tasks over the course of the semester, the most dramatic improvements were made by the students who were able to construct argument maps. Constructing the correct argument map was highly correlated in general with correctly picking out the premises, deciding how these premises are related to each other and the conclusion, and choosing the grounds on which to evaluate the argument.

Our studies also point to future directions for this research. While it is clear that the ability to construct argument diagrams significantly improves a student's argument analysis skills, it would be interesting to consider whether there are other skills that may usefully be labeled "critical thinking" that this ability may help to improve.

In addition, the arguments we used in testing our students were necessarily short and relatively simple. We would like to know what the effect of knowing how to construct an argument diagram would be on a student's ability to analyze longer and more complex arguments. We suspect that the longer and more complex the argument, the more argument diagramming would help.

It also seems to be the case that it is difficult for students to reason well about arguments in which they have a passionate belief in the truth or falsity of the conclusion (for religious, social, or any number of other reasons). We would like to know whether the ability to construct argument diagrams aids reasoning about these kinds of arguments, and whether the effect is more or less dramatic than the aid this ability offers to reasoning about less personal subjects.

In our classes at Carnegie Mellon University, we use argument diagramming not only to analyze the arguments of the philosophers we study, but also to aid the students with writing their own essays. We believe that, for the same reasons that constructing these diagrams helps students visually represent and thus understand better the structure of arguments they read, this would help the students understand, evaluate, and modify the structure of the arguments in their own essays better. We would like to know whether the ability to construct arguments actually does aid students' essay writing in these ways.

Lastly, unlike the relatively solitary activities in which students engage in our philosophy courses—like doing homework and writing essays—there are many venues in and out of the classroom in which students may engage in the analysis and evaluation of arguments in a group

[14] A cognitive tutor is a computer program that tracks a student's progress and provides individualized, real-time tutoring as she solves a problem in a computer-based environment.

setting. These may include anything from classroom discussion of a particular author or topic, to group deliberations about for whom to vote or what public policy to implement. In any of these situations it seems as though it would be advantageous for all members of the group to be able to visually represent the structure of the arguments being considered. We would like to know whether knowing how to construct argument diagrams would aid groups in these situations.

REFERENCES

Armbruster, B.B., & Anderson, T.H. (1982). Mapping: Representing informative text graphically. In Holley, C.D. & Dansereau, D.F. (Eds.), *Spatial learning strategies*. New York: Academic Press.

Dansereau, D.F., Collins, K.W., McDonald, B.A., Holley, C.D., Garland, J., Diekhoff, G., & Evans, S.H. (1979). Development and evaluation of a learning strategy program. *Journal of Educational Psychology*, 71: 64-73.

Descartes, René. (1641). *Meditations on First Philosophy*. Edited and translated by E. Haldane and G.R.T. Ross in *The Philosophical Works of Descartes, Vol. 1*. Cambridge University Press, 1969.

Harrell, Maralee. (2005). Using Argument Mapping Software in the Classroom. *Teaching Philosophy*, 28: 163-177.

Ennis, R.H., (1987), A Taxonomy of Critical Thinking Dispositions and Abilities. In J.B. Baron & R.J. Sternberg (Eds.), *Teaching Thinking Skills: Theory and Practice*, New York: Freeman. Pp. 1-26.

Johnson, D.D., Pittleman, S.D., & Heimlich, J.E. (1986). Semantic mapping. *Reading Teacher*, 39, 778-783.

Kirschner, P.A., Shum, S.J.B., & Carr, C.S. (Eds.). (2003). *Visualizing argumentation: Software tools for collaborative and educational sense-making*. New York: Springer.

Kuhn, D. (1991). *The skills of argument*. Cambridge: Cambridge University Press.

Larkin, J.H., & Simon, H.A. (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science*, 11, 65-99.

Lee, S.P., (2002), *What Is the Argument? Critical Thinking in the Real World*, New York: McGraw Hill.

Moore, B.N., & Parker, R., (2007), *Critical Thinking, 8th Ed.*, New York: McGraw Hill.

Novak, J.D., & Gowin, D.B. (1984). *Learning how to learn*. New York: Cambridge University Press.

Schwartz, R.M., & Raphael, T.E. (1985). Concept of definition: A key to improving students' vocabulary. *Reading Teacher*, 39: 198-205.

Twardy, C.R. (2004). Argument Maps Improve Critical Thinking. *Teaching Philosophy*, 27, 95-116.

van Gelder, Tim. (2001). "How to improve critical thinking using educational technology." In G. Kennedy, M. Keppell, C. McNaught, & T. Petrovic (Eeds.), *Meeting at the crossroads: proceedings of the 18th annual conference of the Australian Society for computers in learning in tertiary education* (pp. 539-548). Melbourne: Biomedical Multimedia Uni, The University of Melbourne.

van Gelder, Tim. (2003). "Enhancing deliberation through computer supported visualization." In P.A. Kirschner, S.J.B. Shum, & C.S. Carr (Eds.), *Visualizing argumentation: Software tools for collaborative and educational sense-making* (pp. 97-115). New York: Springer.

Winn, W. (1991). Learning from maps and diagrams. *Educational Psychology Review*, 3: 211-247.

Appendix A

80-100 Fall 2004 Pre-Test

Consider the following arguments. For each argument:
(a) Identify the conclusion (thesis) of the argument.
(b) Identify the premises (reasons) given to support the conclusion. Restate the premises in the space provided below.
(c) Indicate how the premises are related. In particular, indicate whether they
    (A) are each separate reasons to believe the conclusion,
    (B) must be combined in order to provide support for the conclusion, and/or
    (C) are related in a chain, with one premise being a reason to believe another.
(d) Provide a visual, graphical, schematic, or outlined representation of the argument.
(e) State whether it is a good argument, and explain why it is either good or bad. If it is a bad argument, state what needs to be changed to make it good.

1. Since major historical events cannot be repeated, historians are not scientists. After all, the scientific method necessarily involves events (called "experiments") that can be repeated.
(a) Conclusion:
(b) Premises:
(c) Relationship of the premises. Circle all that apply: (A) (B) (C)
(d) Visual, graphical, schematic, or outlined representation of the argument:
(e) Good or bad argument? Why?

2. The scientific method does not necessarily involve experimentation. For, if anything is a science, astronomy is. But the great cosmic events observed by astronomers cannot be repeated. And, of course, an experiment is, by definition, a repeatable event.
(a) Conclusion:
(b) Premises:
(c) Relationship of the premises. Circle all that apply: (A) (B) (C)
(d) Visual, graphical, schematic, or outlined representation of the argument:
(e) Good or bad argument? Why?

3. Although Americans like to think they have interfered with other countries only to defend the downtrodden and helpless, there are undeniably aggressive episodes in American history. For example, the United States to Texas from Mexico by force. The United States seized Hawaii, Puerto Rico, and Guam. And in the first third of the 20^{th} century, the United States intervened militarily in all of the following countries without being invited to do so: Cuba, Nicaragua, Guatemala, the Dominican Republic, Haiti, and Honduras.
(a) Conclusion:
(b) Premises:
(c) Relationship of the premises. Circle all that apply: (A) (B) (C)
(d) Visual, graphical, schematic, or outlined representation of the argument:
(e) Good or bad argument? Why?

4. Politicians are forever attributing crime rates to policies—if the crime rates are decreasing, to their own policies; if the crime rates are increasing, to the "failed" policies of their opponents. But the fact is that crime rates are best explained in terms of demographics. For crime is primarily a young man's game. Whenever there is a relatively large number of young men between the ages of 15 and 30, the crime rates are high. And whenever this part of the population is relatively small, the crime rates are relatively low.

(a) Conclusion:

(b) Premises:

(c) Relationship of the premises. Circle all that apply: (A) (B) (C)

(d) Visual, graphical, schematic, or outlined representation of the argument:

(e) Good or bad argument? Why?

5. Small commercial fishing operations will continue to flourish only if restrictions on sport fishing are imposed. But the sport fishing lobby is powerful and vocal, for it is the sport of the rich and famous. And the sport fishing lobby does not want any restrictions. Consequently, restrictions on sport fishing activities are not likely in the near future. And, therefore, the small commercial fisherman is in big trouble.

(a) Conclusion:

(b) Premises:

(c) Relationship of the premises. Circle all that apply: (A) (B) (C)

(d) Visual, graphical, schematic, or outlined representation of the argument:

(e) Good or bad argument? Why?

Appendix B

80-100 Fall 2004 Final Exam

Consider the following arguments. For each argument:
(a) Identify the conclusion (thesis) of the argument.
(b) Identify the premises (reasons) given to support the conclusion. Restate the premises in the space provided below.
(c) Indicate how the premises are related. In particular, indicate whether they
    (A) are each separate reasons to believe the conclusion,
    (B) must be combined in order to provide support for the conclusion, and/or
    (C) are related in a chain, with one premise being a reason to believe another.
(d) Provide a visual, graphical, schematic, or outlined representation of the argument.
(e) State whether it is a good argument, and explain why it is either good or bad. If it is a bad argument, state what needs to be changed to make it good.

1. No physical object can travel faster than light. A Hydrogen atom is a physical object, so no hydrogen atom can travel faster than the speed of light.
(a) Conclusion:
(b) Premises:
(c) Relationship of the premises. Circle all that apply: (A) (B) (C)
(d) Visual, graphical, schematic, or outlined representation of the argument:
(e) Good or bad argument? Why?

2. All brain events are physical events, and no physical events can be adequately accounted for in intensional terms, but it is only in terms of intensions that mental states can be adequately described. So, mental states cannot be brain events.
(a) Conclusion:
(b) Premises:
(c) Relationship of the premises. Circle all that apply: (A) (B) (C)
(d) Visual, graphical, schematic, or outlined representation of the argument:
(e) Good or bad argument? Why?

3. John and Robert Kennedy and Martin Luther King, Jr. were, like them or not, this country's last true national leaders. None of John Kennedy's successors in the White House has enjoyed the consensus he built, and everyone of them ran into trouble, of his own making, while in office. In the same way, none of this country's national spokespeople since Robert Kennedy and Dr. King has had the attention and respect they enjoyed.
(a) Conclusion:
(b) Premises:
(c) Relationship of the premises. Circle all that apply: (A) (B) (C)
(d) Visual, graphical, schematic, or outlined representation of the argument:
(e) Good or bad argument? Why?

4. The power set of any set (i.e. the set of all subsets of a given set) must be larger than the original set. The universal set is, by definition, the set of everything. Consequently, the universal set must not be possible, since its power set would have to contain more members than there are things in the universe.
(a) Conclusion:
(b) Premises:
(c) Relationship of the premises. Circle all that apply: (A) (B) (C)
(d) Visual, graphical, schematic, or outlined representation of the argument:
(e) Good or bad argument? Why?

5. Obviously, there is an objective moral law, for every sane person will agree that it is immoral to kill people at will. However, there is an objective moral law only if there is a moral Lawgiver who exists independently of human thinking. Hence, there is a moral Lawgiver who exists independently of human thinking. But God exists if there is a moral Lawgiver who exists independently of human thinking. Accordingly, God exists.
(a) Conclusion:
(b) Premises:
(c) Relationship of the premises. Circle all that apply: (A) (B) (C)
(d) Visual, graphical, schematic, or outlined representation of the argument:
(e) Good or bad argument? Why?

Appendix C

Table A: The distribution of home colleges in each lecture in both Spring 2004 and Fall 2004

Lecture	CFA	CIT	CMU	HSS	MCS	SCS	TSB
Spring 2004 Total	5	40	7	48	12	15	12
Lecture 1	2	10	2	12	5	3	1
Lecture 2	2	5	3	8	4	6	7
Lecture 3	0	13	1	12	1	3	2
Lecture 4	1	12	1	16	2	3	2
Fall 2004 Total	3	37	6	44	18	9	13
Lecture 1	1	13	1	5	0	1	3
Lecture 2	0	6	1	20	3	5	1
Lecture 3	0	7	0	8	4	2	5
Lecture 4	2	5	2	4	7	0	1
Lecture 5	0	6	2	7	4	1	3

Table B: The distribution of instructors, students, men and women in each lecture in both Spring 2004 and Fall 2004

Lecture	Instructor	No. of Students	No. of Women	No. of Men
Spring 2004	*(totals)*	*139*	*46*	*93*
Lecture 1	Lecturer 1	35	13	22
Lecture 2	Lecturer 2	37	18	19
Lecture 3	Lecturer 3	32	10	22
Lecture 4	Lecturer 4	35	5	30
Fall 2004	*(totals)*	*130*	*36*	*92*
Lecture 1	Lecturer 1	24	6	18
Lecture 2	Lecturer 2	36	6	30
Lecture 3	Lecturer 4	26	9	15
Lecture 4	Lecturer 5	21	7	14
Lecture 5	Lecturer 6	23	8	15