HOW TO IMPROVE CRITICAL THINKING USING EDUCATIONAL TECHNOLOGY

Tim van Gelder
Department of Philosophy
The University of Melbourne, Australia
tgelder@ariel.unimelb.edu.au

Abstract
Critical thinking is highly valued but difficult to teach effectively. The Reason! project at the University of Melbourne has developed the Reason!Able software as part of a general method aimed at enhancing critical thinking skills. Students using Reason!Able appear to make dramatic gains. This paper describes the challenge involved, the theoretical basis of the Reason! project, the Reason!Able software, and results of intensive evaluation of the Reason! approach.

Keywords
critical thinking, reasoning, argument, practice, transfer, argument maps, evaluation, pre and post-testing

The Problem

Critical thinking is one of education's most central goals and one of its most valued outcomes. Mission statements routinely aspire to cultivate critical thinking as a core intellectual virtue. And so they should. Fundamentally, critical thinking is just exercising the general forms of thought most conducive to sorting the true from the false – or, more bluntly, "the art of being right." Thus critical thinking is, as the Enlightenment philosophers understood, central not only to intellectual progress but to all forms of social progress as well.

Unfortunately, critical thinking is in short supply. This is apparent from informal observation. As Einstein put it: "Only two things are infinite, the universe and human stupidity, and I'm not sure about the former." Einstein's despair has been supported by detailed empirical studies. In her landmark book *The skills of argument* (1991), psychologist Deanna Kuhn reported intensively studying hundreds of people from all walks of life. She found that over half the population cannot reliably exhibit even the most basic skills of general reasoning and argument. More recently, a DETYA-commissioned study concluded that "employers *value* this skill and *can* find it, but it is *rare*" (Nielson, 2000).

If something is both valuable and rare then we should try to produce more of it. In the case of critical thinking, this is rather difficult to do. There are two standard strategies. One, the indirect strategy, is to expect that students will just pick up critical thinking skills in the course of their studies. Thus the Arts Faculty at the University of Melbourne, which has been referring to itself as "The Critical Faculty," tries to cultivate critical thinking by requiring its students to study politics or English or history. In general, the indirect approach does work, but only slowly. According to the leading specialist in this area, Ernest Pascarella,

> Our best estimate is that the first three years of college [i.e., undergraduate studies at university] provide an improvement in critical thinking skills of about .55 of a standard deviation or 20 percentile points. (Pascarella, in progress).

On a semester-by-semester basis, gains of this magnitude are pretty slight. Consequently there have been many attempts to accelerate the process by teaching critical thinking directly. Most notably, there are the subjects taught every year in philosophy departments going by names such as "Critical Thinking" or "Introductory Logic." Here critical thinking is the content as well as the intended outcome, and one would naturally expect to see hefty gains. Insofar as these subjects have been empirically evaluated, however, they have been equally disappointing (van Gelder, 2000). According to Pascarella,

> Evidence suggests that critical thinking can be taught, although the average effect is based on a rough estimate and is quite modest in magnitude. Students who receive purposeful instruction and practice in critical thinking and/or problem solving skills appear, on average, to gain an advantage in critical thinking skills of .23 of a standard deviation (9 percentile points) over students not receiving such instruction. (Pascarella, in progress).

Average gains of this amount are difficult, if not impossible, for ordinary people to sort out from statistical noise. Many teachers of critical thinking suspect that their efforts are making little if any difference. Doug Walton, a leading informal logician and teacher of critical thinking, wrote recently:

> I wish I could say that I had a method or technique that has proved successful. But I do not, and from what I can see, especially by looking at the abundance of textbooks on critical thinking, I don't think anyone else has solved this problem either. (Walton, 2000).

The Quality Practice Hypothesis and the Need for Computers

Quality Practice

So critical thinking is valuable, rare, and hard to teach. Recognising this problem, the Reason! project at the University of Melbourne has been developing a whole new approach. Rather than tinkering with existing methods, we are building afresh from solid foundations in cognitive science. Reviewing the research literature on cognitive skill acquisition, we found a consensus around the unsurprising idea that cognitive skills, like other skills, improve with practice. However, it must be the right kind of practice. Generically, the practice should be

- *Motivated* – the student should be deliberately practising in order to improve skills
- *Guided* – the student should have some way of knowing what to do next
- *Scaffolded* – particularly in early stages, there should be structures preventing inappropriate activity
- *Graduated* – tasks should gradually increase in complexity
- *Feedback* – the student should have some way of telling whether a particular activity was successful or appropriate

(See Chase & Simon, 1973; Dreyfus & Dreyfus, 1986; Ericsson & Charness, 1994, Ericsson & Lehmann, 1996; Voss, Wiley, & Carretero, 1995). Further, the practice should be suitably *general* in nature. This is a subtle matter. The greatest challenge facing attempts to inculcate general thinking skills is the problem of transfer: skills acquired in one domain or context often do not carry over to other situations. For example, training in formal reasoning (e.g., chess, mathematics, formal logic) yields little benefit outside those domains. To beat the problem of transfer, critical thinking practice must be not merely practice in some situation or other, and not merely practice in a wide variety of situations, but practice in *transfer itself*. That is, students must practice transferring concepts and skills initially experienced in one situation over to a wide variety of other contexts or domains.

We use the term "quality practice" for practice-for-transfer with the generic properties listed above. The "quality practice hypothesis" (QPH) is the conjecture that critical thinking skills improve substantially through, and only through, extensive quality practice.

Situated Cognition?

It is worth mentioning that the QPH directly conflicts with the "situated cognition" perspective on cognitive skills. At least in its more extreme incarnations, this school of thought maintains that there are really no *general* critical thinking skills, and it is pointless to try to teach such skills in the classroom. Rather, all thinking is tied to particular concrete situations; learning doesn't transfer to remote contexts; training on abstract principles is of little use; and learning should take place in complex social environments (see, e.g., Lave, 1988).

This perspective is supported by various kinds of evidence, including the disappointing results of attempts to teach critical thinking (see previous section). Yet it stands soundly refuted insofar as there are unambiguous examples of people acquiring and exhibiting general critical thinking skills. Kuhn (1991) points out that at least some people (in her study, it was graduate students in philosophy) do show flawless general reasoning skills. Further, our work (see below) shows that general critical thinking skills *can* be substantially enhanced in the classroom. The extreme situated cognition position is therefore false. This does not mean there was nothing of value in the broadly "situated" camp; indeed, it offers many valuable insights into the challenges involved in improving general critical thinking skills.

Computers to the Rescue

The QPH, if true, poses an enormous problem. Quality practice appears to require close supervision from an expert coach to help provide the motivation, guidance, scaffolding and feedback students require. However we can't afford to provide such coaching, except at the Australian Institute of Sport. Realistically, the most we can expect is one teacher – who may not be an expert – for every 20-30 students. This is not nearly enough. It thus appears that if the QPH is correct, we will never be able to improve critical thinking substantially on a large scale.

We can think of only one practical way to attack this problem. It is to use computers to help the student, and to help the teacher help the student. It will be impossible, for the foreseeable future, to replace the expert human coach with a computer program. However computers can assume some of the burden of guidance, scaffolding and feedback, so that student practice activities are better quality and the teacher's input has greater effect.

The Reason!Able Practice Environment

Over the past four years we have been developing the Reason!Able software as an environment in which students can do a better grade of practice and thereby improve their reasoning. Reason!Able is a stand-alone Windows program in which students are guided and scaffolded through the complex processes involved in representing and evaluating reasoning on any topic.

One key innovation in Reason!Able is that all reasoning activities are conducted in specially-designed flowchart-like diagrams called argument maps or trees. Reasoning is essentially a matter of working out the logical or evidential relationships among claims. An argument map is a graphical presentation of reasoning in which the relationships among claims are made visually explicit. Figure 1 shows Reason!Able displaying an argument map of a relatively simple piece of reasoning on the topic of fossil fuel use. The various claims are entered in boxes placed on the workspace. The relationships between the claims are signified by three graphical properties:

- **Lines**. A line between two boxes indicates that one is evidence for or against the other
- **Relative position** in 2D space. If one claim is evidence for or against another, they will be appropriately juxtaposed.
- **Colour**. Green indicates a reason (supporting evidence); red indicates an objection (opposing evidence).

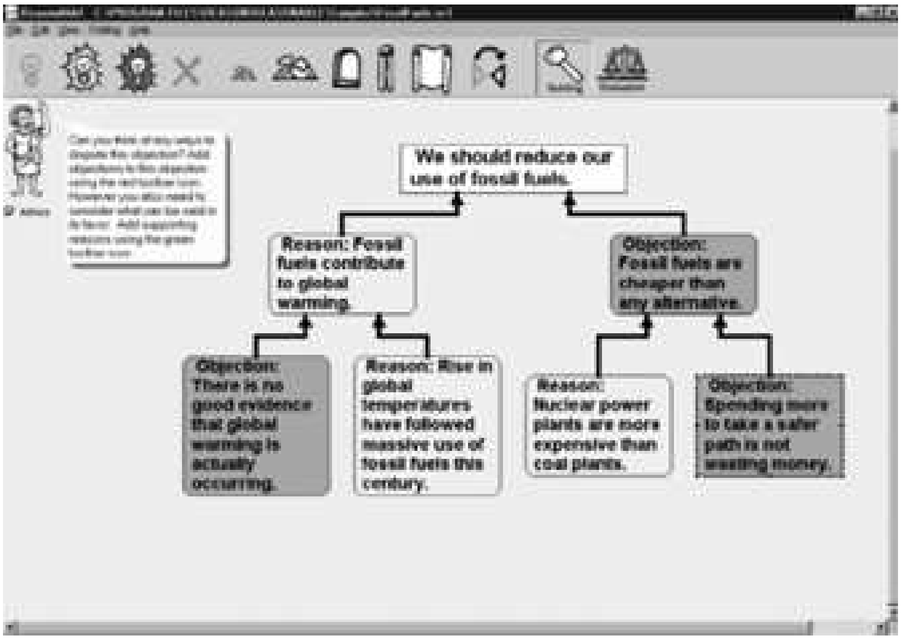

Figure 1: Reason!Able 1.03b screenshot showing an argument map of some simple reasoning on the topic of fossil fuel use. "Socrates" is offering advice in the upper left hand corner.

Using these conventions, a Reason!Able argument tree presents a structure of reasoning in a completely explicit and unambiguous fashion.

Reason!Able argument trees are not static and pre-given. Rather, they are built up by the student herself, who types text into claim boxes and places the claims on the argument tree. The tree structure is manipulable to allow easy reformulating of the reasoning. Thus claims and even whole branches can be torn off the argument tree, or added on at any point, or dragged and dropped to another location.

Reason!Able is designed to be used by novices who have had no prior instruction in the general principles of reasoning and argument. The goal is to help such people gradually acquire the relevant concepts, procedures and skills. Thus there is ever-present context-sensitive guidance. Whenever the student clicks anywhere on the argument tree, the "Socrates" character prompts the student to think about issues pertinent at that point. In Figure 1, the student has clicked on an objection to an objection (i.e., a rebuttal). Socrates is asking the student to consider whether she can dispute the objection – or alternatively, whether she can back it up with evidence.

Developing arguments is an important aspect of critical thinking. Another equally important aspect is evaluating, i.e., determining whether a given argument – particularly an argument you have just developed yourself – is any good. Reason!Able also guides and scaffolds the student through the processes involved in evaluation.

Thus the program has two basic modes of operation: "Building" mode, in which reasoning is articulated, and "Evaluation" mode, in which arguments are assessed for quality. Students switch back and forth between the two modes at any time, using the hammer (build) and scales (evaluate) icons. In evaluate mode, clicking once on any node will cause Socrates to proffer pertinent advice. Double-clicking brings up a dialog allowing the student to enter her considered evaluative judgements. Information about how to make such judgements is available by clicking on the Help button on the dialog. This calls up the relevant page of an extensive Help section, which amounts to an in-built mini-textbook on the theory and practice of informal reasoning and argument.

This brief overview has presented only the most obvious and simple aspects of the software. Other features include:

- Reasons and objections can be "unfolded" to reveal the internal structure of premises (main premise and co-premises or assumptions), with reasons or objections targeted upon specific premises.
- Individual premises can be evaluated for truth, in terms of the relevant reasons or objections, as well as basic "grounds" such as common knowledge.
- The program supports various viewing modes to help users comprehend of complex arguments.

Figure 2: Reason! project members using Reason!Able. The image is projected on a "smartboard" – a touch-sensitive whiteboard allowing many operations to be performed by using a finger instead of the computer's mouse or keyboard. In this image one team member is about to drag a reason box to another location. Using Reason!Able with this technology helps make abstract arguments into concrete, manipulable structures.

Quality Practice in the Reason!Able Environment

Reason!Able is designed to support quality practice. Students practice two main kinds of critical thinking activities using the package. *Critical evaluation* is identifying and then evaluating the reasoning provided by somebody else in some text (e.g., an opinion piece, journal article or book). In *argument production*, students develop their own reasoning to support their own positions. Both activities involve putting together an argument tree representing the reasoning, and then systematically assessing that reasoning. These procedures are much more complicated than most people realise, with an intricate recursive structure.

Reason!Able helps improve the quality of student practice by comparison with ordinary practice activities such as classroom discussions or essay writing. It does this in a number of ways.

1. **Guidance**. Most obviously, Reason!Able provides continual, context-sensitive guidance in the form of promptings from Socrates. Students can "ask" for instructions simply by clicking at any time at any point on an argument tree. Obviously Socrates' "canned" guidance is never going to be as good as advice from a genuine human expert. However it is better than nothing, and may well be better than attempted advice from a human non-expert. And over time we hope to gradually introduce AI techniques to help Socrates become a bit more intelligent. Note also that with experience, students find they can anticipate what Socrates will say and find his always popping up a bit irritating, so they switch off his advice. That's the whole idea; students are then providing their own guidance because they understand roughly what to do.

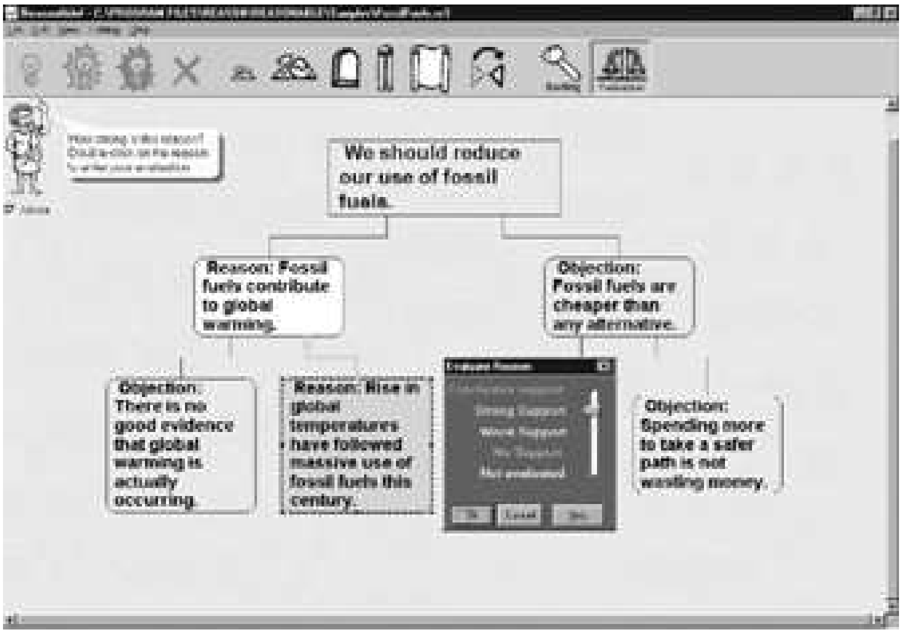

Figure 3. Reason!Able in "evaluate" mode. Students are prompted to enter their own judgements as to the quality of the reasoning. In this case the user is evaluating one reason as providing strong support (bright green) for another (white). Evaluations should be, but might not be, correct or appropriate. Note that all other parts of the argument tree have receded into the background since they are not relevant to the quality of the selected reason The result, when all steps in the process have been completed, is a simple visual representation of a developed argument plus the student's evaluation of the quality of the reasoning.

2. **Scaffolding**. Just as importantly, Reason!Able provides very strict scaffolding for students' thinking. Using Reason!Able, students can *only* create "syntactically" correct argument maps. They rapidly learn that an argument has a main conclusion, and a hierarchical structure of reasons & objections, themselves made up of claims, and so forth. This scaffolding puts strong structural constraints on the way they think about issues and formulate their reasoning. This is a key advantage of Reason!Able over more flexible or free-form frameworks for representing reasoning such as prose or generic diagramming packages such as Inspiration. An added advantage of this approach is that all students in a group are scaffolded to have the *same* understanding of the structure of reasoning, making for more productive communication.

3. **Feedback**. Reason!Able on its own does not provide any significant feedback on the user's activities. Basically, this would require the program to be intelligent enough to understand language and reasoning, and this level of sophistication is currently far beyond the frontiers of computer science. Crucially, however, Reason!Able does provide a context in which students and teachers can generate better feedback than would otherwise be possible. For example, in critical evaluation tasks students can develop their own answers and then compare them with models. Since Reason!Able argument maps represent information in a simple and accessible fashion, students can directly compare their treatments with the recommended answers. By noting differences, and assuming the model answer is correct, students are generating their own positive feedback (when they are in alignment) and negative feedback (when there is a difference). Similarly, when teachers examine students' argument maps, they can see at a glance what students are thinking, and provide more rapid and targeted feedback than is possible in traditional forms of interaction (such as commenting on a piece of written work).

4. **Motivation**. Reason!Able appears to help improve enthusiasm for practice. Many students find Reason!Able's simple, colourful diagrams more appealing than dry prose. Also, the

guidance, scaffolding and improved feedback help them feel that they are making more concrete progress than in more flexible environments, where seems that "anything goes" and students can feel that they are floundering.

The remaining two dimensions of quality practice are (a) graduated exercises, and (b) practice-for-transfer. Reason!Able alone does not enhance these dimensions. It is up to the teacher to provide a suitably graduated and diverse exercise regime.

Does the Reason! Approach Work?

From the outset we have been aiming to develop an approach to critical thinking instruction that *demonstrably* produces *substantial* gains. (A further constraint is that the approach be practical and affordable for widespread use.) The new approach is the Reason! method, which essentially just implements the QPH with Reason!Able as the practice environment. In order to determine the extent to which the Reason! method really works, we have been pre- and post-testing students using the method. This also gives us indirect insight into whether using Reason!Able actually helps students improve their critical thinking.

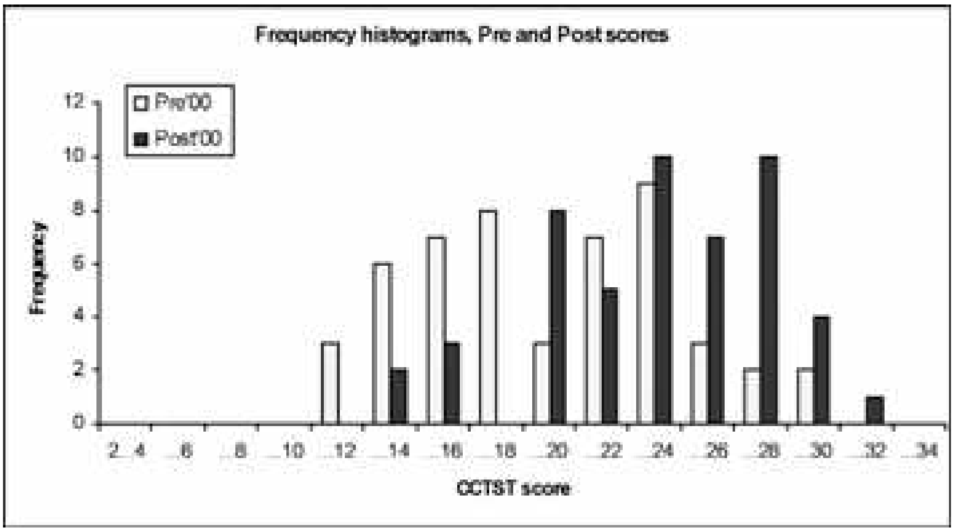

Figure 4: Gains in critical thinking over 12 weeks in 2000 using the Reason! approach, as measured with the California Critical Thinking Skills Test. The distribution of scores shows a substantial upward shift from the pre-test (light blue) to the post-test (dark blue).

Evaluation Method

The context is the University of Melbourne undergraduate subject Philosophy 161003 *Critical Thinking: The Art of Reasoning*. In this subject, students are intensively trained in critical thinking using the Reason! method. We estimate gains in critical thinking by pre- and post-testing using two different tests. The first is the California Critical Thinking Skills Test, probably the best objective (multi-choice) test, and increasingly widely used around the world. The other test is a written task of our own devising, loosely based on the Graduate Record Exam's Writing Assessment. Here students are given a short argumentative text, and are asked to identify and evaluate the reasoning. The written test is scored by two critical thinking experts who are quite independent of our team. They grade "blindly" in the sense that we take all possible precautions to conceal whether any given answer is from the pre-testing or a post-testing. They are given some instructions, but we keep the guidance minimal because we want their independent (and possibly quite different) perspectives on whether critical thinking has been improved. Both tests have two forms and we use a crossover design. Using this design, our evaluation processes are, to our knowledge, the most intensive and rigorous ever deployed to evaluate the efficacy of an attempt to improve critical thinking.

Results

Data is available for studies conducted in both 2000 and 2001. The two studies yielded essentially the same results. The CCTST indicated improvement with an effect size of about 0.84 (mean difference between paired pre- and post-test scores, divided by standard deviation on the pre-test). The written test showed effect sizes of a similar order of magnitude, though there were substantial differences and low correlations between scorers. Collectively these results indicate that the effect is quite robust. Crudely speaking, we are seeing a gain in critical thinking of the best part of one SD in a single semester. (For comparison, an equivalent shift in IQ would be about 12-13 points on average across the group.)

Reason! Approach Versus other Approaches

This is, to our knowledge, the largest gain ever recorded anywhere over a comparable period or for a comparable intervention (see van Gelder, 2000 for an overview of all relevant studies). The Reason! approach appears to significantly outpace other "direct" attempts to teach critical thinking (Figure 5), and appears to dramatically accelerate improvement in critical thinking when compared with the "indirect" strategy (i.e., just being at university) (Figure 6).

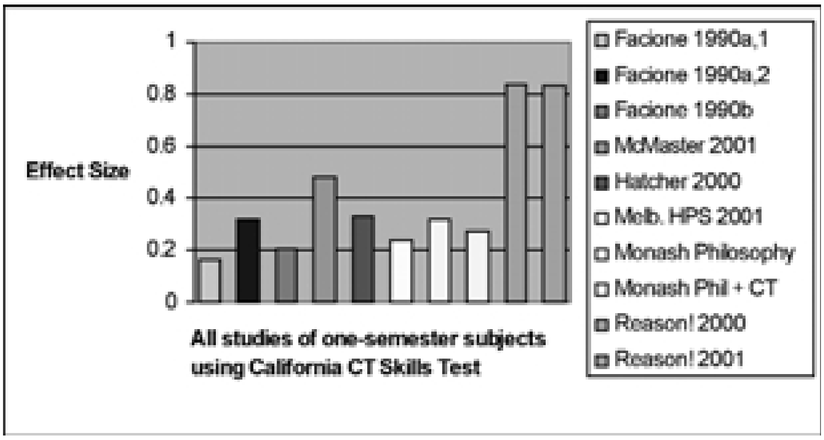

Figure 5. All known CCTST studies of gains in critical thinking in the context of a one-semester subject. This illustrates that the Reason! method is roughly tripling the gross gains of typical one-semester critical thinking subjects. Hatcher (2000) is a two-semester study, pro-rated to one-semester. The next best performance, after the Reason! approach, is McMaster 2000 (unpublished), which was also software-based, though it used a very different package. The latter five studies, including the two Reason! studies, were performed by us and have not yet been published elsewhere.

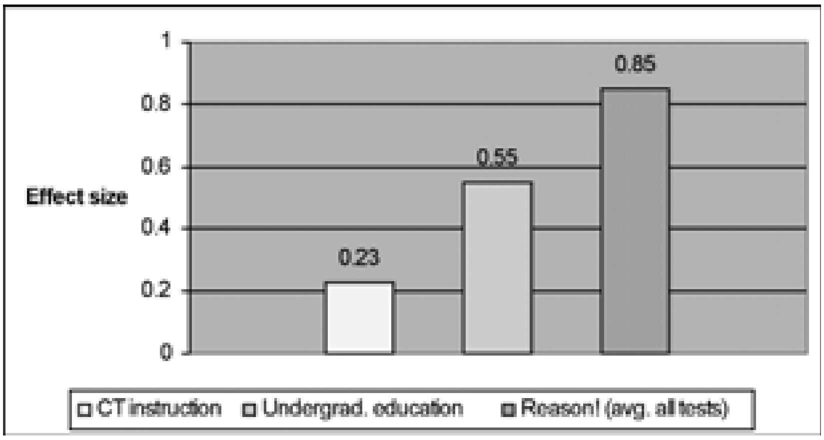

Figure 6. Gain in critical thinking skills, comparing the gain using the Reason! method with standard critical thinking subjects, and the effect of being at university. This comparison suggests that the Reason! method is about three times as effective as standard critical thinking courses, and more effective than three years of undergraduate education.

Effect of Reason!Able

None of this establishes that using the Reason!Able software itself helps improve critical thinking skills. Reason!Able is one element of a total package which includes many other factors which may be causing improvement, such as the :

- intensive practice involved in the Reason! method;
- nature of the practice activities involved;
- particular skills of the instructor(s) involved in teaching 161003; and
- "founder's effect" – the unusual level of enthusiasm and energy shown by the instructor(s).

The substantial gains in critical thinking shown by our students *could* be caused wholly or in large part by these other factors, either individually or in combination.

In theory, to accurately gauge the Reason!Able's distinctive contribution to learning, we would need to compare two courses, equal in all respects other than use of Reason!Able, and with randomised assignment of students to one course rather than another. In practice, this kind of experimental situation is very difficult to arrange. However, although we can't accurately say *how much* benefit Reason!Able confers, it is plausible that the software is making a substantial contribution. After all, it was designed in order to make possible and practical the kind of activity which, according to cognitive science, should enhance general critical thinking skills. Second, students in course evaluation questionnaires agree that the software helps them learn.

In current and future studies, we are attempting to eliminate or at least assess the impact of confounding variables. Additionally, we are trying to obtain a higher level of insight into (a) the QPH, and (b) the benefits of using Reason!Able. The software now records extensive data each student's practice activities, and uploads that data to a central repository. We will be able to process that data together with the improvement to see if there are interesting correlations between use of the software and improvement. Certain patterns of correlation would provide strong evidence both that the QPH is true and that Reason!Able use does contribute substantially to critical thinking skill.

References

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*, (pp. 215-281). New York: Academic Press.

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: The Free Press.

Ericsson, K. A., & Charness, N. (1994). Expert performance. *American Psychologist, 49*, 725-747.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47*, 273-305.

Facione, P. A. (1990a). CCTST experimental validation and content validity: CCTST technical report #1. ERIC, ED 327-549. 1-19. Millbrae, CA: The California Academic Press.

Facione, P. A. (1990b). Technical Report #2: Factors predictive of CT skills. ERIC ED 327 550.

Hatcher, Donald L. (2000). Why we should combine critical thinking with written instruction. *Informal Logic*, *19*.

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

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. New York: Cambridge University Press.

Nielson, A. C. (2000). *Employer satisfaction with graduate skills*. Evaluations and Investigations Program report 99/7. Canberra: Higher Education Division Department of Employment, Education, Training and Youth Affairs. [Online]. Available: http://www.detya.gov.au/archive/highered/eippubs/eip99-7/eip99_7pdf.pdf [27 September 2001].

Pascarella, E. (in progress) *How college affects students revisited: Research from the decade of the 1990s*. This is the new edition of the classic work Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco: Jossey-Bass. I am quoting with permission from the manuscript version of Chapter 4, which has been completed.

van Gelder, T. J. (2000). The efficacy of informal reasoning courses. Preprint No. 1/2000, University of Melbourne Department of Philosophy. [Online]. Available: http://www.philosophy.unimelb.edu.au/reason/ [14 September 2001].

van Lehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology, 47*, 513-39.

Voss, J. F., Wiley, J., & Carretero, M. (1995). Acquiring intellectual skills. *Annual Review of Psychology, 46,* 155-81.

Walton, D. (2000). Problems and useful techniques in teaching argumentation, informal logic and critical thinking. *Informal Logic*, *20*, Teaching Supplement, 35-9.

Acknowledgements

The author acknowledges the assistance of Yanna Rider and Geoff Cumming for data collection and analysis, Andy Bulka for programming, and David Hitchcock for sharing the unpublished data from the McMaster study.