Higher Education Research & Development
Vol. 30, No. 3, June 2011, 387–403

Routledge
Taylor & Francis Group

The effectiveness of a single intervention of computer-aided argument mapping in a marketing and a financial accounting subject

Michal Carrington, Richard Chen, Martin Davies*, Jagjit Kaur and Benjamin Neville

Faculty of Business and Economics, University of Melbourne, Melbourne, Australia

(Received 9 July 2009; final version received 26 January 2011)

An argument map visually represents the structure of an argument, outlining its informal logical connections and informing judgments as to its worthiness. Argument mapping can be augmented with dedicated software that aids the mapping process. Empirical evidence suggests that semester-length subjects using argument mapping along with dedicated software can produce remarkable increases in students' critical thinking abilities. Introducing such specialised subjects, however, is often practically and politically difficult. This study ascertains student perceptions of the use of argument mapping in two large, regular, semester-length classes in a Business and Economics Faculty at the University of Melbourne. Unlike the semester-length expert-led trials in prior research, in our study only one expert-led session was conducted at the beginning of the semester and followed by class practice. Survey results conducted at the end of the semester, show that, with reservations, even this minimalist, 'one-shot inoculation' of argument mapping is effective in terms of students' perceptions of improvements in their critical thinking skills.

Keywords: arguments; computer-aided argument mapping; critical thinking; reasoning

1. Introduction

All disciplines, regardless of their academic nature (empirical, non-empirical, practical, theoretical, text-based etc.) require students to demonstrate critical thinking skills. Critical thinking (CT) is promoted in many universities as a key graduate attribute. The capacity to think critically is highly valued, especially among employers (Graduate Careers Australia, 2006). However, CT is a *learned* skill and – like other learned skills (playing the piano, competence in mathematics) – expertise takes considerable effort, time and dedicated practice (Ericsson & Lehmann, 1996).

While promoting critical thinking as a graduate attribute, many universities do not offer students *explicit* means of learning and practising this core competency. It is not surprising, therefore, that Australian employers, for example, report 'capacity for independent and critical thinking … sets apart successful from unsuccessful applicants … but it is *rare*' (Commonwealth of Australia, 2000, emphasis in original). Moreover, recent research suggests that undergraduate students achieve minimal improvement in their CT skills over the duration of their undergraduate degrees (Hitchcock, 2004). To

*Corresponding author. Email: wmdavies@unimelb.edu.au

ISSN 0729-4360 print/ISSN 1469-8366 online

DOI: 10.1080/07294360.2011.559197
http://www.informaworld.com

address these graduate deficiencies, it is suggested that explicit approaches are required to teach CT (Davies, 2009a, 2009b).

Critical thinking requires, in part at least, the ability to understand and assess *arguments.* This can be enhanced with the use of argument maps. Argument maps/mapping (AM) are visual tools enhancing critical analysis and evaluation of arguments. They differ from 'mind maps' and 'concept maps' in terms of level of detail and degree of precision (Davies, forthcoming). Unlike a mind map or a concept map, an argument map visually represents the structure of an argument, outlining the informal logical connections comprising the argument and informing the conclusion. The AM 'method' of clearly outlining a contention at the top of a map, followed by tiers of reasons and objections can be augmented with Computer-Aided Argument Mapping (CAAM) software programs that aid the mapping process. An example argument map from the discipline of Finance is provided in Figure 1. The argument map shows a contention at the top of the 'tree', followed by a first tier reason (in green) and an objection (in red). The first tier, or layer, of the argument is followed by three second-tier reasons and one rebuttal (in orange) to the first tier objection. A third tier objection (in red) is provided to one of the second tier reasons (NB: Argument maps normally appear colour-coded). To complete the argument 'basis boxes' are provided. These provide terminal support for the reasons, objections and rebuttals. This support includes statistical evidence, expert opinion, quotations and citations to web-based literature.

Evidence from the cognitive sciences shows that visual displays enhance learning (Vekiri, 2002; Winn, 1991). Maps allow the encoding of information in memory in visual as well as propositional form, a phenomenon called 'conjoint retention' or 'dual

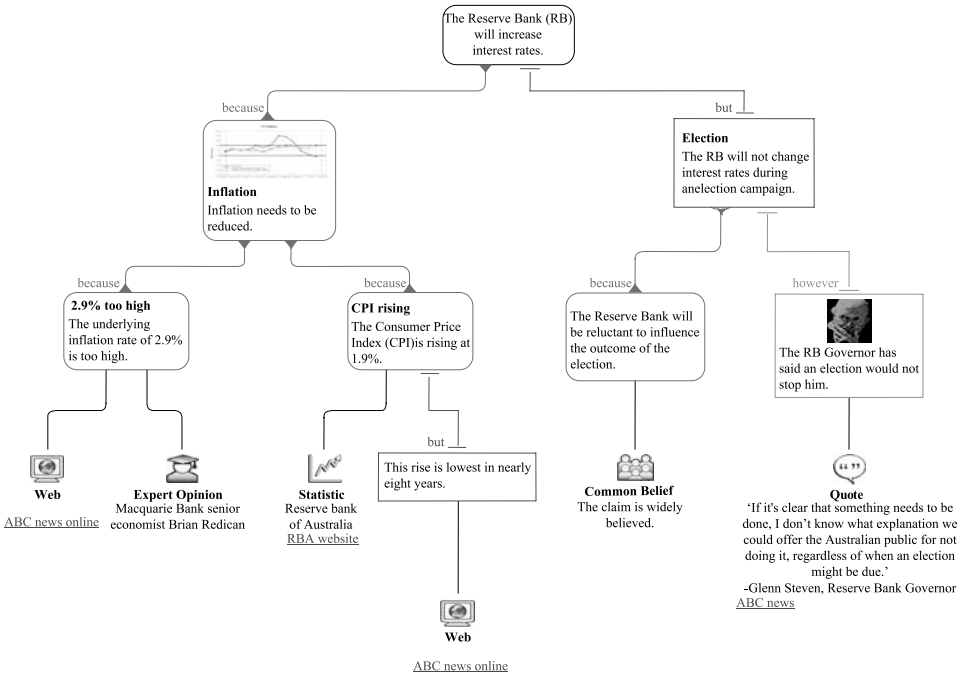

Figure 1. Example of an argument map.
Source: http://www.austhink.com

coding' (Kulhavy, Schwartz, & Caterino, ; Paivio, 1983; Schwatz, 1988). Processing information verbally as well as pictorially helps learning by using more than one modality.

Published studies demonstrate empirical support for the use of CAAM in improving CT skills (for a meta-analysis see Alvarez [2007]). One study showed that one-subject CAAM-based interventions over a 12-week period achieved a 0.8 SD gain as measured by pre- and post-test results in the Californian Critical Thinking Skills Test (CCTST). This is equivalent to the gains usually achieved in three or four years of an undergraduate degree (van Gelder, Bissett, & Cumming, 2004). This CAAM intervention was extensive and involved weekly expert-led lectures on CT and weekly class-based activities involving AM in tutorials. Similar improvements have been found in other studies (*The monash critical thinking study*, 2009; Twardy, 2004). In Twardy's study, the gains were 90% of the gains in the van Gelder et al. study under the same conditions, even after controlling for the 'teacher effect' and substituting an 'inexperienced post-doc', who was unused to AM, to teach the class. In the Monash study, the influence of computer-based feedback in an AM class was trialled over a semester involving weekly 30–40 minute class-based tutorials with regular exercises involving CAAM. They found a CCTST effect size of 0.45 SD and a 14% average improvement in CT skills. Data emerging from these studies are uniform in showing an improvement in skills in using CAAM. However, the effect size differs depending on the nature, type and extent of the trials. Of course, AM may be taught effectively without computers, as found by Harrell (2007). Using software, however, may enhance other aspects of learning, not least of which is student engagement.

An important question is whether AM needs to be introduced in a specialised, semester-length subject in the first years of study or whether it is possible to introduce it within the design of a regular subject. 'Semester-length' in the Australian tertiary context refers to classes conducted over a 12-week period with around three to four hours contact time per week. 'Regular subject' refers to classes undertaken on this basis. While the van Gelder et al. (2004) results indicate the significant benefits of the first option, this is often practically and politically difficult, if not impossible, to institute. Teachers are therefore in a position of 'bolting' AM onto regular subject content and design.

In this paper we trial the teaching effectiveness of AM within two large cohort subjects in the Business and Economics faculty of a large research-intensive university. Our focus was not to replicate the conditions of more extensive, semester-length expert-led trials using AM throughout the semester, rather, we were interested in establishing students' perceptions of whether – given the evidence supporting AM – a one-shot inoculation of AM in a regular class can enhance CT skills given realistic timetabling constraints. We were interested in establishing whether AM would result in improvements in CT based on student self-reports following a single-class intervention in a normal subject stream in two separate classes. This trial involved an initial expert-led session of one hour in duration, followed by subsequent classroom practice by students themselves occurring without expert-led assistance (i.e., it constituted a minimal intervention model). We describe the method used below.

2. Methods

2.1 The initiative

This project involved trialling the integration of AM into the curriculum of two commerce-related disciplines, Accounting and Marketing, both with a large cohort of

students. The two subjects used for the AM intervention were Financial Accounting (FA) and Marketing and Society (MS), both offered in the second semester of 2009. The student cohorts were large, 109 and 182, respectively. The application of AM was largely similar in each subject.

Financial Accounting represents the capstone of student graduate-level studies in financial accounting. The course examines financial accounting from a theoretical perspective so that contemporary accounting issues and practice can be critically analysed and evaluated. It also provides students with the opportunity to develop high-level analytical skills and the ability to authoritatively present arguments for events or issues. This subject differs from their previous financial accounting studies, which are largely numerically-based. Argument mapping was used in this subject trial to guide the students on how to logically structure and map contemporary financial reporting issues.

Similar to FA, MS attracts students in the final stages of their undergraduate marketing studies. Marketing and Society examines the role of marketing in wider society, beyond its relationship with consumers and shareholders. The course looks at the benefits that marketing provides, but also holistically examines the influence of marketing, including its negative impacts upon individuals and society. For many MS students, this subject represents their first experience in critically examining the tenets of the marketing discipline and its practice and evaluating marketing practices using ethical frameworks.

Our intervention began with an expert-led session, 50 minutes in duration, held within the normal timetabled tutorial of the two classes (FA and MS). This class was repeated for different tutorial groups. Each student cohort was divided into smaller tutorial groups of 15–20 students and 20–50 students for MS and FA respectively. These tutorial groups met weekly during semester in a classroom setting to answer tutorial questions with a tutor (in the case of FA, the lecturer was also the tutor). The purpose of tutorials was to reinforce course content delivered in the subject lectures. The AM session consisted of an explanation of AM and an opportunity to use the CAAM software. An outline was first provided of mind mapping and concept mapping and how they differ from AMs. The CAAM software was then demonstrated. Statements were then distinguished from contentions (the latter involving an inference from premise[s]) and a number of simple arguments, consisting of a single reason bearing on a contention, were outlined. This stage involved student participation. This led to a discussion of more complex, multi-strand arguments, with reasons, objections and rebuttals. Students were encouraged to map these more complex examples using the CAAM software. This concluded the class-based intervention.

After the expert-led AM session in the first (FA) and fifth (MS) weeks of the semester, AM was integrated into the FA and MS curricula in two ways:

(1) Weekly tutorial presentations of selected questions using AM: designated tutorial questions were prepared and presented using AM. These weekly presentations gave students an opportunity to prepare their own AM and to discuss AMs provided by others. Each activity provided a different learning experience. Commenting on AMs prepared by others allowed students to build AM skills by identifying the strengths/flaws in others' maps. Preparing their own AM allowed them to test their own understanding of the material. In total, students were required to participate in discussion for 10–11 AMs. In the FA course, the AMs were then submitted to the lecturer for assessment. Constructive feedback was then provided to students to ensure that AM skills were

further enhanced. In MS, tutors provided on-the-spot assessment of the students' AMs and feedback during the tutorial classes. It is important to note that the FA and MS weekly tutorials and AM discussions were led by tutors/lecturers previously inexperienced in AM technique.

(2) A group assignment augmenting written essays with an argument map: the group assignment for both FA and MS was an essay linking complex arguments and requiring students to integrate knowledge from six weeks of classes and to critically defend their own arguments. Each essay was augmented by an AM illustrating the logic of their arguments, citing relevant research.

2.2 Questionnaire design

We examine the effectiveness of AM by analysing data from a self-reported questionnaire in both FA and MS. We also present the initial results of a objective study using the California Critical Thinking Skills Test (CCTST), a well-established test of critical and analytical thinking, in the MS subject. The self-reported questionnaire was identical across both subjects and was conducted in the last two weeks of Semester 2, 2009. It consisted of nine likert-scale questions and three general questions (optional) (see Appendix 1). The nine likert-scale questions adopted a five-point scale, with '1' being 'strongly disagree' and '5' being 'strongly agree', as well as 'not applicable'.

The development of the questionnaire design tested the perceived effectiveness of AM as a tool (Question [Q] 1). Previous research has set out the processes that AM uses to increase or improve CT: visually representing the argument, helping understand the argument, breaking down complex arguments into simple manageable components and building up logical connections between arguments and evidence (Davies, 2009a, 2009b). As each of these components that help in CT skills are conceptually 'concrete singular', as opposed to more complex constructs that require multiple items to cover all facets of the construct (Baumgartner & Homburg, 1996), we followed the C-OAR-SE scale development process (Rossiter, 2002). Concrete singular constructs require single items determined by an expert panel. The four singular items were developed with a panel of two experts from each discipline and one teaching expert to examine their usefulness in the questionnaire (Q2–Q5). Previous research also sets out other benefits of AM (Rider & Thomason, 2008). These benefits are also 'concrete singular' and so, following the same process, Q6–Q7 were set up to test these benefits of AM. Similarly, Q8 was designed to understand whether students found AM useful and would apply it to other areas and Q9 was used to check teaching effectiveness.

Prior to the end of the semester, the questionnaire was distributed in paper form to all students who took FA and MS in Semester 2, 2009. In total, 100 students (out of 109, 91.7% response rate) from the FA group and 123 students (out of 182, 67.6% response rate) from the MS group returned the questionnaire. All students completed the mandatory sections of the questionnaire and all responses were usable. This left a final sample of 223 observations for this study. We consider this sample adequate in statistical terms (see Section 3).

In addition to the questionnaire, we conducted a pilot trial of an objective test of student CT ability using the CCTST. This test is intellectually challenging and onerous, taking approximately 60 minutes to complete. Given the Australian pedagogical context, where students cannot be compelled to complete research tests and course

credit may not be offered, voluntary student completion of the test – which requires a pre-test/post-test design – is likely to be limited. Thus, our pilot of the CCTST in the MS subject resulted in a matched sample of only 25 students completing both the first and second tests (13.7% response rate). Nevertheless, the resulting data emerged as statistically significant, despite the small sample size.

3. Results

We present our analysis separately for FA and MS groups due to the differences in the nature of the subjects. We first use two methods to analyse the students' responses to the nine likert-scale questions: (1) comparing the mean scores and the scale midpoint of 3 and (2) examining the distribution of responses. In addition, we ran a regression analysis on the questionnaire responses to examine the determining factors explaining the overall usefulness of the AM tool. Descriptive comments from the three open questions provided by students were qualitatively analysed and discussed. Finally, to supplement the self-reported responses from the questionnaire, we conducted a pilot CCTST on a sample of MS students.

3.1 Descriptive statistics

The results for the FA group and MS group are reported in Panels A and B of Table 1. The first three columns provide the statistics on the mean, standard deviation and *p*-value drawn from our *t*-test. The last four columns report the distribution of the responses grouped by 'agree' (4 and 5), 'neither agree nor disagree' (3), 'disagree' (1 and 2), and 'NA'. Our *t*-test assesses whether the mean response is statistically different from the scale midpoint of 3. If the mean response is significantly higher (lower) than the midpoint, it indicates that, on average, the students agree (disagree) with the statement. This is a more robust measure of statistical significance than simply testing agree (4 and 5) versus disagree (1 and 2) responses, as it also tests against those who were indifferent (3).

The mean score for Q1 in the FA group is 3.83 and was significantly different from the midpoint of 3 ($p < 0.0001$), indicating that FA students do regard CAAM as an important tool for developing critical thinking. The responses to Q2–5 also scored at least 3.80 and were significantly different to the mid-point ($p < 0.0001$). This provides strong evidence that AM assisted students to visualize (Q2), understand (Q3), break down (Q4) and connect (Q5) arguments. The responses to Q6 and Q7 on essay writing, however, are mixed. Students reported that CAAM did help the students in essay writing (Q6: mean = 3.36, $p = 0.0016$), but the students did not save much time overall (Q7: mean = 3.16, insignificant). Nevertheless, Q8 found that students were planning to apply their CAAM knowledge to other CT areas (Q8: mean = 3.46, $p < 0.0001$). Students were neutral about the difficulties they face in using CAAM due to lack of knowledge (Q9: mean = 2.90, insignificant). Furthermore, distribution of responses demonstrates students' variable experiences using CAAM. While many perceived that CAAM improved their essay writing productivity (Q7) and a lack of knowledge hindered their application of CAAM (Q9), many also did not, resulting in insignificant results. The distribution of responses for Q9, however, is perhaps most telling: 45% of the FA students reported that the knowledge gained from the single-intervention of CAAM was sufficient for them to complete the AM tasks. A minority (36%) reported that the single-intervention was insufficient.

Table 1. Descriptive statistics.

Panel A. Mean score and response distribution for Financial Accounting group (n = 100).

Q	Mean	SD	p-value	%Agree (4 and 5)	%Neutral (3)	%Disagree (1 and 2)	%NA
Q1 Overall effectiveness	3.83	1.10	< 0.0001	68	19	13	0
Q2 Visual representation	3.87	1.03	< 0.0001	72	18	10	0
Q3 Understanding	3.81	0.97	< 0.0001	68	25	7	0
Q4 Component break down	3.80	1.07	< 0.0001	63	27	10	0
Q5 Logical connections	3.83	0.92	< 0.0001	71	19	10	0
Q6 Help essay writing	3.36	1.09	0.0016	42	40	16	2
Q7 Time saving in writing	3.16	1.12	0.1551	37	36	26	1
Q8 Apply to other areas	3.46	1.16	0.0001	53	27	20	0
Q9 Face difficulties	2.90	1.28	0.4348	36	19	45	0

Panel B. Mean score and response distribution for Marketing and Society group (n = 123).

Q	Mean	SD	p-value	%Agree (4 and 5)	%Neutral (3)	%Disagree (1 and 2)	%NA
Q1 Overall effectiveness	3.55	1.05	< 0.0001	59	24	18	0
Q2 Visual representation	3.85	0.98	< 0.0001	72	17	11	0
Q3 Understanding	3.46	1.10	< 0.0001	54	24	22	0
Q4 Component break down	3.64	0.94	< 0.0001	61	27	12	0
Q5 Logical connections	3.54	1.00	< 0.0001	59	24	16	0
Q6 Help essay writing	3.08	1.18	0.4427	37	31	32	1
Q7 Time saving in writing	2.76	1.12	0.0171	28	28	43	0
Q8 Apply to other areas	3.01	1.11	0.8711	38	27	35	0
Q9 Face difficulties	2.44	1.08	< 0.0001	18	25	56	1

Questionnaire results for the MS group are outlined in Panel B. Similar to the FA student cohort, the mean scores suggest that MS students viewed CAAM overall as an effective CT tool (Q1: mean = 3.55, $p < 0.0001$) and AM was effective in assisting them to visualise (Q2: mean = 3.85, $p < 0.0001$), understand (Q3: mean = 3.46, $p < 0.0001$), break down (Q4: mean = 3.64, $p < 0.0001$) and connect (Q5: mean = 3.54, $p < 0.0001$) arguments. The results for MS, however, were not as strong as those for FA. These results are further supported by the distribution of responses, which demonstrates greater variability to that of the FA students. The responses to Q6–8, however, showed some differences to the FA cohort. Unlike the FA students, the MS students did not perceive that CAAM assisted their essay writing (Q6: mean = 3.08, insignificant), disagreed that CAAM assisted their essay writing productivity (Q7: mean = 2.76, $p < 0.0171$) and probably would not apply it to other CT areas (Q8: mean = 3.01, insignificant). Finally, MS students disagreed that they had difficulties with AM due to lack of knowledge (Q9: mean = 2.44, $p < 0.0001$). Indeed, only 18% reported that the single-intervention was insufficient in providing them with knowledge to complete the set AM tasks.

In summary, these results show that both the FA and MS students' perceived that the single-intervention of CAAM was effective in improving their CT skills. The

results for essay writing, however, were mixed. The differences in results between FA and MS could be attributed to the additional four weeks of practice for the FA students and because a greater proportion of the FA students were from non-English speaking backgrounds (NESB). Indeed, 95% of the FA students were international students, compared with a majority of local students in MS. Students from NESB may have found AM more useful for essay writing. For international students, understanding the course material in English can be difficult (Evans, Tindale, Cable, & Hamil Mead, 2009). Essay requirements during each semester comprised a minimum of 20% across the FA and MS and, in terms of assessment, often posed the greatest difficulty for international students. In addition, feedback provided to FA students on their tutorial AMs in the lead-up to their essay assessment was more formalised than feedback provided to students in the MS tutorials. Finally, MS students were mainly from within the marketing discipline and would have had greater prior experience in essay writing and exposure to argument construction. Thus, AM may have been perceived as more useful in the task of essay writing for the FA graduate student cohort.

3.2 Regression analysis

In this section, we run a regression analysis to test the most important variables in explaining the overall effectiveness of AM. The regression model is constructed with the response to Q1 as the dependent variable and the responses to Q2 to Q9 as independent variables. The model is presented as follows:

$$Q1 = b_1 + b_2Q2 + b_3Q3 + b_4Q4 + b_5Q5 + b_6Q6 + b_7Q7 + b_8Q8 + b_9Q9 + e$$

The purpose is to find the determinants for the overall usefulness of AM documented by Q1. If the regression coefficient of an independent variable is positively significant, it means that the independent variable is one of the determining factors. The results are reported in Table 2. For the FA group, Q2 (coefficient = 0.3186, $p = 0.0045$) and Q3 (coefficient = 0.2743, $p = 0.0356$) are the only significant factors determining the effectiveness of AM (Panel A). The results from the MS group (Panel B, Table 2) show two more determining factors than the FA group. In addition to Q2 (coefficient = 0.2432, $p = 0.0090$) and Q3 (coefficient = 0.2236, $p = 0.0154$), the regression coefficients for Q4 (coefficient = 0.3831, $p = 0.0010$) and Q8 (coefficient = 0.1400, $p = 0.0819$) are also positively significant. Overall, the regression results indicate that the two most important AM features that contribute to overall effectiveness are its ability to visually represent the argument (Q2) and understand the argument (Q3). For MS students, the ability of AM to breakdown complex arguments into simple manageable components (Q4) was the strongest determinant, although this was not significant for FA students.

3.3 Analysis of descriptive comments

The students' descriptive comments from the questionnaire provided a qualitative explanation for the perceived strengths and weaknesses of AM. The comments were open-coded and interpretively analysed (Corbin & Strauss, 2008). The codes and resulting frequencies are illustrated in Table 3.

Being able to visually structure an argument was seen by FA and MS students as one of the key advantages of AM. Visual mapping of the argument provided clarity

Table 2. Regression analysis.

Panel A. Financial Accounting group.

		Coefficient	*t*-statistics	*p*-value
Intercept	b_1	–0.0594	–0.1828	0.8553
Q2 Visual representation	b_2	0.3186***	2.9139	0.0045
Q3 Understanding	b_3	0.2743**	2.1349	0.0356
Q4 Component break down	b_4	0.0494	0.5028	0.6163
Q5 Logical connections	b_5	0.1492	1.2235	0.2244
Q6 Help essay writing	b_6	0.1225	1.1260	0.2632
Q7 Time saving in writing	b_7	–0.0173	–0.1689	0.8663
Q8 Apply to other areas	b_8	0.1307	1.2666	0.2086
Q9 Face difficulties	b_9	0.0147	0.2738	0.7849

Panel B. Marketing and Society group.

		Coefficient	*t*-statistics	*p*-value
Intercept	b_1	0.1755	0.5807	0.5626
Q2 Visual representation	b_2	0.2432***	2.6580	0.0090
Q3 Understanding	b_3	0.2236**	2.4600	0.0154
Q4 Component break down	b_4	0.3831***	3.3699	0.0010
Q5 Logical connections	b_5	–0.0530	–0.4993	0.6186
Q6 Help essay writing	b_6	–0.0747	–0.8849	0.3781
Q7 Time saving in writing	b_7	0.1124	1.2920	0.1990
Q8 Apply to other areas	b_8	0.1400*	1.7556	0.0819
Q9 Face difficulties	b_9	–0.0109	–0.1987	0.8429

Note: *, ** and *** indicate significance at the 0.10, 0.05 and 0.01 levels (two tailed), respectively.

and deeper understanding (e.g. 'it brings clarity and logic to thoughts and ideas' [FA], 'it allows a visual representation of an argument and is clear and easy to use' [MS]). The students also suggested that AM provided them with a framework to gather their thoughts and ideas and to then logically construct their own arguments ('it provides a powerful and logical approach for constructing complex and balanced arguments' [FA]). In addition, it was suggested that applying AM to the complex arguments presented in the subject literature enabled them to deconstruct these arguments into their components and assess both the structure of the argument and the supporting evidence (e.g. 'it helps to break an argument down into key components' [MS]).

The students' responses for what they liked least demonstrated that experiences were polarised. Some students had issues with the application of AM. While some students found AM to be 'simple to learn and easy to apply' (FA), others found it overly complex, limiting and inflexible. While some students suggested that the application of AM saved time and helped them with essay writing, others found it time consuming. For example, one FA student felt that it was difficult to use AM for a group assignment as different people in the group had different approaches to the layout of the map itself. Additionally, while some students found AM to be a valuable

Table 3. Coded descriptive comments.

Panel A. Positive responses.

		FA	MS
Theme	Code	No. of Responses	No. of Responses
Visual elements	AM provides arguments with a visual structure	2	14
	Visual representation makes the argument clear and easier to understand	3	18
	Visual representation ensures that arguments are supported with evidence	0	6
	AM is especially effective for visual learners	0	1
Logical structure and flow	AM assists with the linking of arguments	0	5
	AM assists with making arguments flow logically	5	3
	AM helps to coherently structure essay arguments	3	4
	AM provides a powerful and logical approach for constructing complex and balanced arguments	6	13
	AM brings clarity and logic to thoughts and ideas	12	4
Simplifying complex arguments	AM helps to breakdown complex arguments and questions into components	1	9
	AM simplifies arguments	3	2
	AM helps to be succinct and to the point	0	2
AM method	Learning AM has changed my approach and thought processes towards constructing and understanding arguments	1	5
	AM saves time and assists with essay preparation	1	10
	AM was simple to learn and easy to apply	6	9
	AM is a practical and valuable tool	1	2
	AM is transferrable to other subjects	0	3
CAAM software	The Software was useful and easy to use	1	6

Panel B. Negative responses.

		FA	MS
Theme	Code	No. of Responses	No. of Responses
Learning styles	AM does not suit my style of writing and thinking	1	11
	I constructed or changed the AM *after* writing the essay	0	5
	Difficult to do it as a group assignment as different students had different AM	1	0
Teaching and practicing	Required more time practicing and learning AM	1	17
	Needed more information and help on applying AM to the essay	1	11
	Did not have adequate knowledge of the software	0	5

Table 3. (*Continued*).

Panel B. Negative responses.

		FA	MS
Theme	Code	No. of Responses	No. of Responses
	Explicit teaching of AM more appropriate to first-year university students	1	5
AM method	AM is not a beneficial skill or method	2	4
	AM is not necessary to understand a topic or argument	2	8
	AM is unnecessarily complex	4	4
	AM is limiting and inflexible	5	3
	Using AM was time consuming	8	15
CAAM software	Software was not compatible with Macs	1	4
	The software was not user friendly	4	15
	Using the software was time consuming	1	9
	The software was inflexible	2	5

tool that could be transferred to other subject areas, other students viewed AM as not being beneficial.

3.4 Measuring pre- and post-AM skills (MS group)

The MS student cohort was also asked to participate in a pilot of the online California Critical Thinking Skills Test (CCTST) to independently evaluate the effectiveness of the integration of AM techniques. The students completed the CCTST online test twice: initially before AM was introduced to the students (pre-test) and at the end of the semester (post-test), with a period of 10 weeks between tests. Participation was entirely voluntary. Of the 182 students enrolled in the subject, 111 participated in the first test. Participation in the second test was limited to these 111 students and only 25 completed this second test (response rate 13.7%). The results for the 25 students who completed both the pre-test and post-test CCTST were matched and are reported in Table 4. It transpires that, of the 25 students, 12 had first test scores below the total mean and 13 had first test scores above the total mean, indicating that the results were not biased towards either higher or lower achieving students in the second test.

The MS students' CT ability between the first and second tests show a shift from a mean score of 23.40 to 24.44 (out of a maximum possible score of 34, significant at 5% level). This represents an overall improvement in CT ability of 4.44%. Pascarella and Terenzini (2005) suggest that the CT skills of undergraduate students increase approximately 0.5 standard deviations over the duration of their undergraduate career. Thus, the increase in CT skills experienced by the students who completed the CCTST pre- and post-tests over the one-semester AM intervention demonstrates gains normally attributed to half of their 3–4 year undergraduate career. In addition, Pascarella and Terenzini suggest that the greatest gains in CT ability occur in the early stages of undergraduate education. The students in our study were in their final undergraduate year.

Table 4. Descriptive statistics for Test 1 and Test 2 (n = 25).

	TEST 1		TEST 2		
Critical thinking attribute	Mean	SD	Mean	SD	p-value
Induction	13.12	2.32	13.84	2.78	0.0170**
Deduction	10.28	2.61	10.60	2.43	0.1142
Analysis	8.00	1.66	8.36	1.41	0.0425**
Inference	9.68	2.64	10.20	2.27	0.0069***
Evaluation	5.72	1.40	5.88	1.99	0.4482
Total	23.40	4.38	24.44	4.91	0.0209**

Note: ** and *** indicate significance at the 0.05 and 0.01 levels (two tailed), respectively.

This result provides some degree of objective evidence of the enhancement of CT skills for at least some of the students. In this study it was not possible to use a control group. It was deemed pedagogically inappropriate to direct different students to an AM or non-AM group and student choice was impractical for the management of the subjects. However, other studies have done this. Harrell's (this issue) study involved an experimental group of 68 students and a control group of 62 students undertaking the same introductory philosophy course in the fall of 2004. Students elected to participate in either the experimental group (taught AM) or the control group (not taught AM). Both classes involved multiple instructors. Both groups were asked to complete a pre-test and post-test argument schema, which was structurally identical. She found that there was a significant higher post-test score for the treatment group compared to the control group, yet 'the gains from pre-test and post-test of the treatment group were not statistically different from that of the control group'. She explains this 'puzzle' by dividing participants according to their pre-test scores. She concludes that AM intervention leads to greater gains in students with low initial AM skills.

4. Discussion

In this study, we tested student perceptions of the effectiveness of AM/CAAM following a single-class intervention with two relatively large student groups in a normal semester's teaching setting.

Results from both FA and MS groups show good support for AM/CAAM as a CT tool. In general, students felt that AM helped them create a visual map of an argument, gain a clear understanding of arguments, break down complex arguments into manageable components and support arguments with evidence. The results are statistically significant, providing evidence that AM techniques improve students' skills in argumentation. This result is remarkable given the minimalist constraints of the trial: that is, a single class of 50 minutes allocated to the AM intervention and further classroom practice, but no subsequent, expert-led follow-up.

These conclusions are further supported by the overall improvement in CT ability as measured by the first and second CCTSTs conducted on the 25 MS students. In addition, at least half of the students (53%) in the FA group indicated that they plan to use AM in other areas of their studies. This provides encouraging support that FA students *themselves* realise the potential of AM as a CT tool.

Nonetheless, support was clearly not as strong as in earlier, more extensive, published studies. Indeed, given the time allocated to the trial it would be surprising if it was. Some key factors underpinning the disparity in student responses were: (1) the procedure used to integrate AM and CAAM into the subjects and (2) the heterogeneous abilities and motivations of the students.

An understanding of the effectiveness of the integration procedure of AM/CAAM may be garnered from further analysis of the following. Perceived effectiveness is restricted to understanding of the arguments (Q2 to Q5). The responses from these questions are all significantly greater than the scale mid-point of 3. Regression analysis further points to Q2, Q3 and Q4 as being the most significant variables in explaining the overall effectiveness of AM. However, the results on essay writing (Q6 and Q7) are mixed. Financial Accounting students reported some improvement in their essay writing but did not report spending less time with the help of AM. Similarly, MS students did not respond positively to the question concerning essay writing. Students' open comments shed some light upon the apparent inconsistency between students' positive responses about AM and their less than enthusiastic view on the application of AM to essay writing. A significant number of students suggested that they needed 'more time learning and practising' AM and 'more assistance with translating their AM into an essay'. It could be said, therefore, that while the AM method itself was perceived as useful by students, not enough time was given to teaching *how* to use it in relation to essay writing. Given the timetabling constraints of the intervention, this was clearly true. Although students were taught to prepare simple to more complex AMs during class time, they were not given an opportunity to practice writing essays following the preparation of an argument map. Thus, unsurprisingly, they faced difficulties when required to do so for the assignments.

These issues do not necessarily indicate a deficiency in AM or AM/CAAM but they do show a deficiency in classroom design and practice. More time and effort was clearly needed in showing students how to translate from maps into the genre of written assignments. Despite these misgivings, this trial demonstrates that even short-term, minimalist interventions in the context of a normal subject stream can result in improvements in students' understanding and application of arguments. Of course, further embedding the tool into the curriculum by means of follow-up, expert-led practice sessions over several weeks would reinforce student expertise in AM. Nevertheless, our findings align with our purpose of assessing the effectiveness of a one-shot AM intervention in enhancing the CT of students in the context of a regular class.

A second issue is the possibility that different thinking styles determine the effectiveness of AM/CAAM. For example, visualisation of the argument is likely to be of greater assistance to those with a visual thinking style (Zhang, 2003). While some students suggested that AM was a valuable tool, was easy to apply, saved time and assisted with essay writing, an equal proportion of the responses suggested that AM was complex, time consuming and not suited to their learning/thinking styles. This suggests that AM/CAAM might be more appropriate for some thinking styles more than others. Indeed, thinking styles have been shown to influence CT dispositions (Zhang, 2003). This effect and the pedagogical implications should be investigated further.

A limitation of our study is that questionnaire results from both FA and MS student groups constitute self-reported data. The results are an indication of students' perceptions of the AM method upon their own CT skills. These results are, of course, open to the bias associated with self-reported research methods. Of particular relevance for

this study, Kruger and Dunning (1999) found evidence that respondents with little ability in a focal skill (such as CT) have inflated perceptions of their abilities as they are not able to recognise their own errors. The CCTST pre-test/post-test results were an attempt to counter this by providing objective data. These results are suggestive of support for the students' perceptions of their enhanced CT abilities following application of AM. To overcome bias inherent in self-reported responses it is recommended that future studies employ robust independent empirical measurement methods (such as the CCTST) to assess actual shifts in students' CT abilities.

Acknowledgements

This research was supported by a Learning and Teaching Performance Fund grant from the University of Melbourne. We gratefully acknowledge this support.

References

Alvarez, M. (2007). *Does philosophy improve critical thinking skills?* Melbourne: University of Melbourne.

Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: A review. *International Journal of Research in Marketing, 13*(2), 139–161.

Commonwealth of Australia. (2000). *Employer satisfaction with graduate skills.* Canberra, ACT: Department of Education, Training and Youth Affairs.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research.* Thousand Oaks, CA: SAGE.

Davies, W.M. (2009a). Computer-assisted argument mapping: A rationale approach. *Higher Education, 58*(6), 799–820.

Davies, W.M. (2009b). Not quite right: Teaching students how to make better arguments. *Teaching in Higher Education, 13*(3), 327–340.

Davies, W.M. (forthcoming). Mind mapping, concept mapping, argument mapping: What are the differences and do they matter? *Higher Education.* DOI: 10.1007/s10734-010-9387-6.

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.

Evans, E., Tindale, J., Cable, D., & Hamil Mead, S. (2009). Collaborative teaching in a linguistically and culturally diverse higher education setting: A case study of a postgraduate accounting program. *Higher Education Research & Development, 28*(6), 597–613.

Graduate Careers Australia. (2006). *Graduate outlook.* Retrieved October 12, 2007, from http://www.graduatecareers.com.au/content/view/full/52

Harrell, M. (2007). No computer program required: Even pencil-and-paper argument mapping improves critical thinking skills. *Teaching Philosophy, 31*(4), 351–374.

Hitchcock, D. (2004). The effectiveness of computer-assisted instruction in critical thinking. *Informal Logic, 24*(3), 183–218.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence leads to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134.

Kulhavy, R.W., Schwartz, N.H., & Caterino, L.C. (1985). Conjoint retention of maps and related discourse. *Contemporary Educational Psychology, 10*, 28–37.

Monash critical thinking study. (2009). Retrieved October 22, 2009, from http://arts.monash.edu.au/philosophy/research/thinking/04webmap.php

Paivio, A. (1983). The empirical case for dual coding. In J. Yuille (Ed.), *Imagery, memory and cognition* (pp. 307–332). Hillsdale, NJ: Lawrence Erlbaum.

Pascarella, E.T., & Terenzini, P.T. (2005). *How college affects students: A third decade of research.* San Francisco, CA: Jossey-Bass.

Rider, Y., & Thomason, N. (2008). *Cognitive and pedagogical benefits of argument mapping: LAMP Guides to the way to better thinking.* London: Springer.

Rossiter, J.R. (2002). The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing, 19*, 305–335.
Schwatz, N.H. (1988). Cognitive processing characteristics of maps: Implications for instruction. *Educational and Psychological Research, 8*(2), 93–102.
Twardy, C. (2004). Argument maps improve critical thinking. *Teaching Philosophy, 27*(2), 95–116.
van Gelder, T., Bissett, M., & Cumming, G. (2004). Enhancing expertise in informal reasoning. *Canadian Journal of Experimental Psychology, 58*, 142–152.
Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology Review, 14*(3), 261–312.
Winn, W. (1991). Learning from maps and diagrams. *Educational Psychology Review, 3*(3), 211–247.
Zhang, L.-F. (2003). Contributions of thinking styles to critical thinking dispositions. *Journal of Psychology, 137*(6), 517–544.

Appendix 1. Questionnaire paper for argument mapping

Please circle the degree to which you agree with the following statements, where 1 = strongly disagree, and 5 = strongly agree. 'N/A' = the statement is not applicable to you (e.g. because you did not attended these sessions or did not complete the assessment task) .

1. THE USE OF ARGUMENT MAPPING

	N/A	Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
1. AM is an effective critical analytical thinking tool	N/A	1	2	3	4	5
2. AM helps me to have a visual representation of the argument	N/A	1	2	3	4	5
3. AM helps me to understand the argument	N/A	1	2	3	4	5
4. AM helps me to break down complex arguments into simple manageable components	N/A	1	2	3	4	5
5. AM helps me to build up logical connections between arguments and evidence	N/A	1	2	3	4	5
6. Preparing an AM helps me to write the essay	N/A	1	2	3	4	5
7. AM helps me to save time when writing the essay	N/A	1	2	3	4	5
8. I plan to apply the AM knowledge to other critical thinking areas	N/A	1	2	3	4	5
9. I had difficulties in completing the AM assessment task *due to lack of knowledge of argument mapping*	N/A	1	2	3	4	5

2. GENERAL QUESTIONS

Additional Comments: (please express your views fully)

What did you like MOST about AM?

_____
_____
_____

What did you like LEAST about the AM?

_____
_____
_____

In your opinion, what method in addition to AM could help to improve critical thinking skills?

_____
_____
_____