Does doing Critical Thinking AS level confer any advantage for candidates in their performance on other A levels?

Beth Black and Tim Gill Research Division

Critical Thinking can be defined as analytical thinking which underlies all rational discourse and enquiry (Black *et al*., 2008). It is of some interest whether when taught as a separate course, it can be transferred by students to other subject domains and improve their performance in them. In the UK context, Critical Thinking AS level was introduced in schools in 2001 and, as such, represents the catalyst for a large scale introduction of this discipline into schools.

There is now much research that shows that the teaching of Critical Thinking (CT) does indeed improve critical thinking skills. Abrami *et al*. (2008) provides an excellent meta-analysis of studies into the effectiveness of teaching CT. The average effect size was 0.34, indicating that CT interventions tend to have a small to moderate impact upon the development or enhancement of CT skills and dispositions. In one of Abrami *et al*.'s sub-analyses, the 117 studies included in the meta-analysis are divided into one of four types based on the instructional method of the intervention – general, infusion, immersion and mixed[1]. In a 'general' course, CT is taught without any other specific subject matter or domain content – in other words, the main (and only) objectives of the course are to improve CT skills and dispositions. For both 'infusion' and 'immersion' courses, CT is delivered through other subject content, though where they differ is that CT principles and learning objectives are explicit in an 'infusion' approach, while implicit in an 'immersion' approach. Finally, the 'mixed' approach again involves teaching CT through another subject, though it is delivered as an independent track within that subject. The meta-analysis revealed that there were positive effect sizes for all types of intervention. However, immersion (with no explicit CT objectives) was least effective (effect size = 0.09); while the mixed approach was the most effective (effect size = 0.94) with the general and infusion approaches also having moderate to large effect sizes (0.38 and 0.54 respectively).

The result for the 'general' approach is quite interesting given John McPeck's (1981) well-known objections to CT being taught in such a way, as a standalone subject. His point is that one always has to think about something.

> *In isolation from a particular subject, the phrase "Critical Thinking" neither refers to nor denotes any particular skill. It follows from this that it makes no sense to talk about Critical Thinking as a distinct subject and that it therefore cannot be profitably taught as such. To the extent that critical thinking is not about a specific subject X, it is both conceptually and practically empty.*

Thus, Abrami *et al*.'s research appears to contradict this view and show that CT, taught generally as a standalone subject, can improve CT skills.

However, there is less research which shows whether CT skills when taught (in any of the four approaches described above) can be profitably transferred to other subject domains. This is of keen interest since much of the rhetoric around CT as a worthwhile educational goal rests on the notion that it is not just good *in itself* but "being able to think critically is a necessary condition of being educated in a more general sense" (Norris, 1985). Again, there is much speculation as to the best way to deliver CT so as to foster transferable CT skills and dispositions (e.g. Brown, 1997; Halpern, 1998).

In the UK context, from a survey of CT teachers (Black, 2010), we know that CT tends to be delivered separately or discretely – as the 'general' approach, rather than within other subjects. This survey also revealed that the overwhelming majority of respondents (95.7% of all respondents) believed that students did (spontaneously) transfer these skills to other subjects to the benefit of their performance, skills and understanding in other subjects. Of course, the crucial word here is 'believed'. It was the belief or perception of teachers, based on their own (anecdotal) experiences with their students, rather than hard evidence:

> *…the majority [of students] find it quite useful and they now write better essays or think more logically. One said "it has changed my whole way of thinking".*

As well as based upon their understanding of how these skills form a fundamental part of other educational endeavour:

> *[CT] complements analytical requirement in many subjects…Many of our "most-improved" students in year 13 took CT… perhaps due to developing transferable skills.*
>
> *Many subjects call for reasoned arguments. What better way to prepare them?*

Therefore, we were particularly interested to see whether there was any data to support these views that students who have taken CT do better in their other subjects.

This report looks at the performance at A level of candidates who had taken CT AS level, in comparison to candidates who had not taken the CT AS level. It was hypothesised that CT skills are transferable and can be applied to other subjects in a beneficial way. Thus candidates gaining good CT skills at AS level may improve their performance at A level.

The hypothesis that we put forward here is that candidates who took CT, and gained a good grade in it performed better in their A levels than similar candidates who did not take CT. If this is shown to be the case then we can infer that the skills gained by taking the CT AS level were beneficial to the candidates in their other A levels. Of course we cannot prove this association, because many other factors influence how well candidates perform in their A levels.

1. This classification is based upon Ennis's (1989) typology.

Data and methods

Data taken from the NPD databases for 2005 and 2006 were used for this research. These are databases of all exams taken in England and Wales by pupils of different ages. From these it was possible to identify candidates taking CT AS level, and follow them through to their A level results.

The first analysis looked at whether candidates who had performed well in the CT AS level (A and B grade candidates) performed better at A level on average than candidates who did not take CT at all. Candidates getting grades A or B at CT AS level in 2005 were identified in the database. These candidates were then matched to a set of candidates not taking CT by ability (as measured by GCSE mean grade) and the A level results in 2006 for the two different groups were compared.

To choose the matched candidates, the mean GCSE was calculated by converting grades into numbers, with 8 for an A*, 7 for an A and so on down to 0 for U. The distribution of GCSE mean grade for the candidates receiving grades A and B for CT AS level was inspected (n=2208). These pupils were divided up into 20 approximately equal groups (in terms of numbers) by mean GCSE grade. A random sample of candidates was then taken, matched to each of the 20 groups, from the remaining candidates in the database (the non-CT group). For example, the bottom group of CT candidates consisted of 108 pupils with a mean GCSE of between 3.86 and 5.86. A matching sample of 108 was (randomly) taken from the group of non-CT candidates with a GCSE mean grade of between 3.86 and 5.86.

This was done for each of the 20 groups. The 20 random samples were joined together to create one overall matching dataset. The following summary statistics demonstrate that the CT and non-CT groups were well matched:

Table 1: Summary statistics for matched groups

	N	*GCSE Mean*	*SD*	*Min*	*Max*
CT	2208	7.18	0.67	3.86	8.00
Non-CT	2208	7.17	0.69	4.00	8.00

A further analysis was undertaken, comparing the A level performance of all the CT candidates with the performance of all candidates taking A levels in 2006.

Results

Comparison of means

First we looked at the mean A level grades and total A level score for the two groups of candidates (A or B grade CT candidates and non-CT). To calculate the means and totals each grade was transformed into a number, with 10 for a grade A, 8 for a grade B and so on, down to 0 for a grade U. A statistical test (Kolmogorov-Smirnov or K-S test) was used to determine if the difference between the groups in the distribution of their mean or total A level scores was statistically significant or could be attributed to chance[2]. The results are shown in Table 2.

The mean and total A level scores were clearly higher for the high performing CT candidates on average, compared to the non-CT candidates. The difference in the mean A level (9.12–8.68 = 0.44) translates to around a quarter of a grade per A level. The effect is not very large, but would be the equivalent of a grade for a candidate taking four A levels.

2. It was not possible to test for difference in the means using a t-test as the distributions of mean and total grades were not normal.

Table 2: Overall mean A level performance for CT and non-CT candidates

	Group	*N*	*Mean*	*Std. Deviation*	*Sig of K-S Test*
Mean A level	Non-CT	7295	8.68	1.63	<0.001
	CT	7691	9.12	1.20	
Total A level	Non-CT	7295	32.04	9.40	<0.001
	CT	7691	34.39	8.89	

According to the K-S test the differences in the distributions of both the mean and total A level scores were highly significant. The figure in the final column gives the probability of a difference the same as or larger than observed occurring if there was actually no difference between the two groups. A figure of less than 0.05 is generally considered to be significant, and less than 0.01 highly significant.

The direction of the difference can be seen by sketching the cumulative distributions functions of A level grade and total A level score for the two groups:

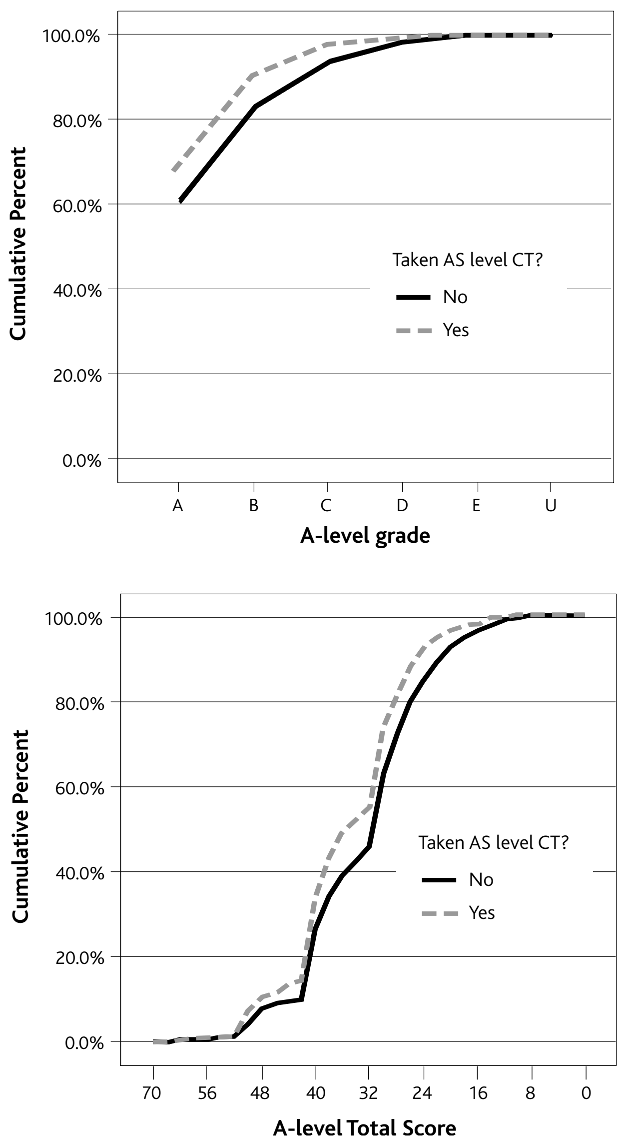


Figure 1: Cumulative frequency distributions of A level grade and total A level score

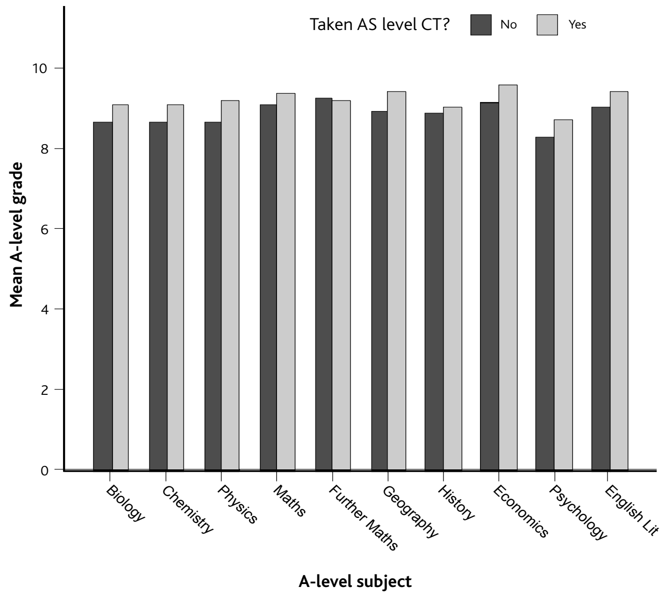


Figure 2 : Mean A level grades for individual subjects

It is clear that the distribution for the CT candidates is further to the left, meaning a larger percentage of this group were towards the top in terms of A level grade than the non-CT students. For instance, in the left hand figure for the non-CT students just over 80% of the grades received were at least a grade B, compared to about 90% of the grades received by the CT students.

We also investigated performance at A level in the most popular subjects individually. For this we selected the candidates from the groups (CT and non-CT) taking each individual subject. This meant that for some of the subjects the candidates in the two groups were no longer exactly matched for ability. Thus some caution should be exercised with the results for these subjects. The mean A level grade for each group in each subject is shown in Figure 2. This data is also displayed in Table 3, along with the number of students and the outcome of the K-S Test. To assist interpretation the GCSE mean grades for candidates in each group taking the subject in question are also listed in the table:

Table 3: Mean A level grades for individual subjects

	GCSE mean grade		*A level candidates*		*A level mean grade*		*Sig of K-S Test*
	Non-CT	*CT*	*Non-CT*	*CT*	*Non-CT*	*CT*	
Biology	7.34	7.33	670	589	8.70	9.10	0.010
Chemistry	7.40	7.44	658	559	8.68	9.14	0.044
Physics	7.37	7.40	364	359	8.70	9.22	0.018
Maths	7.41	7.40	846	801	9.12	9.38	0.065
Further Maths	7.54	7.44	156	182	9.28	9.22	1.000
Geography	7.17	7.21	296	236	8.95	9.44	0.059
History	7.27	7.21	448	681	8.88	9.09	0.559
Economics	7.36	7.44	174	264	9.17	9.60	0.037
Psychology	6.86	6.79	254	267	8.31	8.73	0.333
English Lit	7.25	7.29	505	676	9.09	9.45	0.002

In all the subjects apart from Further Maths the group performing well in CT had a higher mean A level grade in the subject than the non-CT. The K-S test shows there was a significant difference in the distribution of A level grade between groups in several of the subjects. However, we must also consider any differences in the mean GCSE grades.

For Biology and Maths there was virtually no difference between the GCSE mean grades of the two groups, so we can assume they are matched. The K-S test was significant for Biology, so we have evidence that the CT candidates performed better in this subject. However, the test was not significant for Maths, so there was no evidence of improved performance.

For Chemistry, Physics, Economics and English Literature the CT candidates had slightly higher GCSE mean grades, so the significantly better performance of this group at A level was not as high as suggested in the table, and would potentially be non-significant if we had data that matched exactly on prior attainment. However, in each case the difference in GCSE mean grade was small in comparison to the differences in mean A level grade so it is still probable that a significant difference was present.

For Psychology and History, although the difference between the two groups at A level was not significant, the non-CT group had a higher GCSE mean grade. Thus it may be that if the groups were matched more exactly, the performance of the CT group would have been significantly better at A level.

Finally, there was no significant difference between the two groups in their A level Geography performance and, as the CT candidates had a slightly higher GCSE mean grade, there was certainly no evidence that these candidates performed better at A level.

In summary, there is evidence that candidates who achieved high grades in AS level CT performed better overall at A level than candidates who did not study CT at all. There is evidence that this advantage presents itself across a wide range of subjects, in sciences, social sciences and arts subjects. This backs up the hypothesis that CT skills are transferable and applicable to a wide range of subjects.

Regression analysis

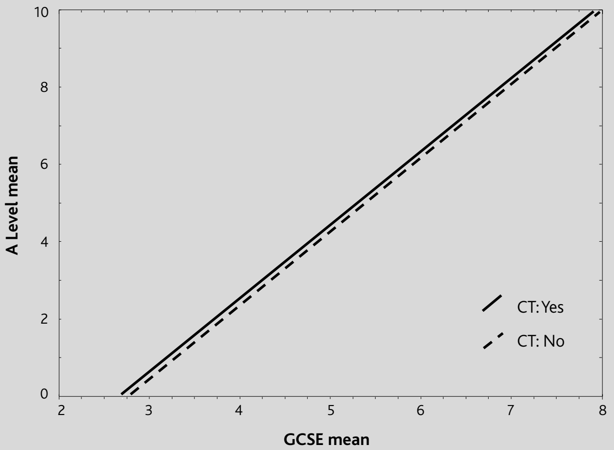
In the previous section we only selected candidates who received a grade A or B in Critical Thinking at AS level. An alternative way of analysing the data is to undertake a linear regression on overall A level performance for *all* candidates. This predicts a mean A level score (and separately a total A level score), based on certain variables. We allowed for previous attainment by including candidates' GCSE mean grade in the model. A variable indicating whether or not the candidate studied AS level CT was also included, which enabled the impact of taking this qualification to be analysed, for a given level of prior attainment. It was also possible to analyse the impact of having received a particular grade on the CT AS level.

Mean A level grade

Figure 3 shows some basic regression output from a model with mean A level grade as the dependent variable and GCSE mean grade and whether or not the candidate had studied CT as the predictor variables.

The R square is a measure of the amount of variation in the dependent variable that can be accounted for by variations in the predictor variables. Thus, 52% of the variation is explained by the regression model, which is reasonable.

We can see from the variables table that both the predictor variables are highly significant (Sig < 0.01). This means we have evidence that changes in these are associated with changes in the mean A level grade. This effect is quantified by B, which is the change in the dependent variable as a result of a unit increase in the predictor variables. Thus the



Variable	B	Sig	R	R Square
Constant	-5.34	<0.01	0.72	0.52
GCSE mean grade	1.92	<0.01		
Critical Thinking	0.20	<0.01		

Figure 3: Regression output on mean A level

model predicts that an increase in GCSE mean grade of 1 unit (equivalent to 1 grade) leads to an increase in A level mean grade of 1.92 (equivalent to just under one grade).

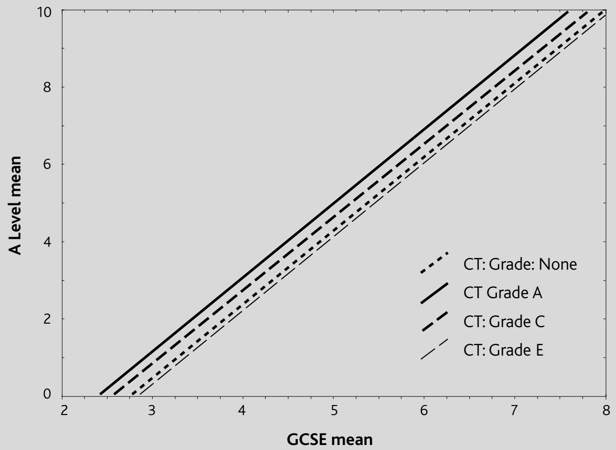
The CT variable is specified as a 1 if the candidate has taken the AS level and a zero if not. Thus, according to the model, having taken the AS level increases (on average) candidates mean A level grade by 0.20, or about one tenth of a grade.

The graph in Figure 3 can help with interpreting the model. This plots GCSE mean against A level mean (as predicted by the model) for the CT and non-CT groups. This demonstrates that for a particular level of GCSE mean, the model predicts a higher A level mean grade for candidates who took the CT AS level, than for those who did not. However, the difference is clearly not very large.

The second model, which is shown in Figure 4, also took into account the grade received by the candidates who took the AS level in CT.

The R square is very similar to the previous model. Once again all of the predictor variables are highly significant. The interpretation of this model is, however, more complicated. The grade received at AS level has been split up into a set of 'dummy' variables, one for each grade (apart from U). The coefficients in the table (B) represent the difference in the predicted mean A level for candidates who have received the particular grade in CT *in comparison to a candidate who received a grade U*. So, a candidate receiving an A grade has a predicted mean A level grade 0.91 higher than a U grade candidate.

To compare the performance of a candidate getting a particular grade on the CT AS level with one not taking the qualification at all, a combination of the coefficient for the grade and the CT coefficient is required. For example, imagine two candidates with the same GCSE mean grade, one having taken the AS level in CT and received a grade C, and the other having not taken CT. The predicted mean A level grade for the candidate who took the CT AS level will be 0.69–0.27 = 0.41 higher than the candidate not taking CT. Thus the overall effect is an increase in predicted mean A level grade of 0.41, or around one fifth of a grade. For a candidate with a grade B the overall predicted increase is 0.67 (0.94–0.27) and for a grade A candidate it is 0.64 (0.91–0.27), both of which are about one third of a grade. For a candidate who took three A levels this amounts to around one grade overall.



Variable	B	Sig	R	R Square
Constant	-5.25	<0.01	0.72	0.52
GCSE mean grade	1.91	<0.01		
Critical Thinking	-0.27	<0.01		
Grade A	0.91	<0.01		
Grade B	0.94	<0.01		
Grade C	0.69	<0.01		
Grade D	0.51	<0.01		
Grade E	0.22	<0.01		

Figure 4: Regression output on mean A level, including grades achieved on CT

Again, a graph can aid interpretation of this result. Figure 3 plots the GCSE mean against predicted A level mean for candidates achieving different grades in their AS level CT, and for the non-CT group[3]. Thus a grade E candidate had a very slightly lower predicted mean A level grade than a non-CT candidate with the same prior attainment. Candidates getting a grade A or C had a higher predicted mean A level than a non-CT candidate with the same prior attainment.

It is worth noting the unexpected result that the coefficient for a grade B in CT is higher than that for a grade A. Inspection of the distribution of mean A level showed this to be a 'ceiling' effect. Of the candidates who received an A or a B at CT a large proportion (1,414 out of 3,357) received all grade As at A level, giving them the maximum mean A level score of 10, and many others had a mean A level grade of 9 or more. Thus the level of discrimination was not enough to be able to distinguish between the CT grade A and grade B candidates.

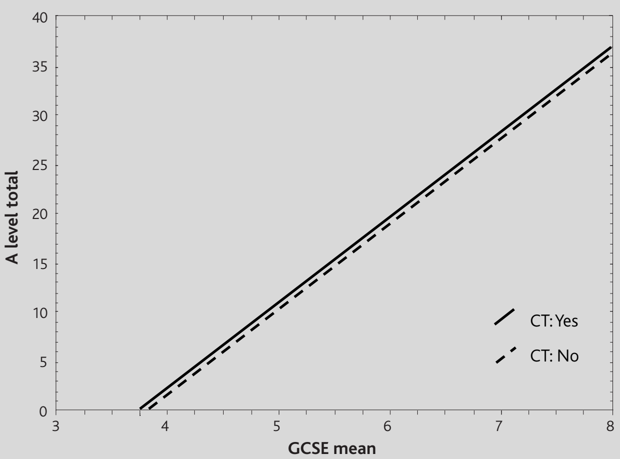
Total A level grade

We repeated both models using total A level score as the dependent variable. Figures 5 and 6 have the output from the two models with the same predictor variables as above.

Both models had reasonable R square values and all coefficients were significant. For the overall model, having taken CT increased the predicted total A level score by 0.64, or about one third of a grade. In terms of the individual grades, getting a grade U reduced the predicted total by 1.69, compared to not taking CT, and a grade E reduced it by 0.58. For all other grades the predicted score increased compared to not taking CT, in ascending order of grade. Having a grade A increased it by 3.71, equivalent to almost two grades.

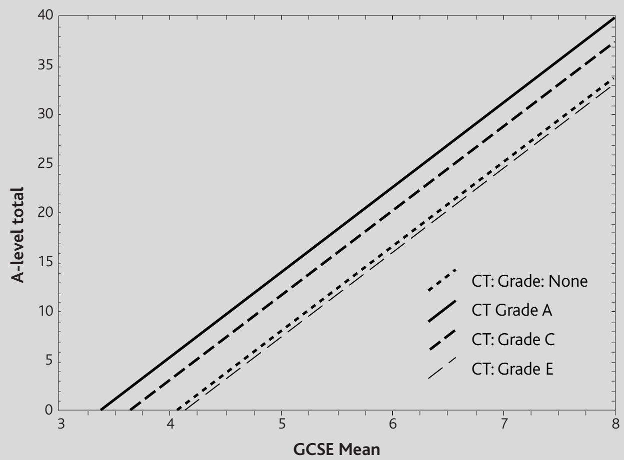
Note that in this case, the grade A coefficient was larger than the grade B coefficient, so as expected getting a grade A gave more of an advantage in terms of total A level score than getting a grade B. This was because

3. The lines for grade A, C and E only are shown to avoid crowding the graph too much.



Variable	B	Sig	R	R Square
Constant	-32.97	<0.01	0.75	0.56
GCSE mean grade	8.66	<0.01		
Critical Thinking	0.64	<0.01		

Figure 5: Regression output on total A level score



Variable	B	Sig	R	R Square
Constant	-32.51	<0.01	0.72	0.52
GCSE mean grade	8.59	<0.01		
Critical Thinking	-1.69	<0.01		
Grade A	5.39	<0.01		
Grade B	4.69	<0.01		
Grade C	3.14	<0.01		
Grade D	2.21	<0.01		
Grade E	1.12	<0.01		

Figure 6: Regression output on mean A level, including grades achieved on CT

there is more discrimination at the top end with regards to total A level score, since the candidates with a mean A level score of 10 (all grade A) are split into those with a total A level score of 30, 40, 50 and 60.

Thus we have further evidence that candidates taking the AS level CT, and getting a reasonably good grade, perform better overall on their A levels the following year. According to the model presented here the improvement for the top candidates is around one third of a grade on the mean A level and around two grades when looking at total A level score.

Discussion

We should note some caveats of this research. Although we have shown an association between taking CT at AS level and performing well at A level, we cannot be sure that the former *causes* the latter. It may be that candidates who perform well on CT do so because they already possess the skills and attributes to perform well academically more generally (although this had not differentially benefitted them at GCSE level).

Secondly, since not all schools offer CT, there may be a school effect that we have not been able to identify. For instance, perhaps only the better schools offer it, in which case the candidates in these schools are likely to perform better overall. An alternative analysis would be to use data over time, and see if centres that started teaching CT saw an improvement in the progress of pupils from GCSE to A level in subsequent years, whilst similar centres that did not teach CT improved less or not at all.

However, if we accept the interpretation that studying CT AS level can improve performance in other subjects, it is worth reflecting on this a little further. Piecing these findings together with the survey data (Black, 2010), we might be surprised by any discernible transfer effect for a number of reasons:

- Teachers often reported little or no training in improving their own CT skills or how to teach the discipline.
- Limited resourcing of the courses in terms of amount of dedicated timetabling as well as other resources (e.g. teaching materials).
- A significant minority of centres reported low motivation of staff and students where the CT course was obligatory rather than optional.
- Teachers tended to report that their overall agenda or aim was for students to achieve a good grade in the CT exam, rather than to foster transferable skills and dispositions.

Therefore, if this study does indeed present evidence for the transferability of CT, it might almost be viewed as an unintended (though serendipitous) consequence of delivering the CT AS level.

This research also suggests that it would be of some interest to investigate the mechanisms by which transferability is best fostered *within* this general or standalone approach to teaching.

References

Abrami, P.C., Bernard, R.M., Borokhovski, E., Wade, A., Surkes, M.A., Tamim, R. & Zhang, D. (2008). Instructional Interventions Affecting Critical Thinking Skills and Dispositions: A Stage 1 Meta-Analysis. *Review of Educational Research*, **78**, 4, 1102–1134.

Black, B. (2010). "It's not like teaching other subjects" – the challenges of introducing Critical Thinking AS level in England. *Research Matters: A Cambridge Assessment Publication*, **10**, 2–8.

Black, B., Chislett, J., Thomson, A., Thwaites, G., & Thwaites, J. (2008). A definition and taxonomy for Critical Thinking. *Research Matters: A Cambridge Assessment Publication*, **6**, 30–36.

Brown, A. (1997). Transforming schools into communities of thinking and learning about serious matters. *American Psychologist*, **52**, 399–413.

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher*, **18**, 3, 4–10.

Halpern, D.F. (1998). Teaching critical thinking for transfer across domains. Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist*, **53**, 4, 449–55.

McPeck, J. (1981). *Critical thinking and education*. Oxford, UK: Martin Robertson.

Norris, S.P. (1985). Synthesis of Research on Critical Thinking. *Educational Leadership*, **52**, 8, 4.