

**Socrates in the schools from Scotland to Texas:**

**Replicating a study on the effects of a Philosophy for Children program**

Frank Fair<sup>a</sup>

[psy\\_fkf@shsu.edu](mailto:psy_fkf@shsu.edu)

Lory E Haas<sup>a</sup>

[leh015@shsu.edu](mailto:leh015@shsu.edu)

Carol Gardosik<sup>b</sup>

[gardosc@stthom.edu](mailto:gardosc@stthom.edu)

Daphne D Johnson<sup>a</sup>

[ddjohnson@shsu.edu](mailto:ddjohnson@shsu.edu)

Debra P Price<sup>a</sup>

[EDU\\_DPP@shsu.edu](mailto:EDU_DPP@shsu.edu)

Olena Leipnik<sup>a</sup>

[SOC\\_OVL@shsu.edu](mailto:SOC_OVL@shsu.edu)

**Abstract**

In this article we report the findings of a randomised control clinical trial that assessed the impact of a Philosophy for Children program and replicated a previous study conducted in Scotland by Topping and Trickey. A Cognitive Abilities Test (CAT in the UK or CogAT in the USA) was administered as a pretest and a posttest to randomly selected experimental groups (N = 363, 186 seventh graders + 177 eighth graders) and control groups (N = 177, 79 seventh graders + 98 eighth graders). The students in the experimental group engaged in philosophy lessons in a setting of structured, collaborative inquiry in their language arts classes for one hour per week for a number of weeks. The control group received the standard language arts curriculum in that one hour. The study found that the seventh grade students who had experienced the P4C program showed significant gains relative to those in the seventh grade control group at a high level of statistical significance, but the eighth grade students in the experimental group did not show such gains over the eighth grade control group. It was discovered that the seventh grade teachers started the program early in the school year and continued it for a period of 22 to 26 weeks, while the eighth grade teachers started much later and used the program for only 4

---

<sup>a</sup> Sam Houston State University, Texas

<sup>b</sup> University of St Thomas, Texas

to 10 weeks. Our findings suggest that the P4C program must involve students in activities for a significant period of time before the program shows results, but that a meaningful impact on students' cognitive abilities can be achieved in about 24 weeks of lessons, less than half the time evidenced by the study by Topping and Trickey.

### **Key Words**

critical thinking, philosophy for children, cognitive abilities, community of inquiry, thinking through philosophy, reflective thinking

### **Introduction**

Several faculty members at a Texas university decided to replicate a study conducted in Scotland by Topping and Trickey on the effects of a Philosophy for Children (P4C) program. The group chose to replicate that study (a) because it was one of the very few randomised controlled clinical trials assessing the impact of a P4C program and (b) because the study showed large, durable gains in cognitive abilities by children who experienced a program of structured weekly philosophical group discussions. At the beginning of the 2010-2011 school year the Cognitive Abilities Test (CAT in the UK or CogAT in the USA) was administered as a pretest to the randomly selected experimental groups (N = 363, 186 seventh graders + 177 eighth graders) and the control groups (N = 177, 79 seventh graders + 98 eighth graders). At the end of the school year it was administered as a posttest. The students in the experimental group engaged in philosophy lessons in a setting of structured, collaborative inquiry in their language arts classes for one hour per week for a number of weeks. The control group received the standard language arts curriculum in that one hour. When the changes in the group mean scores on the CogAT were analysed, the seventh grade students who had experienced the P4C program showed significant gains relative to those in the seventh grade control group at a high level of statistical significance, but the eighth grade students in the experimental group did not show such gains over the eighth grade control group. Subsequently, we discovered by asking the teachers that the seventh grade teachers started the program early in the school year and continued it for a period of 22 to 26 weeks, while the eighth grade teachers started much later and used the program for only 4 to 10 weeks. The contrasting outcomes for the two groups indicate that the P4C program must involve students in activities for a significant period of time before the program shows results; however, while in the Scottish study the P4C

lessons continued for a total of 58 weeks, it appears that a meaningful impact on students' cognitive abilities can be achieved in about 24 weeks of lessons, less than half the time.

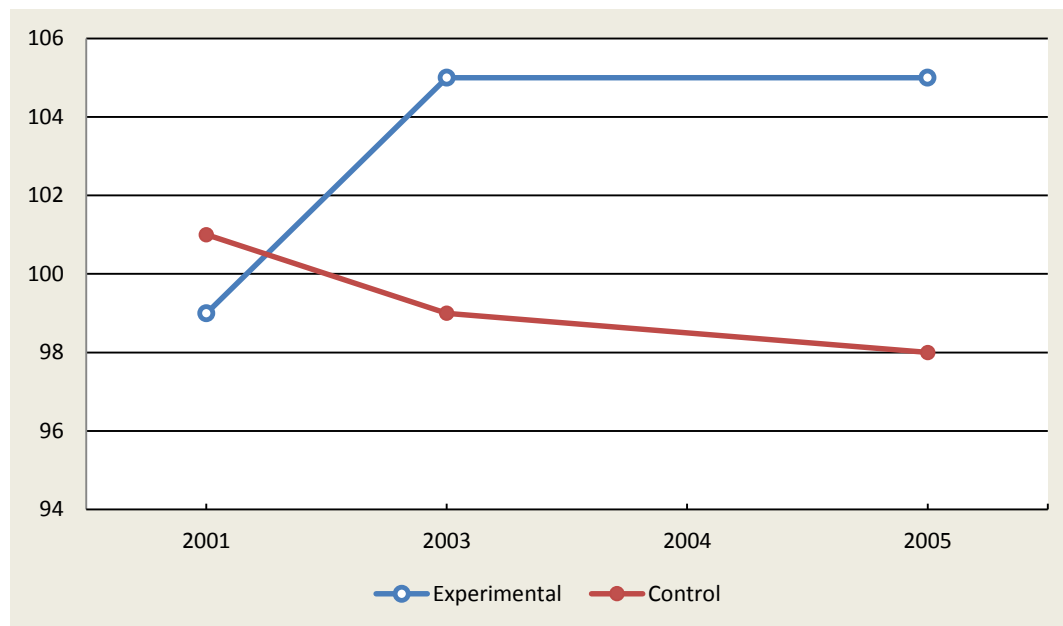
### **The beginning: Deciding to replicate an interesting study from Scotland**

In 2009, two of us at a university in Texas who had been interested in the Philosophy for Children (P4C) movement for some time joined with three other faculty members to form a P4C Interest Group. Partly from some of the group members' past experience with P4C, we thought that the P4C movement had great promise, but we wanted to see what evidence was available about its educational impact.

We noted that there were many studies and much anecdotal evidence, but we found one study in particular that merited great attention. It was a ground-breaking study on the effects of a P4C program that had been done several years before in the Clackmannanshire Educational Authority in Scotland by Dr Keith Topping and Dr Steven Trickey (Topping & Trickey 2007a). The Scottish study was one of only a few genuine randomised controlled trials that had been done on the effects of a P4C program, and that fact gave its results a solid claim on our attention. Furthermore, the particular P4C program involved (Cleghorn & Baudet 2002) was extremely inexpensive both in terms of instructional time required—one hour per week—and in terms of the cost of materials. Most importantly, the results were remarkable. Not only did the scores on the Cognitive Abilities Test (the CAT in the United Kingdom, the CogAT in the USA) of the students who received the one-hour-per-week philosophy lessons improve significantly over the scores of the control group but, even more interesting, the improvements were still visible in a follow up study done two years later (Topping & Trickey 2007b). Depicted in Figure 1 is the impact of P4C Program in Scotland on students' Cognitive Abilities Test scores after 58 weeks of instruction from 2001 to 2003 and on their test scores two years later in 2005.<sup>1</sup>

---

<sup>1</sup> Trickey, personal communication 2010



**Figure 1:** Students' scores on the CAT from the Scottish study

We decided to replicate the Topping and Trickey study at a local middle school. The school has students in Grade 7, whose usual starting age is 12, and students in Grade 8, whose usual starting age is 13. The study we proposed followed the Scottish study by being concerned with the impact of a structured program of collaborative philosophical inquiry chiefly (but not exclusively) on students' cognitive abilities. In our original plan, the study involved having several seventh grade teachers devote one hour per week to discussions of philosophical topics with their students. The P4C program materials would be the same as those used in Scotland, namely *Thinking through philosophy, Book 4* by Paul Cleghorn and Stephanie Baudet (2002). To measure the impact of those lessons on the students we proposed to compare the pretest and posttest scores on the CAT/CogAT of the students in the treatment group with the CAT/CogAT scores of the students in the control group. We used the CAT/CogAT as the assessment instrument because Topping and Trickey used it in their study. They used it because 'The CAT is the most widely used test of reasoning abilities in the UK, with close to one million students assessed each academic year', and the CAT has a number of separate subtests 'which are aggregated into three batteries, providing standardised measures of verbal, quantitative, and nonverbal reasoning abilities' (Deary, Strand, Smith & Fernandes 2007, p. 14). One particular reason Topping and Trickey gave for using it, is that there is evidence that improvement on the CAT/CogAT score is associated with improved subject matter learning in several areas as measured by students' performance on the General

Certificate of Secondary Education (GCSE) exams given in some parts of the United Kingdom at 15-16 years of age and covering English, Science, Mathematics, History, Geography and more. To corroborate this, Topping and Trickey referenced the 2001 handbook for the CAT Version 3 developed by Smith, Fernandes and Strand (Topping & Trickey 2007a, p. 278).<sup>2</sup> We should note that, since we were using the CogAT, we asked the test developer Dr David Lohman, if the CogAT and the CAT were indeed the same, and he assured us that they were, differing only in minor matters of language.<sup>3 4</sup>

One thing that should be noted is that while this article will present the results of the P4C program on the students' cognitive abilities, the program itself of one hour per week dialogue aims to create 'a community of inquiry'. The impact of forming that community on students' social and emotional behavior was investigated by Trickey and Topping (2006). They report the effects of the P4C program on students' expressions of empathy, self-confidence and emotional self-regulation, and we replicated that aspect of their study and will report those results in a separate article.

The Clackmannanshire educational district is comparable to the local school district in our study in that it is not a wealthy area. Indeed, the median household income for our district in 2009 was \$33,359 in contrast to the overall Texas median income of \$49,433 (United States Census Bureau 2014). Not too surprisingly, 60.2% of the district's students were classified as 'Economically Disadvantaged' (The Texas Tribune 2015). One difference between the two districts is that Clackmannanshire is ethnically homogenous, whereas the local district has a more diverse student population with percentages of students in 2010 at 46% white, 26.9% African American, and 25.8% Hispanic (ProximityOne 2015).

There were two other differences between our proposed project and the Scottish study that we thought could be important. First, one of the authors of the *Thinking through philosophy* teaching materials, Paul Cleghorn, was on site in Scotland as a consultant and, naturally, he would not be available in Texas. Second, and more worrisome, we understood that in Scotland, the students experienced the P4C program for 58 weeks over the course of two school years. In contrast, we would

---

<sup>2</sup> See also Deary et al. 2007 for a report of a massive study relating CAT scores to GCSE scores.

<sup>3</sup> Personal communication

<sup>4</sup> For further information about these tests the reader should visit either <http://www.glassessment.co.uk/products/cat3-cognitive-abilities-test-third-edition> (for the CAT) or <http://www.riverpub.com/products/cogAt/details.html> (for the CogAT).

have to work within the confines of a single school year, so there simply could not be the same number of P4C lessons for the students in our study.

Early in 2010, once we had secured the support of the Principal of the middle school and of the Superintendent of the school district, and once our Institutional Review Board had given the necessary approval, we were fortunate to be able to invite one of the authors of the original study, Dr Steven Trickey, to come to our university to consult with us about how his study had been conducted. Then, with everything in place, we submitted a grant proposal to a foundation in Houston, and we were pleased when the proposal was accepted in June of 2010. One result of having foundation funding in hand was that our university Office of Research and Sponsored Programs subsequently found additional funds, and this influx of new funding effectively doubled the size of the project by allowing us to add the eighth grade students to the study, whereas we had originally planned to conduct the study with seventh grade students only. We then made plans for a teacher-training workshop on one of the teacher in-service days in August 2010 before classes began at the school. The additional funding turned out, as we will describe later, to create a situation in which some very interesting results were obtained.

## **Teacher training**

### *Selecting the teachers*

By agreement with the Principal, we scheduled an all-day teacher-training workshop that would occupy one of the teacher in-service trainings that are held a few days before classes begin in the fall. Instruction at the school is conducted by multidisciplinary teacher teams in each grade, and we determined that the easiest way to randomly assign teachers to the experimental and control groups would be to randomly select the teams. Also, after consulting with the Principal, we decided that within each team it would be the two Language Arts teachers who would be charged with conducting the once-a-week philosophy sessions. We selected the Language Arts teachers because we thought that, on the whole, they would be more accustomed to conducting discussions of contrasting ideas and contrary opinions about pieces of literature. By a stroke of good fortune, Dr Steven Trickey was able to join us for the training day, and we felt that his presence would add a degree of authenticity to our attempt to replicate his study. To keep in touch with the teachers throughout the project we decided to establish an electronic discussion board where the teachers could share their impressions, ask questions, make comments, and receive feedback from our group members as the project unfolded.

The teacher-training day, 18 August 2010 came, and at 8:30 in the morning, after describing the study to the teaching staff assembled in the 'cafetorium', we did a random drawing for the two teams in each grade whose Language Arts teachers would conduct the weekly one hour philosophy discussions in their classes. As a result there were four seventh grade and four eighth grade Language Arts teachers selected.

### *The training*

Immediately after the random selection, the eight Language Arts teachers and the rest of their team members from disciplines such as Math, Science, and History—24 teachers in all—went with us to the classroom in which the P4C training would take place. At the training session, we began from 8.30-9.45 with an introduction by Dr Trickey, who provided information about the study he had conducted and about the notion of establishing a community of inquiry in the classroom. For readers unfamiliar with this notion, the introductory pages of *Thinking through philosophy, Book 4* give this guidance:

A community of enquiry has a rational structure and a moral structure. The former is about how to go about exploring ideas through dialogue ... It includes such aspects as ensuring that participants:

- Ask open and inviting questions
- Give evidence and examples
- Make comparisons
- Summarise and evaluate
- Seek clarification

The moral structure includes the application of emotional intelligence and could be called the 'spirit of enquiry.' It also includes the rules of behaviour necessary for a group activity such as one person being allowed to speak at a time ... ensuring that pupils:

- Focus attention on the speaker
- Don't 'put down' others
- Are not forced to speak
- Respect others' views

- Are truthful
- Are open minded

(Cleghorn & Baudet 2002, p. 4; cited with permission of Paul Cleghorn)

From 10:00-10:50 members of our university P4C group provided an overview of the structure of the Philosophy discussion sessions (PowerPoint slides are available on request from Dr Fair). We covered the seven stages that were to occur in each session (from *Thinking through philosophy: Book 4*):

### **The focusing exercise**

This is a simple but extremely powerful exercise that helps people focus their attention and be 'in the present.' Its simplicity is its strength and also its weakness. It is easy to do, but also easy for children to think they are doing it when in fact they are not. The exercise consists of simply 'giving' attention out through the senses for two or three minutes. This creates a highly alert yet peaceful condition. For the short time (literally two minutes) that it takes to do it, in value for time it is probably the most powerful aspect of the programme!

...

### **Linking with the previous week**

This is simply good practice and reinforces in the mind what has taken place the week before, thereby strengthening memory. It also provides an opportunity for children to bring forward new evidence and experience from during the week. Thinking is not something that only takes place on Wednesdays from ten until eleven!

### **Pair/group work**

This provides an opportunity to check that children have understood the literal meaning of the story or poem, and more importantly, is where there is a planned focus on one or more of the thinking skills. These are built into the programme. Also important about this stage is that it is very 'inclusive.' Children who may not initially speak in a whole class forum can gain confidence that their ideas are important and accepted in a small-scale setting.

### **The stimulus**

The story or poem is read aloud by the teacher.

### **Dialogue**



The key to the whole programme. This is the Socratic method of questioning or 'dialectic.' Prof. Matthew Lipman says this is not 'mere' conversation, but 'an inquiry, an exploration of ideas - a quest. It follows a line of investigation like a detective.' The 'Questions for Thinking' are used to stimulate this process, but it is important to note that these are not to be slavishly followed through to the end if the dialogue is flowing. The section on 'Questioning' in these notes will help build the skill of being a good facilitator of a dialogue. It is not to present yourself as the expert in philosophy (and it is assumed that you are not), or yet to move the group to think as you think. It is to coordinate and enhance the dialogue by modelling open questioning, by encouraging all to take part, by not being content always with first answers and knowing how to allow the child to think more deeply by further questioning. Look at the criteria in the 'Assessment of a Philosophy Session' for further examples. Dialogue is important because it stimulates a deeper engagement between pupils and teacher and can take learning to a deeper level of understanding. It requires the teacher to move from being a dispenser of information to a facilitator of learning - the method can be used to great effect across the curriculum. Dialogue stimulates thinking and emotional intelligence, bringing with it self-confidence.

### **Closures**

These are brief ways of closing the dialogue and in essence, provide a variety of ways of drawing the attention of the children to how their thinking has progressed during the session.

### **Thought for the week**

Each week a practical idea drawn from the theme of the story or poem is highlighted to provide some 'homework.' In the main it is looking for evidence in real situations of an aspect of the theme. This should be encouraged but entirely natural and unforced! Nothing will put children off more quickly. It is useful if the 'Thought for the Week' is displayed on the classroom wall for reference and as a reminder.

(Cleghorn & Baudet 2002, pp. 6-7; cited with permission of Paul Cleghorn)

Next, from 11.00 to 11.50 am we did role playing with the teachers so that they could see concretely what the discussion process would be like. In this 'hands on' Session #1 the group was presented with the two poems, 'Myself' and 'Me' from the teaching

materials in *Thinking through philosophy, Book 4*. The group as a whole went through the seven stages outlined above. Before breaking for lunch six teacher volunteers were recruited to lead the afternoon sessions.

After lunch, we had two more sessions. In hands on Session #2 from 1.10 to 2.00 pm three groups of 12 teachers were assembled at random, each with one of the teacher volunteers as facilitator. The three groups followed the seven stages of the community of inquiry model and dealt with the story 'Walking away' as the discussion prompt. Finally, in Session #3 from 2.10 to 3.00 pm, three different teacher volunteer facilitators, each assisted by a university P4C group member, followed the community of inquiry model in dealing with the prompt provided by the mystery story 'The broken windows'. When we completed the exercises, we adjourned the training session, and we promised the teachers that we would keep in touch with them via a discussion board.

### **Challenges arise**

One thing that immediately went awry was our attempt to keep in communication with the eight language arts teachers (four at the seventh grade level and four at the eighth grade level). Because of the formidable 'fire walls' installed in the school district's computer system, the teachers were unable to join the electronic discussion board from within the school. As a response to this problem, we agreed to visit the school every two or three weeks to meet with the participating teachers during their regular team meetings to answer questions, to give encouragement, and in general to get some sense of how the P4C lessons were going.

Of course, the lessons could not start until after the pretesting with the CogAT was complete. This meant that the philosophy discussions could not begin until mid-September because the CogAT pretest had to be administered to all 800+ students in the school, that is, every student in the seventh and eighth grades, within the first few weeks the school term which began in late August. This presented a logistical challenge both because of the number of students involved and because of the timeframe needed to administer the CogAT. Therefore, this required the test to be administered in two separate sessions. In addition the teachers needed, understandably, to delay starting the program because they wanted to have time to establish a rapport with their classes before trying something new to both them and their students.

When we met with the teachers during the course of the school year, we began to notice a pattern. Several of the teachers, especially those from the seventh grade, met with us regularly, and they reported enthusiastically about how the lessons were going with their students. They told us how they had adapted the lesson material and which of the lessons their students found most interesting. They also commented on various aspects of the structure of the lessons. For example, one element of the lesson structure involved beginning with the relaxation exercise. Some of us were doubtful about the impact of the relaxation exercise, but we were assured by the teachers that, during a busy school day, this was important because it helped the students to look forward to and to enjoy the philosophy sessions.

However, at these meetings we also noticed that some of the teachers, particularly those from the eighth grade, seldom attended and, when they were there, they often did not say much. We were not totally surprised to learn later from one of those teachers that she did not begin the lessons until January 2011. As it turned out, this meant that there was relatively little time for the lessons because in mid-March everything in the school began to be focused on preparation for the state's standardised testing program. The philosophy sessions stopped and, by the time the standardised state testing was completed, all that could be done was to administer the CogAT posttest in May 2011 to all the seventh and eighth grade students.

### **Encouraging comments from surveys of teachers and students**

As the school year was ending we surveyed the participating teachers and, for those teachers who were willing and able, we supplied a survey instrument for their students; an instrument patterned after the one Dr Steven Trickey used in the original study in Scotland. The survey data provided an interesting array of reactions to and perceptions concerning the one-hour-per-week philosophy sessions. Some of the teachers rated their experience with the philosophy sessions as 10 on a ten-point scale with 10 the highest positive rating. They were very pleased with the program, and they saw positive changes in their students' ability to provide reasons for their views and to listen carefully to others' opinions. Here are some of the teachers' comments:

- A seventh grade teacher: 'I was so impressed that my students started incorporating this style of thinking during our regular literature time. They started asking these amazing philosophical questions. Far beyond just about the plot, they started digging deeper.'

- An eighth grade teacher in response to a survey question about highlights of the experience: ‘When some students began encouraging others to speak their minds.’
- A seventh grade teacher: ‘Students became more thoughtful about others’ ideas; they became more talkative and not afraid to express ideas. I really enjoyed participating in the project. I would like to continue to use the lessons with future classes.’
- An eighth grade teacher: ‘It was interesting to see some students develop new and dynamic ways of thinking. When the class was focused on one side of a subject, sometimes other students brought forth an entirely new view on the situation.’

The student reactions ranged from ‘it’s boring’ to one claiming that ‘I hit my sister less’ to ‘thoughtful notes’ about how they had come to better appreciate their classmates’ points of view. Some student comments noted on the survey (with errors included) were:

- ‘I like being able to hear how other people opinion can be the same as yours.’
- ‘I like it because you can really think about things and learn you are really smart if you just think.’
- ‘I think the lessons helped us listen to each other more.’
- ‘In the lessons some people started to become more open and accepting of others opinions.’
- ‘The class as a whole discuss it better instead of arguing about it.’
- ‘I learned how to show facts and find evidence and how to use it during the disputes.’
- ‘It helped me understand that close mindedness is not a good thing.’
- ‘It helps you think about choices you make and view the world different.’
- ‘I learned that to put your opinion out there you have to make sure you know what you’re saying.’
- ‘I learned that you should give people a chance to talk because they might change your mind.’

(Please note that a more complete and detailed presentation of these surveys and the qualitative data will be presented in a later article.)

**A serendipitous outcome: Two experiments instead of one**

While this qualitative data was encouraging, we still did not yet have a quantitative analysis of the impact of the philosophy sessions on students' CogAT scores. A counsellor at the school, who had been supportive of the project from the outset and who was finishing her doctoral degree, agreed to do the statistical analysis. We looked first at the reports of the changes in mean scores for all students in the experimental group (both seventh and eighth grade students who had the philosophy sessions) versus the reports of changes in mean scores for all of the students in the control group (both seventh and eighth grade students who had not had the sessions) and our initial impression was discouraging. There did not appear to be much of a difference between the two groups. However, when the data were analysed by comparing just the seventh grade students who had the sessions with those seventh graders who had not, a striking result emerged. The mean increase in CogAT scores for the seventh grade treatment group was significant and showed a large positive effect comparable to the effect size of 0.75 found by Topping and Trickey. In contrast, for the eighth graders, there was no statistically significant difference to be found between the CogAT scores of students in the treatment group and those in the control group. What follows in the next section is a description of the analyses that showed this divergence in the results for the two different grades and, in the subsequent section, there is an explanation for the divergence.

***Analyses of the quantitative data***

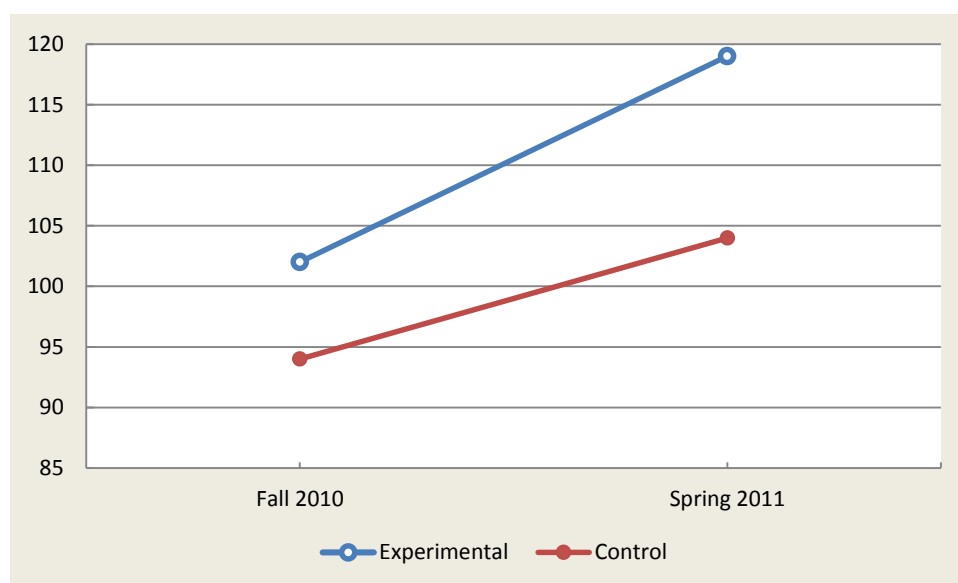
Analyses of data were conducted to determine whether a statistically significant difference was present between the pretest and posttest scores of experimental group and control group participants in the seventh and eighth grade, as well as the presence of a statistically significant difference between the pretest and posttest scores among groups for each grade level. Due to differences in group sizes parametric analysis of covariances (ANCOVA), were conducted to determine whether differing sample sizes among the two groups of participants was a factor in the outcome of posttest scores. Levene's Test of Equality of Error Variances was violated for each analyses and caution may be noted, however the ANCOVAs were performed to control for the effect of variance in group membership and is considered a robust statistical procedure (Field 2009).

### *The pretest and posttest scores for the seventh and eighth grades combined*

The difference in posttest scores for the combined experimental group of seventh and eighth grade students and for the combined seventh and eighth grade control group was statistically significantly different,  $p < .001$ , after controlling for the potential effects the covariate may have on the outcome, revealing a small effect size,  $\eta^2 = .07$  (Cohen 1988). Refer to Table 1 and Figure 2 for descriptive statistics and differences in pretest and posttest scores for the combined groups.

**Table 1:** Descriptive statistics for pretest and posttest scores for seventh and eighth grade students combined.

Participants	Pretest - <i>M</i>	Posttest- <i>M</i>	Pretest- <i>SD</i>	Posttest- <i>SD</i>
Experimental ( $n = 363$ )	102.19	119.38	32.69	31.74
Control ( $n = 177$ )	93.86	104.23	36.99	35.32



**Figure 2:** Combined seventh and eighth grade pretest and posttest scores on the CogAT

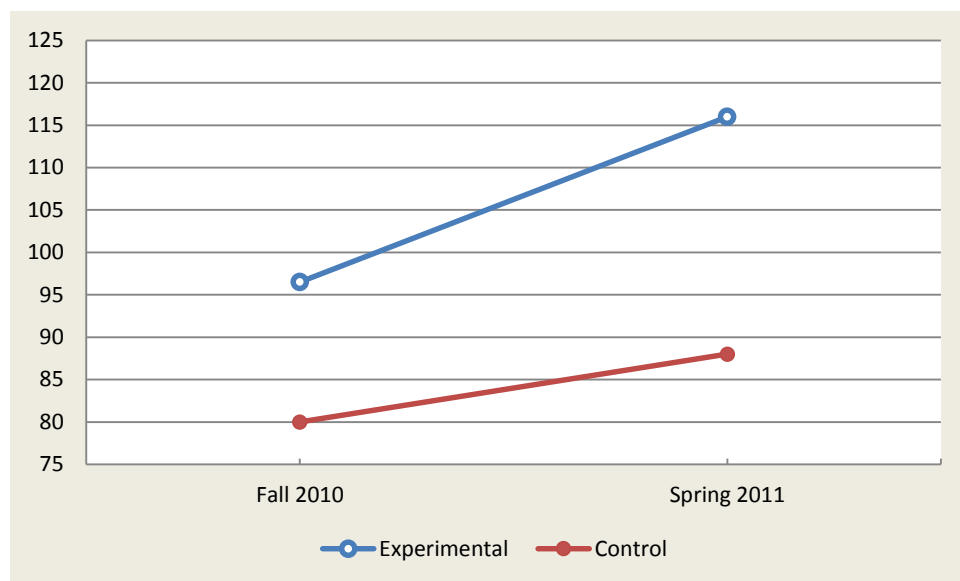
Further analyses were conducted to determine whether a statistically significant difference was present between pretest scores and posttest scores of experimental group and control group participants within each grade level. A statistically significant difference was noted for seventh grade students,  $p < .01$ , revealing a moderate effect,  $\eta^2 = .21$ , (Cohen 1988). However, a statistically significant

difference in scores for eighth grade students was not indicated,  $p = .19$ . Presented in Table 2 are the pretest and posttest scores for seventh and eighth grade students.

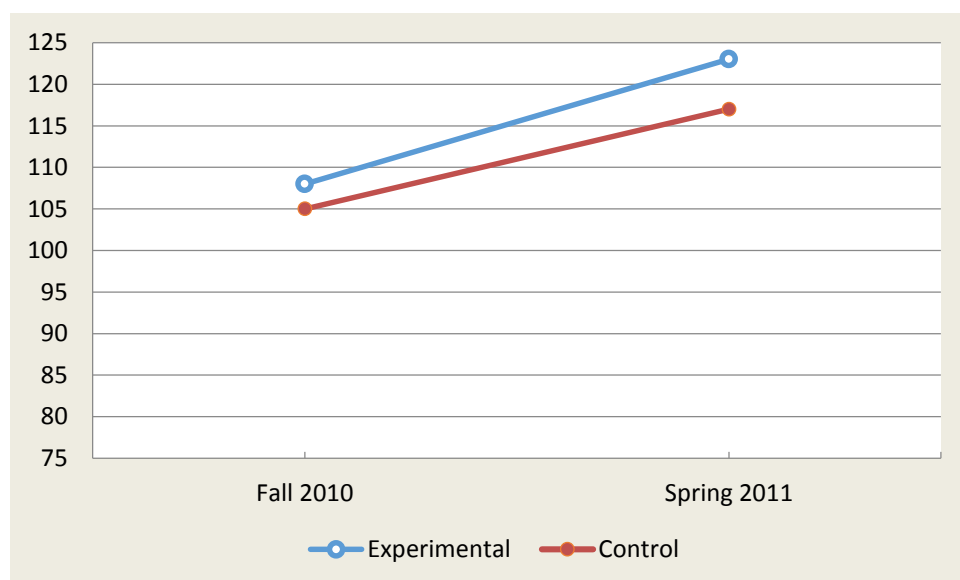
**Table 2:** Descriptive statistics for pretest and posttest scores by grade level

Participants	Pretest - <i>M</i>	Posttest- <i>M</i>	Pretest- <i>SD</i>	Posttest- <i>SD</i>
Seventh Grade				
Experimental ( $n = 186$ )	96.53	116.13	31.14	28.74
Control ( $n = 79$ )	80.15	88.05	38.10	32.48
Eighth Grade				
Experimental ( $n = 177$ )	108.14	122.79	33.32	34.36
Control ( $n = 98$ )	104.91	117.28	32.23	32.1

To determine the extent of differences in pretest and posttest scores within the two groups, parametric dependent samples *t*-tests were conducted for both groups of seventh grade participants. For participants in the control group the paired samples analysis yielded a statistically significant difference in pretest and posttest scores,  $t(78) = -2.56$ ,  $p = .01$ , Cohen's  $d = 0.22$ , a small effect size. For participants in the experimental group a statistically significant difference in pretest and posttest scores was indicated,  $t(185) = -15.40$ ,  $p < .001$ , Cohen's  $d = 0.65$ . According to Cohen (1988) the effect size for the difference was moderate to large. In reference to pretest and posttest scores for seventh grade students, a greater difference in posttest scores was present for participants in the experimental group than for participants in the control group. Presented in Figures 3 and 4 are the differences in pretest and posttest scores for each group by grade level. Additionally, the length of time the experimental groups participated in the philosophy discussions is noted, indicating significant differences in participation between the seventh and eighth grade groups.



**Figure 3:** Seventh grade CogAT scores after 22-26 weeks of P4C



**Figure 4:** Eighth grade CogAT scores after 4-10 weeks of P4C

#### *A plausible explanation for the different outcomes*

After further examination of the differences in pretest and posttest scores, a very plausible explanation for these differing outcomes quickly emerged. The counsellor met individually with each of the teachers—the four in the seventh grade and the four in the eighth grade—to confirm how many weeks each had been able to devote to the philosophy sessions. Their responses tell the tale: the four seventh grade teachers did 22, 22, 24, and 26 weeks of sessions, while the four eighth grade teachers



did the sessions for only 4, 5, 8, and 10 weeks. The informal meetings the counsellor had with the teachers also provided additional insight. The teachers who devoted time weekly to the P4C sessions stated that the students loved the discussions and couldn't wait each week to participate. Additionally, the teachers who included the sessions weekly were excited to see the students so engaged and eager to share their thoughts and listen to opinions of others. The teachers who did not include the sessions in their weekly routine shared that they did not feel comfortable with the open discussions and felt a lack of control in the classroom. Noteworthy is that two teachers were new to the classroom and felt they were not prepared to have the open discussions where they didn't have a specific plan. Another interesting response was from a veteran teacher who shared that she likes things to be very structured in the classroom and was not comfortable with students in control and having to develop their own opinions and that there was not a correct answer or outcome at the end of the sessions. This was an interesting response as the teachers who were enthusiastic about the discussions felt that was the best part, that the students had their own ideas and were able to critically think about their responses.

Obviously, no one would expect that just a few weeks of the P4C program would have a marked effect on students and, since it would take some repetition to accustom students to the practices of giving reasons for their views and of critically evaluating the reasons that they and others give for contrasting views, we would not be surprised to find a threshold for the minimum number of times that a typical student has to engage in those practices before they become consolidated.

### **Conclusions and looking ahead**

We believe that these results have real educational significance. When we began this study as a replication of the Scottish study, we knew at the outset that there was one potentially important condition we could not replicate. The Scottish study was conducted over a period of a year and a half and Dr Trickey has confirmed to us that this meant there were 58 weeks of philosophy sessions. In contrast, we had to confine ourselves to the standard American school year, and that meant that, even in the best case, there would be far fewer than 58 sessions. So we are greatly encouraged by these results because they appear to show that a program of regular, weekly, structured one-hour philosophy group sessions can have an educationally important impact—with far fewer than 58 sessions! This, naturally, makes such a program all the more likely to be an instructional option that can be a practical one in the American school setting.

We believe that our study provides evidence that supports the main contentions of the study by Topping and Trickey, namely that regular, one hour per week, structured community of inquiry P4C sessions are a relatively powerful educational intervention which boosts students' cognitive abilities significantly while doing so at a very small cost both in materials needed and in instructional time. Let us stress this last point by noting:

1. Staff development time needed is minimal. One day of in-service sufficed.
2. The materials are inexpensive. For example, at the beginning of 2013 a downloadable pdf copy of *Thinking through philosophy, Book 4* can be purchased from Educational Printing Services in the UK for £17.50 (about \$25.00 USD).
3. The P4C lessons require only one hour per week of instructional time.
4. The extent of time required for the program to have an effect is reasonable—25 weeks or so; well within what can be done in one standard school year.

Additionally, comments from students and teachers showed that this intervention was popular with them. There was no testing over the material, just relatively relaxed conversation about ideas, about differing perspectives, and about how strong the reasons were for thinking that one idea is better than another.

One may wonder at this point, however, about why this simple and relatively cheap intervention has the effects that our study and Topping and Trickey's Scottish study showed. We can hazard an opinion that the impact is due to the confluence of (a) the philosophical subject matter and (b) the structured discussion process that establishes a community of inquiry. Philosophical questions that relate to deep topics are notoriously open to multiple answers from a variety of perspectives but, at the same time, those multiple answers are not to be simply immediate reactions. They are to be *considered* opinions—opinions which are supported by reasons, and those reasons in turn are to be subjected to public scrutiny to see how well they do, in fact, support the opinions offered. To make this more concrete, the first story in the *Thinking through philosophy, Book 4* (Cleghorn & Baudet 2002) is 'Walking away', and in it a boy, Barry, is with a new friend, Alex, when Alex shoplifts a mobile phone. This presents a moral dilemma for Barry, and the invitation to the students is to discuss various aspects of the situation. Accordingly, Cleghorn and Baudet suggest questions for thinking such as 'What were the arguments going on in Barry's mind after he had seen Alex hide the phone in his jacket?' and 'When he was in the shop Barry had to decide what he was going to do about the situation. Has anyone ever had a similar decision to make?' (p. 24).

So the philosophical subject matter of the pieces that are used to prompt discussion lends itself to the students feeling entitled to have and to explain their own individual views on the matter. This all relates to what Cleghorn and Baudet call the 'rational structure' of the community of inquiry. But equally important is what they term the 'moral structure' of the community. Students are led to understand that they are not to 'put down' others and they are to respect others' views, to be truthful and open-minded (p. 4). At the same time the students get practice in giving and critiquing reasons, their own and others, and doing so in public. Thus, after everyone has been given some thinking time, the students are put into pairs where they share their thinking, give reasons for their choices, and are able to challenge those reasons, provided their challenges themselves are based on evidence. After talking in pairs, the whole class joins in sharing their thoughts.

With this activity occurring regularly over many weeks, it seems likely that a number of the students will grow in confidence in their thinking abilities, and all of them will gain practice in using reflective critical thinking skills. These are the sorts of skills that psychologists such as Keith Stanovich (2009) and Daniel Kahneman (2011) have studied extensively; skills that can be acquired and that comprise much of what they label 'Type 2' thinking. It does not seem much of a stretch then to assume that, when confronted by the problems to be solved in the CogAT, these more practiced reasoners are better equipped to solve them. Accordingly, we believe the promise of this approach to enhancing students' reasoning abilities is bright.

## References

- Cleghorn, P & Baudet, S (2002) *Thinking through philosophy, Book 4*. Educational Printing Services, Blackburn, UK.
- Cohen, J (1988) *Statistical power analysis for the behavioral sciences*. 2nd edn, Lawrence Erlbaum, Hillside, NJ.
- Deary, I, Strand, S, Smith, P & Fernandes, C (2007) Intelligence and educational achievement. *Intelligence*, 35, pp. 13-21.
- Field, A (2009) *Discovering statistics using SPSS (Introducing statistical methods)*. SAGE, Thousand Oaks, CA.
- Kahneman, D (2011) *Thinking, fast and slow*. Farrar, Straus and Giroux, New York, NY.

- ProximityOne (2015) *Huntsville ISD, TX (4824030): DP1 general demographic characteristics*. Available from [http://proximityone.com/acs/dptx/dp1\\_4824030.htm](http://proximityone.com/acs/dptx/dp1_4824030.htm)
- Stanovich, K (2009) *What intelligence tests miss: The psychology of rational thought*. Yale University Press, New Haven, CT.
- The Texas Tribune (2015) *Public schools explorer: Huntsville ISD*. Available from <http://www.texastribune.org/public-ed/explore/huntsville-isd/>
- Topping, K & Trickey, S (2007a) Collaborative philosophical enquiry for school children: Cognitive effects at 10–12 years. *British Journal of Educational Psychology*, 77, pp. 271–288.
- Topping, K & Trickey, S (2007b) Collaborative philosophical inquiry for schoolchildren: Cognitive gains at 2-year follow-up. *British Journal of Educational Psychology*, 77, pp. 787-796.
- Trickey, S & Topping, K (2006) Collaborative philosophical enquiry for school children: Socio-emotional effects at 11-12 years. *School Psychology International*, 27(5), pp. 599-614.
- United States Census Bureau (2014) *State median income*. Available from <http://www.census.gov/hhes/www/income/data/statemedian/>