Socrates in the Schools: Gains at Three-year Follow-up

Frank Fair

psy\_fkf@shsu.edu

Lory E Haas

leh015@shsu.edu

Carol Gardosik

gardosikc@stthom.edu

Daphne Johnson

ddjohnson@shsu.edu

Debra Price

EDU\_DPP@shsu.edu

Olena Leipnik

SOC\_OVL@shsu.edu

Frank Fair, Lory Haas, Daphne Johnson, Debra Price, and Olena Leipnik are all from Sam Houston State University, Texas

Carol Gardosik is at the University of St. Thomas, Texas

Socrates in the Schools: Gains at Three-year Follow-up

**I. Introduction**

When our Philosophy for Children (P4C) study group decided to replicate Topping and Trickey’s study of the impact of a P4C program on students in Clackmannanshire, Scotland (Topping & Trickey, 2007a), we had several reasons for our choice. First, Topping and Tricky used the Cognitive Abilities Test (CogAT in the US, CAT in the UK) as their tool to measure the impact of the P4C program, and thus they used an instrument that was well-regarded and widely-studied. Scores on the CogAT/CAT have been the subject of much study, most notably, of how scores on it correlate well with scores on General Certificate of Secondary Education (GCSE) exams given in a variety of subjects such as English, Science, Mathematics, and History to 15-16 year-old students in the UK (Deary et al., 2007). Second, Topping and Trickey’s study employed a control group and took care to approximate the conditions of a randomized controlled clinical trial for the CogAT/CAT pretests and posttests. Third, it was very important that the P4C intervention they employed was *cheap*. It was cheap both in terms of instructional time—one hour per week—and in terms of the materials used--*Thinking Through Philosophy* --at about $25.00 USD per copy. Nor did it take extensive and expensive staff preparation.

In addition to the materials for reflection provided by *Thinking Through Philosophy* *Book 4* (Cleghorn & Baudet, 2002), the classroom teachers in Topping and Trickey study followed a lesson plan that involved seven stages The lessons began with (1) a focusing exercise to create a relaxed, meditative frame of mind and (2) a brief link with the previous week’s discussion. (3) Then came the stimulus for discussion—usually reading a story from *Thinking Through Philosophy*. (4) Next the students work in pairs discussing the story and reflecting on some open-ended questions suggested by the story. (5) This was followed by dialogue in larger groups where the teacher has encouraged the students to form a *community* *of inquiry* by

(a) communicating their views in response to the questions at hand,

(b) supporting their views with reasons,

(c) listening respectfully to the views of others,

(d) indicating whether they agree or disagree with those views,

(e) providing alternative viewpoints, and

(f) gradually developing a process of dialogue.

(6) Finally, the teacher brings closure by encouraging the students to reflect on the discussion and how their thinking might have progressed, and (7) by providing a “thought for the week” that highlights an idea to serve as a basis for “homework” to be reflected on in order to relate the idea to situations outside the original stimulus (Topping & Trickey, 2007b; for more see Cleghorn & Baudet, 2002).

So when Topping and Trickey reported achieving substantial pretest to posttest gains in CogAT/CAT scores for their experimental group versus their control group, those results were worth serious attention. But then Topping and Trickey did what few others had done, namely, they conducted a follow-up study two years later. This was important because of the tendency for the impact of many educational interventions to fade over time (Cascio & Staiger, 2012). In their follow-up study Topping and Trickey were clearly able to document the persistence and durability of the effects of the P4C intervention, and that was very encouraging (Topping & Trickey, 2007b). The results of the original study and the two year follow-up can be seen in this graph:

***Figure 1:***

**II. Our Initial Study**

Our study of the impact of the same P4C program on 7th graders (12-13 years old), took place during the school year of 2010-2011. We structured the program to emulate, as closely as we could, the intervention that Topping and Trickey employed in their study. Thus we used the same materials from *Thinking Through Philosophy Book 4,* and we trained the teachers to follow the same 7-stage process for each discussion. Additionally, we were fortunate to have Dr. Trickey consult with us and take part in the teacher-training session held shortly before school began in late August of 2010.

In addition to the program with the 7th graders, we were able, thanks to some additional funding, to extend the program to the 8th graders, ages 13-14 at the school. As we reported earlier (Fair et al., 2015), because the 8th grade teachers used the program for a relatively short period of time (4-10 weeks) and the 7th grade teachers, by contrast, used the program much longer (22-26 weeks), the result was, in effect, two different studies. The 8th grade study yielded little evidence of any impact on CogAT/CAT scores, whereas the study of the 7th graders yielded evidence of a significant impact. Here are the graphs of the pre-test and post-test CogAT/CAT test scores for the 7th and 8th graders:

***Figure 2:***

***Figure 3:***

As one can see--and the statistical analysis bears this out--the 22-26 weeks of one-hour per week Philosophy discussions in the 7th grade had an appreciable impact, whereas the 4-10 weeks of discussions in the 8th grade did not (Fair et al., 2015).

Naturally, we were happy that our results for the 7th graders replicated the results that Topping and Trickey had obtained (Topping & Trickey, 2007a), and we were not too surprised that in the 8th grade we did not see those results, presumably because so many fewer weeks were used for discussions. We also were happy that 22-26 weeks was a sufficient period of time to produce the results, because that length of time could be easily accommodated in one school year in the USA. In contrast, the original Topping and Trickey study concerned the effects of a 58-week program of weekly discussions—somewhat over twice as many weeks as in our study. That led us to wonder whether the results of our program, one involving substantially fewer weeks of discussions, would have a similarly durable impact, or whether the results would be more susceptible to fading over time.

**III. Our Three-year Follow-up Study**

Fortunately, we were also able to conduct our own follow-up study. A large number of students who had been 7th graders during the original study in 2010-2011 were sophomores, that is, they were in their second year at the local high school, during the 2013-2014 school year. The local school district administered the CogAT/CAT to all of the sophomore class students, and we were allowed to acquire data from the school district archives, specifically the test scores for those sophomores who had either been in the experimental group or the control group in the original study. Of course, the scoring of the CogAT/CAT was blind as far as the students’ participation in our original study was concerned. When the scoring was done, the result was that we had the 10th grade CogAT/CAT scores for 133 out of the original 186 members of the experimental group and 50 out of the original 79 members of the control group. These numbers compare well with the numbers that Topping and Trickey were able to obtain for their two-year follow-up--71 out of 105 from their experimental group and 44 out of 72 from their control group (Topping & Trickey, 2007b).

Besides wondering about the difference in impact that 58 weeks versus half that time might make, we also wondered if having our follow-up at the three-year mark, instead of two years for the Trickey and Topping study, might allow for greater “fading” of the results. But, as it turned out, we need not have worried. Here is what we found.

*Analyses of the quantitative data*

Analyses of data were conducted to determine whether a statistically significant difference was present between the pretest scores of seventh grade students and posttest scores of experimental group and control group participants three years later as tenth grade students. You will note that the means are slightly adjusted than as plotted on the previous figure for the first and second round of testing for seventh grade students. This is due to changes in the number of participants in the final round of testing for comparison of all three rounds of testing.

To determine the extent of differences in pretest and posttest scores within the two groups, parametric dependent samples *t*-tests were conducted for pretest scores of both groups as seventh grade participants and posttest scores of both groups as tenth grade students. For participants in the control group the paired samples analysis yielded a statistically significant difference in pretest and posttest scores, *t*(49) = -2.67, *p* = .01, Cohen’s *d* = 0.28, a small effect size. For participants in the experimental group a statistically significant difference in pretest and posttest scores was indicated, *t*(132) = -10.28, *p* < .001, Cohen’s *d* = 0.68. According to Cohen (1988) the effect size for the difference was moderate to large. In reference to pretest and posttest scores for tenth grade students, a greater difference in posttest scores was present for participants in the experimental group than for participants in the control group.

**Table 1:** Descriptive statistics for pretest and posttest scores by group

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Participants** | **Pretest -*M*** | **Test 2-*M*** | **Posttest-*M*** | **Pretest-*SD*** | **Test 2-*SD*** | **Posttest-*SD*** |
| Experimental (*n* = 133) | 100.09 | 117.25 | 122.53 | 30.41 | 27.59 | 35.25 |
| Control (*n* = 50) | 89.60 | 92.4 | 100.26 | 37.40 | 31.64 | 39.09 |

***Figure 4:***

**IV. Discussion**

In the article describing the results of their follow-up study, Topping and Trickey write: “This study provides evidence that gains in cognitive ability which opened up over the intervention period in primary school can be maintained over the subsequent 2 years into secondary school in the same measure, even when pupils have not had any further experience of collaborative inquiry in the secondary school” (Topping & Trickey, 2007b, pp. 793-794). Our follow up study strengthens and extends their work. First, our study strengthens their results by showing that the program of collaborative philosophical discussions that they used could have the same powerful effect on students who were not only half a world away from Scotland but, furthermore, were more ethnically diverse. (Fair et al., 2015) Second, it strengthens their results by showing that, instead of 58 weeks, a smaller number of weeks (22-26) would suffice to produce the effects. Third, our study demonstrates the portability of the program since, while Dr. Trickey did help with the teacher training day, neither he, nor Dr. Topping, nor Paul Cleghorn, the author of the discussion materials, were present during the weeks the program ran. The discussions were conducted in the local public school seventh grade classrooms by regular seventh grade language arts teachers. Neither they nor the students were hand picked, but they succeeded wonderfully nonetheless. Finally, the fact that we could detect a significant difference between the two groups of students after three years, not just two, gives greater credence to the likelihood that the intellectual effects produced by the one-hour-per-week philosophy discussions are as durable as one could wish.

**V. Conclusion**

Philosophy for Children was inspired by thinkers such as John Dewey and Lev Vygotsky, but the lion’s share of the credit in recent years must be given to the pioneering work of Matthew Lipman who, along with colleagues such as Margaret Sharp and Frederick Oscanyan, established the Institute for the Advancement of Philosophy for Children (IAPC) at Montclair State University. (See Lipman et al., 1980, and Lipman, 2003, for further background.) Since the IAPC was founded, other programs have come into existence, both in the USA and abroad, but Philosophy for Children remains, in the American context, a very peripheral enterprise as far as the public schools are concerned. We hope that by confirming and strengthening the results demonstrated in Topping and Trickey’s seminal study, that we can give new energy to the movement to incorporate philosophical discussion in the schools in our home state and the rest of the nation. The powerful intellectual impact of the one-hour-per-week philosophy discussions is clear. But, importantly, discussing philosophical issues is not just about sharpening students’ minds, it is about nurturing their spirits as well.

**References**

Cascio, E. & Staiger, D. (2012) ‘Knowledge, tests, and fadeout in educational interventions’, National Bureau of Economic Research Working Paper 18038. Retrieved from <http://www.nber.org/papers/w18038> on June 15, 2015.

Cleghorn, P. & Baudet, S. (2002) *Thinking through philosophy, Book 4*. Blackburn, UK: Educational Printing Services Ltd.

Cohen, J. (1988) *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillside, NJ: Lawrence Erlbaum.

Deary, I., Strand, S., Smith, P., & Fernandes, C. (2007) ‘Intelligence and educational achievement’, *Intelligence*, 35, pp. 13-21.

Fair, F., Haas, L., Gardosik, C., Johnson, D., Price, D., Leipnik, O. (2015) ‘Socrates in the schools from Scotland to Texas: Replicating a study on the effects of a philosophy for children program’, *Journal of Philosophy in Schools*, 2(1), pp. 18-37.

Field, A. (2009) *Discovering statistics using SPSS (Introducing statistical methods).* Thousand Oaks, CA, SAGE.

Lipman, M., Sharp, A, & Oscanyan, F. (1980) *Philosophy in the Classroom*, 2nd ed*.* Philadelphia, PA: Temple University Press.

Lipman, M. (2003) *Thinking in Education* (2nd ed.). Cambridge, UK: Cambridge University Press.

Topping, K. & Trickey, S. (2007a) ‘Collaborative philosophical enquiry for school children: Cognitive effects and 10-12 years’, *British Journal of Educational Psychology*, 77, pp. 271-288.

Topping, K. & Trickey, S. (2007b) ‘Collaborative Philosophical enquiry for schoolchildren: Cognitive gains at 2-year follow-up’, *British Journal of Educational Psychology*, 77, pp. 787-796.