Research in Science Education

Bringing CASE in from the Cold: The Teaching and Learning of Thinking

--Manuscript Draft--

Manuscript Number:	RISE-D-14-00081
Full Title:	Bringing CASE in from the Cold: The Teaching and Learning of Thinking
Article Type:	Manuscript
Keywords:	thinking skills; metacognition; cognitive conflict; pedagogy
Corresponding Author:	Mary Colette Oliver, PhD University of Nottingham Nottingham, UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Nottingham
Corresponding Author's Secondary Institution:	
First Author:	Mary Colette Oliver, PhD
First Author Secondary Information:	
Order of Authors:	Mary Colette Oliver, PhD
	Grady Venville, PhD
Order of Authors Secondary Information:	
Abstract:	Thinking Science is a two-year program of professional development for teachers and thinking lessons for students in junior high school science classes. This paper presents research on the effects of the Thinking Science on students' levels of cognition in Australia. The research is timely with the general capability of critical and creative thinking in the newly implemented F-10 curriculum in Australia. The design of the research was a quasi-experiment with pre and post-intervention cognitive tests conducted with participating students (n = 654) from nine cohorts in seven high schools. Findings showed significant cognitive gains compared with an age matched control group over the length of the program. Noteworthy, is a correlation between baseline cognitive score and the school's Index of Community Socio-Educational Advantage (ICSEA). We argue that the teaching of thinking be brought into the mainstream arena of educational discourse and the principles from evidence-based programs such as Thinking Science be universally adopted.

Bringing CASE in from the Cold: The Teaching and Learning of Thinking

Mary Oliver,

School of Education, The University of Nottingham

and

Grady Venville

Graduate School of Education, University of Western Australia

Key words: thinking skills, metacognition, cognitive conflict, pedagogy

Corresponding Author
Mary Oliver

The University of Nottingham
School of Education
Jubilee Campus
Wollaton Road
Nottingham NG8 1BB

tel: +44 (0)115 748 4791
fax: +44 (0) 115 846 6600
mary.oliver@nottingham.ac.uk

Acknowledgements
This research was supported by a grant from the Australian Research Council (DP1093877). The ideas presented in the paper are those of the authors and not the funding institution. We acknowledge the late Professor Philip Adey whose wisdom and encouragement over the years of this research was invaluable.

Bringing CASE in from the Cold: The Teaching and Learning of Thinking

Abstract
Thinking Science is a two-year program of professional development for teachers and thinking lessons for students in junior high school science classes. This paper presents research on the effects of the *Thinking Science* on students' levels of cognition in Australia. The research is timely with the general capability of critical and creative thinking in the newly implemented F-10 curriculum in Australia. The design of the research was a quasi-experiment with pre and post-intervention cognitive tests conducted with participating students (n = 654) from nine cohorts in seven high schools. Findings showed significant cognitive gains compared with an age matched control group over the length of the program. Noteworthy, is a correlation between baseline cognitive score and the school's Index of Community Socio-Educational Advantage (ICSEA). We argue that the teaching of thinking be brought into the mainstream arena of educational discourse and the principles from evidence-based programs such as *Thinking Science* be universally adopted.

Key words: thinking skills, metacognition, cognitive conflict, pedagogy

Introduction

While critical and creative thinking are dispositions that are desirable in students across all subject areas, teachers' pedagogical expertise for developing these dispositions within their students is capricious. Moreover, evidence for popular approaches to the teaching and learning of thinking in schools and classrooms is often non-existent (Adey, 2012), discredited (Stephenson, 2009) or lack the "standardized and intervention-specific outcome measures" (Burke & Williams, 2008, p. 104) that evidence effectiveness. The lack of clarity that surrounds the term 'thinking skills' is problematic, particularly when curriculum documents specify that these are cross-curricular or core to teaching and learning programs. Familiar to many teachers and educators is Bloom's taxonomy of thinking skills, which suggest a hierarchy of thinking patterns, from knowledge, comprehension through to synthesis and evaluation. Indeed, teachers will recognize that more difficult questions for students tend to be those requiring explanations, understanding and application of concepts rather than recall. Demands for students being able to demonstrate a deep understanding of science subjects has led to the call for "less about what and more about how" (Leyser, 2014, p. 45).

The focus of this paper is the impact on Australian students of a cognitive acceleration or thinking program involving teacher professional learning and a classroom intervention. Over two years, the professional learning was targeted at schoolteachers of science to develop their theoretical understanding and pedagogy in teaching thinking skills to their students. The Cognitive Acceleration through Science Education (CASE) program was originally developed at King's College, London, in the United Kingdom (UK) and published commercially as *Thinking Science* (Adey, Shayer & Yates, 1990). The *Thinking Science* intervention has accumulated significant evidence of effects, both on students' cognitive development and school achievement over the last three decades (for example, Adey & Shayer, 1990; Babai & Levit-Tori, 2009; Endler & Bond, 2008; Author et al., 2012; Shayer, 1999). The findings have shown that it is possible to improve high school students' achievement in science, with evidence of long-term and far-transfer effects (Shayer, Adey & Wylam, 1981, Shayer, 2000).

While the *Thinking Science* program was developed some time ago, the general support and knowledge of the importance of developing thinking skills in students remains high in England. For example, in a current review of the English national curriculum, The Department for Education acknowledge that "improving students' thinking and reasoning skills is of high interest to teachers" (Department for Education, 2012, para 1). In Australia, critical and creative thinking is a cross curricular general capability in the newly implemented F-10 national curriculum (ACARA, 2012). The *Australian Curriculum* clearly states that the development of thinking skills, together with the imparting of knowledge, are the primary purposes of education and that critical and creative thinking are embedded across all learning areas. However, in Australia there are few professional learning programs for teachers to support their implementation of this new cross curricular general capability and uncertainty as to what is meant by critical and creative thinking. Some schools have used this opportunity to implement 'brain-based' programs in order to develop thinking in students in the absence of evidence (see for example, Stephenson's commentary on Brain Gym®, 2009).

The purpose of this paper is thus threefold: 1. To describe the implementation of the Cognitive Acceleration through Science Education (CASE) or *Thinking Science* program in Australia; 2. To detail the relationship between the cognitive levels of student and the school's Index of Community Socio-Educational Advantage (ICSEA); 3. To present data showing the impact of the program on students' cognitive development.

The Challenges for Thinking Programs

Prevalent in educational institutions are a number of myths regarding classroom-based thinking programs, activities and approaches that are supposedly related to research on the brain (Adey, 2012; OECD, 2007,). For example, teachers and curriculum materials often arrange lessons around different learning styles that students might have including visual, auditory or kinaesthetic; or around multiple intelligences including logical-mathematical, spatial, linguistic, musical, or interpersonal intelligences. Even when a person's preferred learning style is used, there is no evidence of educational improvement (Pashler, 2008). Mainstream psychology has consistently provided considerably more evidence to support a high correlation between different aspects of intelligence, or a general intelligence quotient, *g*, rather than multiple intelligences (Visser et al., 2006). By contrast, thinking programs that include the development of metacognition in students is effective at raising student achievement (McGuinness, 1999, Higgins et al., 2005, 2007).

Despite our concerns about teachers' use of classroom pedagogies for which there is little evidence (Stephenson, 2009), detailed analyses of the large body of literature in the field of education indicate that a limited number of programs do improve students' thinking and the performance of students on cognitive and curriculum-based tests (Higgins, Baumfield, & Hall, 2007). One of these programs is the Philosophy for Children (P4C) program developed in the US by Matthew Lipman (1976) which engages children in philosophical inquiry in a collaborative manner to ensure the development and growth of 'reasonableness'. By 'reasonableness', Vansieleghem and Kennedy (2011) claim that the "emphasis is on analytical reasoning as a guarantee for critical thinking" (p. 177). The P4C program requires students to participate in non-judgmental dialogue, thinking, listening and reflecting; activities that are quite different from the passive listening and copying of notes that often results from a traditional didactic approach to teaching and learning.

An example of a program involving the stimulation of cognition at the tertiary level that is supported by published evidence was introduced by the Physics Nobel Laureate, Carl Wieman. Wieman criticises teaching and learning that is dominated by the memorization of facts and information and suggests teachers address key pedagogical strategies: "reducing cognitive load ... addressing beliefs and stimulating and guiding thinking" (Wieman, 2007, p. 13). Large effect sizes were reported when comparisons were made between student learning outcomes from a traditional lecture and a teaching and learning program grounded in Wieman's application of cognitive psychology and physics education. The conclusion that "deliberate practice teaching strategies can improve both learning and engagement in a large introductory physics course" (Deslauriers, Schelew, & Wieman, 2011, p. 864) augurs well for improving learning at the tertiary level.

The Theory and Pedagogy of the *Thinking Science* Program

The *Thinking Science* intervention involves 30 'thinking' lessons delivered over two years, usually about one every two weeks during school term. In the UK the program is implemented in Year 7 and Year 8, the first two years of secondary school when students are between 11 and 13 years of age. Each thinking lesson focuses on a specific reasoning patterns (or schemata) including controlling variables, ratio and proportionality, compensation and equilibrium to analyse process, correlation, probability, classification, formal models of thinking and compound variables. Groups of lessons spiral through increasing levels of complexity related to the reasoning patterns.

The theoretical framework underpinning *Thinking Science* was strongly influenced by the developmental psychology of Piaget (Shayer, 2002) and the socio-cultural psychology of Vygotsky (Moll, 1990). *Thinking Science* lessons each have five central stages or pillars: 1. concrete preparation, 2. cognitive conflict, 3. social construction, 4. metacognition, and 5. bridging (Shayer, 2003). *Concrete preparation* involves the teacher describing the problem, setting the scene, and clarifying the vocabulary relevant to the thinking lesson. For example, in a lesson exploring the relationship between the variables of electric current and thickness of wire, some exploratory 'talk' about what is meant by current helps focus the students' thinking about what to measure rather than the nature of an electric current. Data are collected during this phase, and students and teachers often refer to this as the 'doing' part of the lesson.

Cognitive conflict is a deliberately introduced, non-intuitive element of the lesson that is surprising for the students because it does not make sense when they use their current thinking patterns to try to understand the phenomena. Cognitive conflict is considered the driver of cognitive growth because a mental struggle is required by the students to move beyond their current ways of thinking. For example, one activity in *Thinking Science* firstly helps students to establish a relationship between two variables and then presents them with data where no relationship can be identified. Student cognition is stimulated by this moderately difficult intellectual challenge which is accompanied by group questioning, discussion, and problem solving drawing on the Piagetian idea of equilibration and the Vygotskian idea of a zone of proximal development (ZPD). *Social construction* occurs as students work together in small groups in an attempt to solve the challenge then sharing the development of ideas and explanations in a whole class discussion. Teachers are pivotal in facilitating the whole class discussion, asking for contributions from all groups. At various points throughout the lesson, teachers ask specific *metacognition* questions to develop students' abilities to reflect on their own and each other's thinking.

Metacognition is about the students becoming aware of how they were thinking and how others were thinking when they discussed and/or solved the problem, and aware of what they learned that is different to what they understood and could do prior to the lesson. Finally, the *bridging*, or transfer part of a Thinking Science lesson is used by teachers to relate the reasoning pattern to everyday science lesson, or real life. For example, having worked through the lessons on probability in *Thinking Science*, teachers might discuss with students the probability of getting lung cancer from smoking, or they might actively transfer the thinking patterns learnt into genetics when students are solving Mendelian genetics problems that require an understanding of probability. Sometimes the pillars of cognitive acceleration are discernable as discrete and sequential within a particular lesson, however, frequently they are highly integrated. Anecdotal evidence suggests that as teachers become skilled at using the pillars they adopt them in their regular science lessons and provide opportunities for students to draw upon the problem-solving strategies and ways of thinking developed during the *Thinking Science* lessons. Metacognition with transfer to other lessons has been identified as "two of the most significant concepts in the field of teaching thinking" (Leat & Lim, 2003, p. 386).

The Impact of *Thinking Science* on Students' Cognitive Development

In the original trial and experimentation with the CASE intervention in the UK, students in CASE schools achieved statistically significantly higher results than their peers in control schools in the British General Certificate of Secondary Education (GCSE), the national examination taken when students are 16 years of age, three years after the intervention. Moreover, the statistically significant finding was found not only in the science subject area, but also in mathematics and in English Language (Adey & Shayer, 2002). The improved student achievement in subjects other than science has been attributed to CASE having an effect on general intellectual growth, or perhaps "a fundamental effect on students' general ability to learn, and that they can then turn this generally enhanced learning ability to bear on all school subjects" (Shayer, 2000, p. 9) as well as on science-related thinking skills (Adey & Shayer, 1994). Improving cognitive ability was evident across all ability ranges with independent meta-analyses and reviews supporting these findings (Higgins et al., 2005; McGuinness, 1999; Ofsted, 2000). Summaries can be found in Shayer and Adey (2002) and Shayer (1999).

Due to the reported impact on student cognition and achievement in science cognitive acceleration programs have been developed in other subject areas including mathematics (Shayer & Adhami, 2007, 2010), as well as technology (Backwell & Hamaker, 2004) and the arts (Gouge & Yates, 2002). Moreover, a series of *Let's Think!* programs based on the same theory and pillars have been developed for primary school-aged children (e.g. Author et al., 2002, 2003). The collection of cognitive acceleration programs have been reported in a meta-analysis to show a mean effect size of 0.61 (Trickey & Topping, 2004, in Higgins et al., 2005, p. 31). Cognitive acceleration programs have been successfully adapted to educational contexts in countries outside the United Kingdom including China (Author et al., 2003), Malawi (Mbano, 2003), Finland (Hautamäki, Kuusela, & Wikström, 2002), Oregon (USA) (Endler & Bond, 2008), Pakistan (Iqbal & Shayer, 2000) and Ireland (Gallagher, 2008; McCormack, 2009). In a trial in Israel a compacted intervention using a small number of the CASE lessons was effective in promoting Year 9 students' "reasoning abilities and attainment in science, particularly in regard to the control of variables" (Babai & Levit-Dori, 2009, p. 445). The hypothesis that intelligence is modifiable and can be "enhanced by appropriate curriculum intervention" (Author, 2012, p. 212) resonates with findings about neuroplasticity and learning (Author, 2011).

Purpose and Research Questions

In 2010 the authors initiated a medium scale cognitive acceleration intervention in Australia using the *Thinking Science* professional learning materials and classroom 'thinking' lessons from the UK. The intervention involved six days out-of-class professional learning with participating teachers and in-class observation and feedback. Due to the school structure in Australia, the *Thinking Science* lessons were implemented with students when they were in Years 8 and 9 (12 to 14 years of age) compared with the typical Years 7 and 8 in the UK when they are about 6 months younger. The 'thinking' lessons were incorporated alongside the standard curriculum with students participating in a 'thinking' lesson about every two weeks as a replacement of a regular science lesson over the two-year period of Year 8 and Year 9.

The purpose of the research presented in this paper was to determine the effect on participating high school students of implementing the Cognitive Acceleration through Science Education (CASE) or *Thinking Science* program in the educational context of Australia. More specifically, the research question was: What was the effect of the cognitive acceleration program on participating students' cognitive development over the two-year program? To inform the potential expansion of the intervention within Australia, we also were interested in how the program impacted students in different schools; the general range of cognitive development evident in Australian school students; and the degree to which students' cognitive development correlated with the socio-educational status of their school.

Research Design and Methods

The design of this research was a quasi-experiment with 62 teachers and 654 students from seven high schools, including nine cohorts of students participating in the *Thinking Science* intervention and 120 students forming the comparison control group. Mixed methods of data collection were used including cognitive testing of students prior to and after the *Thinking Science* intervention, and qualitative surveys and focus group interviews with teachers participating in the *Thinking Science* intervention. Data from the interviews are not presented here.

Participants

Data were collected in seven high schools whose administration and science teachers volunteered to participate in the *Thinking Science* intervention. The data collection involved 62 teachers and 654 students when they were in Year 8 and Year 9 (ages 12 -14) over the period when *Thinking Science* was implemented in their science lessons. The schools included one small rural school and one regional school, with the remaining schools located in a state capital city. Five schools were government funded and two were private schools. One of the government schools was an academic select school. Table 1 provides an overview of the participating schools.

Insert Table 1 about here

Australian schools are identified with a value of Index of Community Socio-Educational Advantage (ICSEA) developed by the Australian Curriculum, Assessment and Reporting Authority (ACARA). Variables used to determine the ICSEA are derived from the Australian Bureau of Statistics (ABS) and include location of the school (rural, regional metropolitan), parental education, occupation and income, proportion of students with languages other than English and proportion of Indigenous students. The average ICSEA value is 1000 and standard deviation is 100 points. Schools' ICSEA values are reported publicly on the Australian Government My School website (www.myschool.edu.au) and are subject to small changes in value reflecting the school population from year to year. The participating schools are representative of a range of ICSEA values as shown in Table 1.

Quantitative Measure of the Cognitive Levels of Participating Students

Piagetian Science Reasoning Tasks (SRT) were used to measure and determine the levels of thinking from early concrete to formal operations in the school population. SRTs were developed to assess the non-verbal, general reasoning capability of students. The history, development, validity and reliability of these Piagetian-based and Rasch-scaled tasks have been described by Shayer, Küchemann and Wylam (1976), Wylam and Shayer (1978), Shayer and Adhami (2007) and Shayer (2008). Results from these studies using the SRT detail the levels of thinking in the school-aged population, distribution of levels of thinking at different ages and provide a reference point for researchers and educators. The tests arose from the interviews conducted by Piaget in seeking to elicit the reasons for children thinking in a particular way and categorising their thinking patterns within a developmental or Piagetian framework. Data from the SRT have been correlated with other nonverbal reasoning tasks to establish reliability (Shayer, Küchemann & Wylam, 1976) and were used to determine the effectiveness of the *Thinking Science* intervention in England (Adey & Shayer, 1990, 1994; Shayer & Adey, 1992). The cognitive level of a sample of 10,000 students aged between 9 and 14 years was determined using the SRT (Shayer, et al., 1976). From these data, early adolescence was identified as being a period of "rapid development in concrete thinking", p. 164 with approximately 20% of children using formal operations (Shayer, et al., 1976; Styles & Andrich 2004).

Other Rasch-scaled tests have been developed which both measure the thinking levels of students and correlate well with the SRT including Bond's Logical Operations Test (BLOT) (Endler & Bond, 2001, 2006) and Raven's Matrices (Styles, 2008). Data from both Bond's and Shayer's work suggest there exists in schools, "a broad range of cognitive development evident at average ages 13, 15 and 17 years, but that range decreased little (if at all) over the five years of high school" (Endler & Bond, 2006, chapter 4, p. 3). More recently, students' scores on the SRT have been highly correlated with scores on the Essential Secondary Science Assessment (ESSA) test), used in Australian state of New South Wales (Millar, *pers. comm*) (see http://www.schools.nsw.edu.au/learning/712assessments/essa/index.php). Raven's Matrices attempt to measure the reasoning ability component of general intelligence, *g* or general intelligence, where the task is to identify a missing element of a picture. Results on the SRT and Raven's matrices are highly correlated, with the Raven's providing a "finer level of scale" (Styles, 2008, p. 96) than the SRT, and both providing information about cognitive development using a non-reasoning task.

General reasoning ability is a predictor of scientific reasoning and not reflective of instructional quality or maturation (Wiliam, 2007). By contrast, when science (defined in terms of knowledge) is tested, scores reflect instructional quality and opportunity to learn among other variables. Similar reasoning patterns may not always be reflected in similar patterns of knowledge content. A comparative study of college level physics students in the US and China showed few differences in the distribution of reasoning despite quite different approaches to school education in both countries and very large differences in levels of content knowledge (Bao et al., 2009).

Researchers working with teachers in the study reported in this paper determined that use of the BLOT or ESSA tests as measures would exclude many students from the data collection due to the literacy demands of these tasks. By contrast, the SRT use familiar laboratory apparatus to show students the activities of pouring water, weighing small items on a scale, using a ruler and balancing a beam, activities that could be readily demonstrated by teachers in the participating science classes. Because of the demonstrations the literacy demands on the students are low. To standardise the process, teachers were provided with a video and power point presentation prepared by the School of Isolated and Distance Education in Western Australia initially for use with students in remote parts of Australia. Piagetian Science Reasoning Tasks (SRT) were used to determine students' level of cognitive development before and after the intervention of

Thinking Science. The SRT (volume and heaviness) was administered to all Year 8 students prior to the implementation of *Thinking Science* program and a different SRT (equilibrium and balance) was administered on completion of the full program at the end of the second year. Teachers in their science classes using the available video, power point and classroom equipment administered the tests. The test papers were scored independently by researchers.

Only twice tested students from each participating school were included in the data set. Published data with control and experimental groups have been available for researchers to use for comparative purposes, particularly in the absence of particular populations. We drew on these data (Adey & Shayer, 1990) as in an earlier study to determine the effect of a pilot study with one school cohort (Author et al., 2012). The control data served as a comparison to gauge the effect of the intervention. The control data were drawn from a population of aged matched students who did not participate in the *Thinking Science* intervention but were twice-tested at equivalent time points at the start and end of the program. As children mature, their levels of cognition increases (Shayer, Kuchemann & Wylam, 1976), so the gains made over the course of the program are more reflective of the effectiveness of the program rather than the actual raw scores. Cognitive gains made by the participating students were compared with those who did not experience the intervention using a t-test of significance. To determine the effect of the intervention, effect sizes were calculated as suggested by Allen and Bennett (2008) and Cohen's *d* was used to indicate the magnitude of the differences in cognitive gain between the intervention and control groups. Using Cohen's (1988) conventions as a guide, *d* of .20 can be considered small, *d* of .50 is medium, and *d* of .80 is large.

Findings

The findings are structured into two main sections. The first section presents findings with regard to the relationship between cognitive levels of Australian students and the socio-educational status of their schools as well as the range of cognition evident within a particular school at the start of the intervention. The second section presents findings related to the effect of the intervention on students' cognitive development..

Cognitive Levels of Students in Australian Schools

Figure 1 present the data from the participating schools with the mean baseline score for the Year 8 students and the schools' Index of Community Socio-Educational Advantage ICSEA. These data are taken from the large data set on Year 8 students tests in at the start of the intervention. The correlation between the students' levels of thinking and the school ICSEA value is positive (r =0.71).

Insert Figure 1 about here

Students' levels of thinking were determined using a Piagetian SRT. Figure 2 shows the range of levels of cognitive development within one cohort of Year 8 at one school (School 5).

Insert Figure 2 about here

The Effect of the Intervention on Australian Students' Cognitive Development
Table 2 presents the data of cognitive gains from students in each of the nine cohorts in the seven schools and the control sample as reported by Adey and Shayer (1990). A total of 654 students were twice tested from the initial schools' sample of more than 1200 Year 8 students. These students started at a lower mean cognitive level compared with the control population, but made greater cognitive gains over the intervention period.

Insert Table 2 about here

The mean gains made in each cohort and overall are significant at the .05 level when compared with the control group with one exception (School 1, Cohort 1b). The overall mean effect size of 0.56 compares with the gain made by the control group and falls within what Hattie (2009) described as being 'worthwhile' and comparable to the gains reported in a pilot case study reported earlier (Author et al., 2012). The *Thinking Science* intervention had a differential impact on students from different school cohorts with effect sizes ranging from 0.2 to 0.995 (Table 2). The smallest effect was found with Cohort 1b School 1, a small rural school and the largest in School 7, an academic select school.

Discussion

Levels of thinking are closely correlated with the schools' ICSEA and this reflects a degree of social inequity. It is not the place here to explore that issue or conundrum but to present the data as one variable that may enable intervention programs to be successfully implemented, sustained and developed in schools. The more 'disadvantaged' in this data set appear to make gains compared with the control group but not as much as the students in more 'advantaged' schools, although positive 'teacher' effect' has been identified in the data in a low SES school, where greater fidelity to the program was observed. The schools with higher ICSEA values, at least in this sample, had greater stability in terms of student population, staffing, and participation in the professional learning opportunities. There was high attrition from the data set, with the school with the smallest gain having the greatest attrition of both students and teachers involved in the professional learning program and, conversely, the school with the greatest gain had the lowest attrition of both students and teachers. Schools experience different and changing priorities, with varying rates of student attendance and teaching staff turnover. Such factors inevitably impact on the effectiveness of an intervention program, and raise questions about scalability and sustainability (see Lee & Krajcik, 2012 for a discussion and overview), not to mention some of the less tractable problems of social equity, resource allocation and access to what we might call 'high quality teaching'. The findings presented here are nevertheless of interest as in optimal conditions, with a stable student population, high rates of school attendance and a science department that embedded the intervention practices into the teaching and learning program, the effects were clear: students show a large gain in their levels of reasoning compared with students in other schools (Author, *under review*).There may be other influences such as the school environment, families, peers and other resources not considered in this study that support interventions such as this one in schools and impact on individual students. Other studies on cognitive acceleration interventions have shown the effects of individual teachers on students' thinking (Author, 2002), which points to the non-homogeneous impact in the schools participating in this study. Indeed, teachers exert considerable effect on students' learning, and gains in achievement (Taylor, Roehrig, Hensler, Connor, & Schatschneider, 2010). Understanding the impact of high quality teaching is a likely driver of policy development and the monitoring of teaching standards.

Findings from the early work reported from the CASE project (Adey & Shayer, 1990) showed that males made greater gains than females. It was suggested then as a possible explanation for the differential impact on students, that brain maturation occurs at different rates and this has subsequently been confirmed by Andrich and Styles (1994) and Lenroot and Giedd (2010). Work is currently underway to establish whether starting the program a year earlier (Year 7 in Australian schools) in a girls' school will result in greater gains for females.

Key to the success of *Thinking Science* are the cognitive conflicts set within a specific reasoning pattern, the pedagogy that drives the discussion of ideas in student groups and metacognition. These instructional strategies when used together have the capacity to improve the reasoning ability of students. The results of the pilot study reported earlier (Author, 2012), demonstrated that participating students' achievement in science between Years 7 and 9 showed greater gains than other students in the state of Western Australia as measured by the statewide monitoring standards in education tests (WAMSE, see http://www.scsa.wa.edu.au/internet/Years_K10/WAMSE). We anticipate continued impact of CASE lessons on scholastic achievements, and data will show whether there are effects with long term and transfer across the curriculum. The overall effect size of 0.56 certainly warrants closer examination of the CASE practices and impact on different students.

Improving the thinking of teenagers has consequences for their performance in school and beyond in terms of equity, economics and life course (OECD, 2010). The teenage years are of particular interest to educators as they include the second period of considerable intellectual growth spurts (Andrich & Styles, 1994 being more recently confirmed using imaging by Dosenbach et al., 2010; Ramsden et al., 2011; Styles, 2008). It is from adolescence that development of formal operations is manifest in reasoning. The goal of CASE, through its rich pedagogy, is to develop formal operational thinking in all students regardless of their maturation or schooling. We should be optimistic about the effects of teaching on this age group on students who show varying degrees of aptitude for, and attitude towards their learning: that students are not set on a specific intellectual trajectory. Interventions like the *Thinking Science* program investigated in this study can make a difference to their cognitive capacity and subsequently their scholastic achievement (Author, 2002).

The control group used in this study some attention. Matched in detail for age and duration of testing program, they were disparate both in time and space. Such differences in the 'starting points' between the experimental and control groups have been addressed through a long-term study of the cognitive levels of children in the UK. These data show that compared with an age matched cohort tested 30 years apart, fewer of today's early adolescents use formal operations than their counterparts in 1976 (Shayer, Ginsburg, & Roe, 2007; Shayer & Ginsburg, 2009). In contrast to the received wisdom of the Flynn effect, Shayer documented that current day students leaving primary schools are less capable of reasoning than the previous generation. The case for a CASE intervention appears to be compelling.

Implications

There is a real tension between implementing an educational intervention with fidelity (Andrews, 2012) and allowing teachers to have the "freedom, space, and resources to create next [best] practice" (Hargreaves & Fullan, 2012, p. 51). This tension resonates with the need for teachers to "adapt materials in ways that align to standards and support learning goals" (Penuel & Fishman, 2012, p. 295), and reflects the reality of the "complex interaction between the innovation content, the local working conditions and sense making by the school team" (März & Kelchtermans, 2013, p. 15). Indeed, further research is needed to explore the rationale for teachers to choose how to develop professionally, and at the same time, offering teachers professional learning for highly effective intervention programs. Given that "differences in teacher effectiveness account for a large proportion of differences in student outcomes" (Jensen, 2011, p. 6) and subsequently on economic opportunities, programs that do make a "difference in educational improvement to the most disadvantaged students" (AERA, February 2014, p. 2) need to be supported by policy makers and administrators. Universities have a role to play in disseminating evidence of best practice, supporting teacher development and informing policy direction (Connor, Alberto, Compton, & Connor, 2014; Lee & Krajkic, 2012). The suite of cognitive acceleration programs not only have a body of literature to merit their consideration and adoption, an acceptance that "teaching for thinking is a very special case of thinking" (Adey, 2006, p. 56), and the associated professional development have also been well articulated. The 'why' of teaching or changing practice needs to be at the heart of the debate followed by the 'how' and 'what' does it take to get us there? We suggest that there is a moral imperative to bring CASE back from the cold and situate the theory, practice and impact into the current debate about pedagogy.

Conclusion

The overall impact of the *Thinking Science* intervention on the cognition of 654 students in seven high schools in Australia was positive and was represented by an effect size of 0.56 when compared with a control group. The findings indicate that the *Thinking Science* intervention had different impact in different schools with effect sizes ranging from 0.2 to 0.995. Overall, the findings support the wider implementation of cognitive acceleration pedagogy in Australian schools to support the general capability of critical and creative thinking of the Australian Curriculum. There is, however, tension between the need to implement an intervention such as *Thinking Science* with fidelity and the professional freedom of teachers.

Acknowledgements
This research was supported by a grant from the Australian Research Council (DP1093877). The ideas presented in the paper are those of the authors and not the funding institution. We acknowledge the late Professor Philip Adey whose wisdom and encouragement over the years of this research was invaluable.

References

ACARA Australian Curriculum, Assessment and Reporting Authority. (2012). *The Australian Curriculum v4.0*. Accessed 10th April 2014 from http://www.australiancurriculum.edu.au/

Adey, P. (2012). From Fixed IQ to Multiple Intelligences. *Bad Education: Debunking Myths in Education.* P. Adey & J. Dillon. Maidenhead, Open University Press**:** 199-214.

Adey, P., & Shayer, M. (1990). Accelerating the development of formal thinking in middle and high school students. *Journal of Research in Science Teaching, 27*(3), 267-285.

Adey, P., & Shayer, M. (1994). *Really raising standards: cognitive intervention and academic achievement*. London, Routledge.

Adey, P., Robertson, A., & Venville, G. (2002). Effects of a cognitive acceleration programme on Year 1 pupils. *British Journal of Educational Psychology 72:* 1-25.

Allen, P., & Bennett, K. (2008). *SPSS for the health and behavioural sciences.* Melbourne, Australia: Thomson.

Andrews, D., (2012) In search of feasible fidelity. *Better Evidence-based Education, 4*(2) 22-23.

Author et al., (2012). *International Journal of Science Education*.

Author (2011). *Studies in Science Education.*

Author et al., (2002). *British Journal of Educational Psychology*.

Babai, R., & Levit-Dori, T. (2009). Several CASE Lessons Can Improve Students' Control of Variables Reasoning Scheme Ability. *Journal of Science Education and Technology, 18*(5), 439-446.

Bao, L., et al. (2009). Physics: Learning and Scientific Reasoning. *Science 323*(5914): 586-587.

Chen, J.-Q. & McCray, J. (2012). A Conceptual Framework for Teacher Professional Development: The Whole Teacher Approach. *National Head Start Association Dialog. 15*(1) 8-23.

Chin, C. (2007). Teacher questioning in science classrooms: Approaches that stimulate productive thinking. *Journal of Research in Science Teaching, 44*(6), 815-843.

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

Connor, C. M., Alberto, P. A., Compton, D. L., & O'Connor, R. E. (2014). *Improving Reading Outcomes for Students with or at Risk for Reading Disabilities: A Synthesis of the Contributions from the Institute of Education Sciences Research*

Centers. In N. 2014-3000). (Ed.). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.
Department for Education. (2012). *Themes: Thinking skills*. http://www.education.gov.uk/schools/toolsandinitiatives/tripsresearchdigests/a0013261/themes-thinking-skills.
Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved Learning in a Large-Enrollment Physics Class. *Science, 332*(6031), 862-864.
Dosenbach, N. U. F., et al. (2010). Prediction of Individual Brain Maturity Using fMRI. *Science 329*(5997), 1358-1361.
Endler, L., & Bond, T. (2001). Cognitive Developmnet in a Secondary Science Setting. *Research in Science Education, 30*(4), 403-416.
Gallagher, A. (2008). *Developing thinking with four and five year old pupils: The impact of a cognitive acceleration programme through early science skill development*. Master's Thesis. Dublin City University.
Gouge, K. & Yates, C., (2002) *Creating a Cognitive Acceleration Programme in the Arts: the Wigan LEA Arts project* in Shayer, M., & Author, P. (eds). Learning Intelligence. Cognitive Acceleration across the curriculum from 5 to 15 years. Open University Press, Philadelphia.
Graham, S., Perin, D., (2007). A Meta-Analysis of Writing Instruction for Adolescent Students. *Journal of Educational Psychology, 99* (3), 4445-476.
Hargreaves, A., & Fullan, M. (2012). *Professional Capital: Transforming Teaching in Every School*. Abingdon, Oxon: Routledge.
Hattie, J. (2003). *Teachers make a difference: What is the research evidence*. Paper presented at the Australian Council for Educational Research Conference.
Higgins, S., Hall, E., Baumfield, V., & Moseley, D. (2005). *A Meta-Analysis of the Impact of the Implementation of Thinking Skill Approaches on Pupils.* London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
Higgins, S., Baumfield, V., Hall, E. (2007) Learning skills and the development of learning capabilities. *In: Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
Higgins, S., Kokotsaki, D., & Coe, R. (2011). *Toolkit of Strategies to Improve Learning - Summary for Schools, Spending the Pupil Premium*. Durham.
Jensen, B., Hunter, A., Sonnemann, J., & Burns, T. (2012). *Catching up: Learning from the best school systems in East Asia*: Grattan Institute.
Jensen, B., & Reichl, J. (2011). *Better teacher appraisal and feedback: Improving performance*. Melbourne: Grattan Institute.
Lee, O., & Krajcik, J. (2012). Large-scale interventions in science education for diverse student groups in varied educational settings. *Journal of Research in Science Teaching, 49*(3), 271-280.
Lenroot, R. K., & Giedd, J. N. (2010). Sex differences in the adolescent brain. *Brain and Cognition, 72*(1), 46-55.
Leyser, O. (2014). What Should be in the Biology Curriculum? *School Science Review, 95*(352), 43-45.

Leat, D. and M. E. I. Lin (2003). "Developing a Pedagogy of Metacognition and Transfer: Some signposts for the generation and use of knowledge and the creation of research partnerships." British Educational Research Journal **29**(3): 383-415.
Lee, O., & Krajcik, J. (2012). Large-scale interventions in science education for diverse student groups in varied educational settings. *Journal of Research in Science Teaching, 49*(3), 271-280.
Lipman, M. (1976). Philosohy for Children. Metaphilosophy 7:1, 17-39.
Lipman, M. (2003). *Thinking in Education* (2nd ed.)New York: Cambridge University Press.
Maaske, R. J. (1949). What Is Good Teaching? *The Educational Forum, 13*(2), 226-227.
Marchand, H. (2012). Contributions of Piagetian and post-Piagetian theories to education.
März, V., & Kelchtermans, G. (2013). Sense-making and structure in teachers' reception of educational reform. A case study on statistics in the mathematics curriculum. *Teaching and Teacher Education, 29*(0), 13-24.
Matsumura, L. C., Garnier, H. E., & Spybrook, J. (2012). The Effect of Content-Focused Coaching on the Quality of Classroom Text Discussions. *Journal of Teacher Education, 63*(3), 214-228.
McCormack, L. (2009). *Cognitive acceleration across the primary-secondary level transition*. PhD Thesis. Dublin City University.
McGuinness, C. (1999). *From Thinking Skills To Thinking Classrooms*: Department for Education and Employment (UK).
Mourshed, M., Chijioke, C., & Barber, M., (2010). *How the World's Most Improving School Systems Come Out on Top.* McKinsey & Company, London.
OECD. (2007). *PISA 2006: Science Competencies for Tomorrow's World*. Paris: OECD.
OECD. (2009). *Creating Effective Teaching and Learning Environments: First Results from TALIS*. Paris: OECD.
OECD. (2010). *The High Cost of Low Educational Performance*. Paris: OECD.
Pedagogy. (2008). In *Key Concepts in Education*: Sage UK
Penuel, W. R., & Fishman, B. J. (2012). Large-scale science education intervention research we can use. *Journal of Research in Science Teaching, 49*(3), 281-304.
Pollard, A. (2010). *Professionalism and Pedagogy: A contemporary opportunity. A Commentary by TLRP and GTCE*. London: TLRP.
Ramsden, S., Richardson, F. M., Josse, G., Thomas, M. S. C., Ellis, C., Shakeshaft, C., et al. (2011). Verbal and non-verbal intelligence changes in the teenage brain. *Nature, 479*, 113-116.
Shavelson, R.J., & Towne, T. (2002) *Scientific Research in Education*. National Academy Press, Washington, DC.
Shayer, M, Adey, P. & Wylam, H. 1981). Group Tests Of Cognitive Development Ideals and a Realization. *Journal of Research in Science Teaching 18*(2): 157-168.
Shayer, M. and P. S. Adey (1992). Accelerating the development of formal thinking in middle and high school students III: Testing the permanency of effects. *Journal of Research in Science Teaching 29*(10): 1101-1115.
Shayer, M. (2000). *GCSE 1999: Added-value from Schools Adopting the Case Intervention* London: King's College.

Shayer, M. (2008). Intelligence for education: As described by Piaget and measured by psychometrics. *British Journal of Educational Psychology, 78*, 1-29.

Shayer, M., & Adhami, M. (2007). Fostering Cognitive Development Through the Context of Mathematics: Results of the CAME Project. *Educational Studies in Mathematics, 64*(3), 265-291.

Shayer, M., & Adhami, M. (2010). Realizing the cognitive potential of children 5–7 with a mathematics focus: Post-test and long-term effects of a 2-year intervention. *British Journal of Educational Psychology, 80*(3), 363-379.

Shayer, M., Küchemann, D. E., Wylam, H. (1976). The distribution of Piagetian stages of thinking in British middle and secondary school children *British Journal of Educational Psychology 46*, 164-173.

Stephenson, J. (2009). Best Practice? Advice Provided to Teachers about the Use of Brain Gym [R] in Australian Schools. *The Australian Journal of Education 53*(2): 109-124.

Styles, I. (2008). *Linking Psychometric and Cognitive-Developmental Frameworks for Thinking about Intellectual Functioning* in Uses and Abuses of Intelligence: Studies Advancing Spearman and Raven's Quest for Non-Arbitrary Metrics. J. Raven and J. Raven. New York, USA, Royal Fireworks Press**:** 69-98.

Taylor, J., Roehrig, A. D., Hensler, B. S., Connor, C. M., & Schatschneider, C. (2010). Teacher Quality Moderates the Genetic Effects on Early Reading. *Science, 328*(5977), 512-514.

Thomson, S., Bortoli, L. D., Nicholas, M., Hillman, K., & Buckley, S. (2011). *Challenges for Australian education: results from PISA 2009: the PISA 2009 assessment of students' reading, mathematical and scientific literacy*. Camberwell Victoria, 3124: Australian Council for Educational Research.

Vansieleghem, N., & Kennedy, D. (2011) What is Philosophy for Children, What is Philosophy with Children – After Matthew Lipman? *Journal of Philosophy of Education, 45* (2), 171-182.

Visser, B. A., Ashton, M. C.,Vernon, P. A. (2006), *g* and the measurement of Multiple Intelligences: A response to Gardner, *Intelligence* 34 (5): 507–510.

Vygotsky, L. S. (1978). *Mind in society*. Edited by M. Cole, V. John Steiner, S. Scribner and E. Souberman. Cambridge, MA: Harvard University Press.

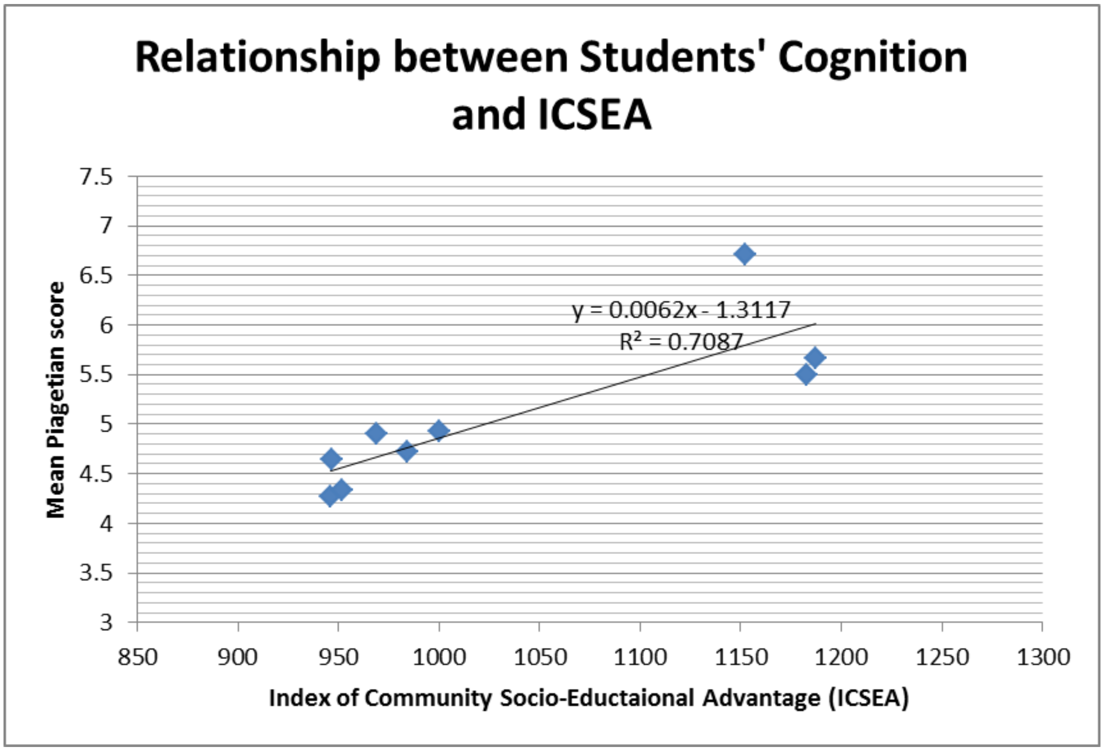
Wegerif, R., Mercer, N., & Dawes, L. (1999). From social interaction to individual reasoning: An empirical investigation of a possible socio-cultural model of cognitive development. *Learning and Instruction, 9*(6), 493-516.

Wiliam, D. (2007). Three practical, policy-focused procedures for determining an accountability test's instructional sensitivity: III: An index of sensitivity to instruction. *American Educational Research Association,* . Chicago, IL.

Wood, W. B. (2009). Innovations in Teaching Undergraduate Biology and Why We Need Them. *Annual Review of Cell and Developmental Biology, 25*(1), 93-112.

McCormack, L. (2009). *Cognitive acceleration across the primary-secondary level transition*. PhD Thesis. Dublin City University.

Figure 1
Participating schools' mean Year 8 baseline Piagetian Science Reasoning Task (SRT) scores compared with the schools' Index of Community Socio-Educational Advantage (ICSEA).



For review purposes only, these are the data from the schools

	ICSEA	Mean baseline score
1	947	4.64
2	946	4.27
3	984	4.72
4	952	4.33
5	969	4.90
6	1187	5.67
7	1000	4.93
8	1152	6.71
9	1183	5.49
10	989	4.95

Table 1
An overview of the schools that participated in the CASE intervention

School	Sector	Location	ICSEA[1]	Students (n = 654)	Teachers (n = 63)
School 1 (Cohort 1a)	Public	Rural	946	68	6
School 1 (Cohort 1b)	Public	Rural	946	27	4
School 2	Public	City	984	63	7
School 3	Public	City	952	32	5
School 4	Public	Regional	969	94	9
School 5 (Cohort ~~5b~~5a)	Catholic	City	1187	91	5
School 6	Independent	City	1000	64	8
School 7	Public	City	1158	144	12
School 5 (Cohort 5b)	Catholic	City	1183	71	6

[1] School Index of Community Socio-Educational Advantage in the year data were collected

Figure 2
Distribution of levels of thinking in one Year 8 cohort

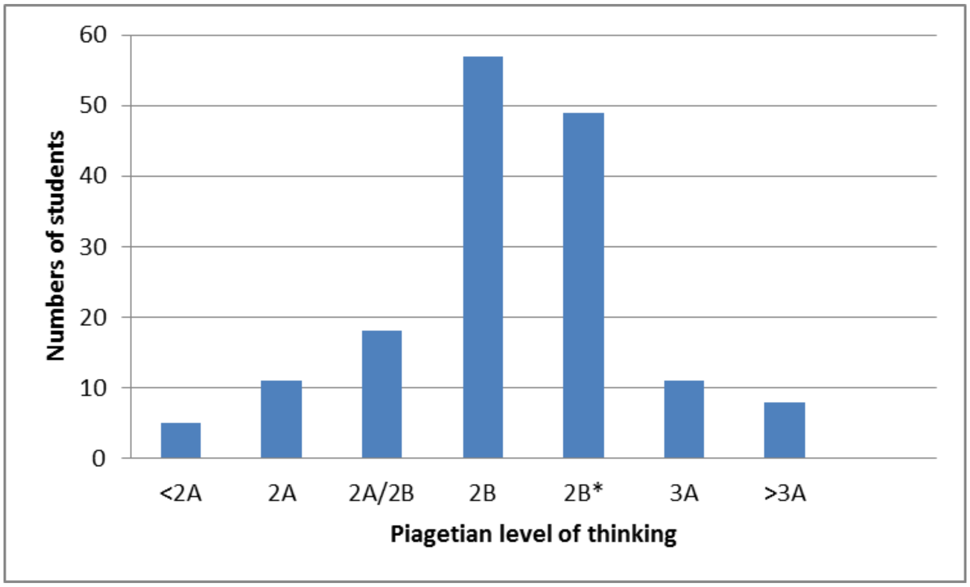


Table 2
Schools' mean pre/post cognitive gains as measured by Science Reasoning Tasks (SRT) over the period of the two-year intervention.

	n	Pre-test mean (SD)	Post-test mean (SD)	Gain	Effect size (Cohen's *d*)
Control	120	6.17 (1.03)	6.64 (1.36)	0.46 (1.09)	
School 1 (Cohort 1a)	68	4.82 (0.94)	5.75 (0.77)	0.94 (0.95)	**0.47**
School 1 (Cohort 1b)	27	4.90 (1.07)	5.60 (0.66)	0.7 (1.56)	**0.2 (ns)**
School 2	63	4.99 (0.97)	5.99 (0.90)	1.00 (1.12)	**0.49**
School 3	32	4.61 (0.91)	5.90 (0.69)	1.46 (1.33)	**0.82**
School 4	94	4.80 (1.35)	6.03 (1.05)	1.23 (1.20)	**0.67**
School 5 (Cohort 5a)	91	5.72 (1.15)	6.71 (1.07)	.995 (1.22)	**0.46**
School 6	64	5.06 (1.06)	5.89 (0.96)	.84 (1.02)	**0.36**
School 7	144	6.23 (0.94)	7.89 (1.20)	1.65 (1.30)	**0.995**
School 5 (Cohort 5b)	71	5.5 (1.3)	6.52 (0.92)	1.03 (1.05)	**0.53**
All students	654	5.18 (1.08)	6.25 (0.91)	1.09 (1.2)	**0.56**

All gains were significant to 0.05 with the exception of one school cohort.