2005

Volume 17, No. 4

MSU Digital Commons Citation

Thinking

I am thinking about fish.

NOR

Table of Contents

Picture Credits

Cover: "I am thinking about fish." The drawings in this issue were made by
children in the Kindergarten classroom of Ellen Cahill at Bradford Academy, a
public school in Montclair, New Jersey, who engage weekly in philosophical
discussion. They were drawn in response to the teacher's request to "draw
yourself thinking," and then to "draw an idea."

Thinking in Stories

by Gareth B. Matthews
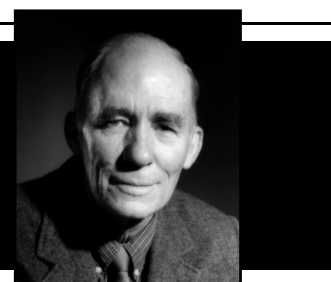
Jim Benton, *Franny K. Stein: Mad Scientist – The Fran that Time Forgot*, New York: Simon & Schuster, 2005.

Franny K. Stein is both a little girl and also a mad scientist. She created a special kind of dog food with zero gravity to make walking her dog easier and more fun. She invented cannibalistic broccoli so that kids who hate vegetables would never have to eat them.

Franny has a regular mom, a regular dad, and a regular little brother. But she herself is anything but a regular little girl. No other member of her family is even interested in science. She is mad about it.

When Franny hears of the science fair that is to be held at her school this year, she has mixed feelings. She loves creating new things to show them off to other kids. But she already has a long list of things she wants to invent. And she doesn't want to interrupt her inventing schedule. "You can't have your cake and eat it, too," her mother tells her. "Oh, can't I?" Franny thinks to herself; "I'll see about that."

Franny invents a "time-warp" dessert plate. Using this plate Franny can eat her cake and then make a small zone on the plate to go back to the time just before she had eaten the cake.

At the school science fair Franny wins first prize for her time-warp dessert plate. But the certificate she receives gives her full name, 'Franny Kissypie Stein,' rather than just her name with the middle initial, 'K,' which she prefers. All the school kids laugh when they hear that first prize goes to Franny Kissypie Stein. Franny is humiliated. Then she becomes infuriated.

Franny resolves to extend her time-warp invention so that she herself can go back to the time when she was a newborn baby to change her middle name. Confronting her baby self in the hospital nursery, she goes to the chart on the crib and erases her middle name, except for the initial 'K.' But what middle name should she give herself in place of 'Kissypie'? She settles on 'Kaboom.'

On her way back to the present time, Franny decides to keep on going past the present into the future. The teenage Franny she meets in her own future frightens her. Moreover, she finds the inventions of teenage Franny downright repulsive, and the person she is to become, evil.

When Franny returns to the science fair, she discovers that the other kids are, if anything, even more amused by 'Kaboom' than they were by 'Kissypie.' But before she flies into another rage it dawns on Franny that she didn't have to change her name in the first place. What she had to change

was how she felt about people laughing at her name. Moreover, she now resolves not to turn into the evil teenager Franny she had seen on her journey into her future.

* * *

Some philosophers think that the idea of time travel is logically incoherent. After all, for Franny to go back to the time of her birth she has to become, in a way, two people: the time traveler alongside the baby she visits. But surely one person cannot become two.

The greatest logician of the 20th Century, Kurt Gödel, dismissed this objection. Obviously thinking that the idea of time travel is coherent, Gödel went so far as to offer a formula for calculating the amount of energy that would be required to go back a little or a very long period of time. As I once heard the late Elizabeth Anscombe put the matter, Gödel's formula makes it clear that a wealthy government would be able to send someone a very long way back into the past, but even the wealthiest government would not be able to afford the cost of sending someone back to yesterday!

Jim Benton, the author of this 'Franny K. Stein' story (it is one of a series), uses the idea of time travel as a sort of thought experiment. Suppose you could go back in time and change your middle name, or do something else that, you might now think, would make your present life better. Or suppose that you could travel into your future to see what you will become, given the direction in which your life is now headed. How might altering your past or seeing into your future affect your present plans and attitudes?

In this story Franny learns from her travel into the past that she can achieve the result she wants, not by altering something done long ago, but rather by changing her present attitude towards what happened then. She also learns, by anticipating what her present self is likely to turn into, what she needs to do now to become the person she really wants to become later on.

This story is fast-paced, outrageously zany, and genuinely funny. But, despite the inventiveness and sheer fun on every page, it is also a morally serious story. Best of all, it manages to be morally serious without ever becoming moralistic. That feat is almost as mind-boggling as time travel.

— Gareth B. Matthews —

Empirical Research in Philosophy for Children: Limitations and New Directions

ALINA REZNITSKAYA

Philosophy for Children (P4C) is a non-traditional educational approach that places dialogic inquiry at the center of its pedagogy (e.g., Lipman, 1988; Splitter & Sharp, 1996). There is an expansive literature accompanying P4C practice, including curriculum materials (e.g., Lipman, 1981; Lipman, 1982), theoretical research (e.g., Lipman, 1991, 1996a; Lipman & Sharp, 1994; Lipman, Sharp, & Oscanyon, 1980; R. F. Reed & Sharp, 1996) and empirical investigations (e.g., Iorio, Weinstein, & Martin, 1984; Kyle, 1987; Niklasson, Ohlsson, & Ringborg, 1996; Terry, 1988; Yeazell, 1981).

Despite the volume of published work in connection with P4C pedagogy and the generally positive educational outcomes associated with the program (for reviews, see García-Moriyón, Rebollo, & Colom, in press; IAPC, 1982; IAPC, 1991), there remains a substantial need for theoretically-driven rigorous empirical studies of this educational approach. Many empirical investigations of P4C present largely unsystematic reflections on the goals and practices of the practice, typically supported with exemplary excerpts from discussions and quotes from students and teachers (e.g., Berrian, 1984; Fisher, 2001; Gordon, 1983; Jenkins, 1986; Kyle, 1983; Leeuw & Mostert, 1987). While interesting and thought-provoking, these studies are essentially anecdotal accounts, as they do not follow and/or report a thorough, planned, methodical process of data collection, analysis, and interpretation. In the most recent and extensive meta-analytic review of research on the effectiveness of P4C pedagogy (García-Moriyón et al., in press), the authors had to exclude the vast majority of studies for not meeting the minimum criteria related to research design, data analysis, and reporting.

In this paper, I will discuss several new possibilities for conceptualizing and conducting empirical studies in P4C. While educational goals of P4C pedagogy are diverse and multidimensional (e.g., R. F. Reed & Sharp, 1996), I will focus on a single objective of the program: the development of argumentation and reasoning. In the following discussion, I will present a psychological theory, capable of providing a better guide for empirical investigations of P4C objectives related to promoting argumentation in students. Next, I will review empirical studies of P4C, relating them to a broader context of research on argumentation development. I will point out the areas in need for further investigation and describe effective technological tools and research strategies that can enhance the quality of data collection, analysis, and interpretation.

Integrating Theoretical Perspectives: An Argument Schema Theory

The conceptualization, development, and implementation of P4C as a classroom practice has been greatly influenced by the works of several scholars within the socio-cultural tradition in psychology and education, including Vygotsky, Mead, Dewey, and Bruner (e.g., Lipman, 1988, 1996a; Lipman et al., 1980). The founder of the P4C approach, Mathew Lipman writes: "the conclusions I found in Vygotsky's *Thought and Language* (1962) and *Mind and Society* (1978) showed me how to apply, specifically to the relationship between teaching and mental development, the views I had arrived at earlier as a result of repeatedly dipping into Peirce and Dewey…" (Lipman, 1996a, p. xiii). Whether explicitly or implicitly, Lipman and others often refer to the socio-cultural frameworks when addressing theoretical and practical aspects of P4C pedagogy

*Alina Reznitskaya (*reznitskayaa@mail.montclair.edu*) is Assistant Professor of Educational Foundations at Montclair State University, and interested in researching educational environments that promote the development of argumentation and reasoning. She has authored and co-authored several publications on the influence of oral discussion on written argument and on the role of argument schema in learning to reason.*

(e.g., Lipman, 1996b; Lipman et al., 1980; Splitter & Sharp, 1996; Sprod, 1995).

The general presence of socio-cultural influences in P4C literature, however, does not address the need to have a well-articulated psychological theory, capable of providing necessary orientation for the design and analysis of empirical research. In such theory, cognitive and social mechanisms underlying student learning would be specified and linked to particular educational objectives, allowing for generation of falsifiable predictions regarding the acquisition and transfer of argumentative knowledge.

Argument Schema Theory (AST) (Reznitskaya & Anderson, 2002) represents an initial effort to develop such a framework.

AST integrates multiple, and largely independent research traditions, namely, the socio-cultural theoretical orientation (Lave & Wegner, 1991; Mead, 1962; Vygotsky, 1981) and the schema-theoretic perspective (Anderson & Pearson, 1984; Mishra & Brewer, 2003; Rumelhart, 1980). Consistent with the socio-cultural views (e.g., Bakhtin, 1981; Luria, 1981; Mead, 1962; Rogoff, 1990; Vygotsky, 1962; Wertsch, 1985; Wertsch & Bivens, 1992), AST emphasizes the priority "in time and in fact" of social interaction in cognitive development. The educational potential of a social activity comes from its dialogic organization (Bakhtin, 1981, 1986; Kuhn, 1992; Mead, 1962; Vygotsky, 1981). "It is in argumentation, in discussion, that the functional moments appear that will give rise to the development of reflection" (Vygotsky, as cited in Wertsch, 1985, p. 112). Through participation in dialogic interactions, children observe, experience, try out, and eventually internalize various "psychological tools" (Vygotsky, 1981) that advance their cognitive development to higher levels.

Socio-cultural theories present a powerful conception of learning, but their generality makes it difficult to operationalize the major tenets and to subject them to empirical tests. Several researchers criticized socio-cultural theories for lacking the desirable level of detail, explicitness, and clarity (e.g., Anderson et al., 2001; Kucan & Beck, 1997; Webb & Palincsar, 1996; Wells, 1999; Wertsch & Bivens, 1992). Anderson et al. (2001), for example, raised various

concerns about the mysterious process of internalization: "What gets internalized or appropriated and under what circumstances? How does the process work?" (p. 2).

To further develop and specify socio-cultural perspectives, AST combined them with a separate theoretical tradition, namely, schema theory. One of the major contributions of the schema-theoretic orientation comes from the idea of symbolic knowledge representation. According to this view, declarative and procedural knowledge is organized and stored in memory symbolically via generic structures, or schemas (e.g., Anderson, 1977; Bartlett, 1932; Rumelhart & Ortony, 1977; Schank & Abelson, 1977). While the definitions of a schema vary, it is typically described as an abstract knowledge structure that can be instantiated with particular, context-specific details. Learning involves generation and modification of these symbolic structures, and successful transfer, or performance in new situations, entails accessing and applying relevant schemas (Gentner, 1989; Gick & Holyoak, 1987; S. K. Reed, 1993). In a more contemporary reformulation of the schema-theoretic perspective, common elements of the schema can be connected by a unifying, explanatory theory, which accounts for and justifies existing relationships among the elements (Mishra & Brewer, 2003).

Schema theory gained its credibility through numerous empirical investigations, explaining a wide range of empirical findings (e.g., Anderson & Pichert, 1978; Bransford & Johnson, 1972; Cheng & Holyoak, 1985; Oulette, Dagostino, & J., 1999). While not without some criticism (e.g., Bigenho, 1992; Brewer, 2000; Sadoski, Pavio, & Goetz, 1991), schema theory provides a useful theoretical structure for representing and investigating cognitive development.

Applying schema theory to the context of argumentation, AST delineates the elements and functions of an *argument schema.* Based on normative models of a rational argument (e.g., Govier, 1985; Hollihan & Baaske, 1973; Jensen, 1981; Toulmin, 1958; van Eemeren, Grootendorst, & Henkenmans, 1996; Walton, 1996), an argument schema should include such elements as the statement of belief, reasons, grounds, warrants, backing, modifiers, counterarguments, and rebuttals. It
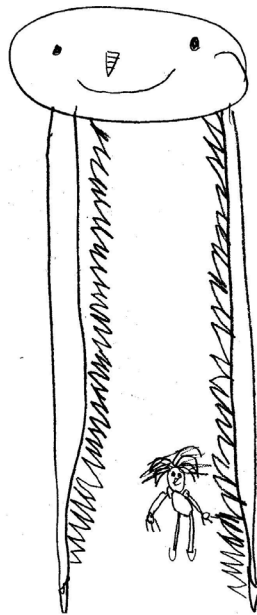
contains the understanding of the rhetorical structure of an argument, its properties, functions, and conditions for use. It includes knowledge about the inferential rules of reasoning, as well as cognitive and social practices appropriate for argumentation. Importantly, the argument schema is more than a simple collection of individual elements. Rather, the elements and their relationships are supported through a set of epistemological beliefs, which constitute an "explanatory framework" (Mishra & Brewer, 2003) for the schema. An *evaluative* type of epistemology (Kuhn, 1991) represents the normative structure. The evaluative view assumes that knowledge is relative and contextual, while also recognizing that some judgments are more reasonable than others (cf., Hofer & Pintrich, 1997; King & Kitchener, 1994; Perry, 1970).

According to Anderson et al. (2001), an argument schema can be further broken down into recurrent patterns, or *argument stratagems.* Argument stratagems are specific rhetorical and reasoning moves utilized in argumentation. "A complete argument stratagem is comprised of information about (a) the purpose or function of the stratagem, (b) the conditions in which the stratagem is used, (c) the form the argument takes, (d) the consequences of using the stratagem, and (e) the possible objections to this form of argument" (Anderson et al., 2001, p. 2).

AST assumes that it is possible to postulate general, "field-invariant" characteristics of an argument. While different domains (i.e., moral, scientific, legal) may have their own argumentation standards (Toulmin, 1958), even these "field-dependent" rules can be generalized across multiple contexts. Thus, we can think of argumentative knowledge as an aggregation of field-invariant and field-dependent rules, principles, and informal heuristics, which together comprise an argument schema. Generalizing from research on various types of schemas and discourse structures (Anderson & Pichert, 1978; Bransford & Johnson, 1972; Brewer & Treyens, 1981; Chambliss, 1995; Cheng & Holyoak, 1985; B. J. Meyer, Brandt, & Bluth, 1980; S. K. Reed, 1993), the hypothesized functions of a developed an argument schema include: (1) allocating attention to argument-relevant information; (2) directing retrieval of argument-relevant information from memory and permitting inferential reconstruction; (3) organizing argument-relevant information; (4) providing the basis for anticipating objections and for finding flaws in one's own arguments and the arguments of others; and (5) facilitating argument comprehension, construction, and repair (Reznitskaya & Anderson, 2002).

To explain the acquisition of an argument schema, we return to the socio-cultural perspective on learning. Argument schemas are developed through socialization into argumentative discourse in a dialogic collective setting. Pedagogically effective group discussions allow participants to use the discourse of reasoned argumentation in a variety of situations. While contextually different, these discussions share important structural elements, providing students with multiple instances from which to abstract the rules and practices of argumentation.

Abstract properties of a schema should enable its application to new situations, or transfer of argumentative knowledge. Just like entering a new restaurant activates a restaurant schema (Schank & Abelson, 1977) abstracted from multiple prior experiences with eating out, an encounter with a task requiring the use of argumentation should trigger a set of cognitive and social practices that constitute an argument schema. The richness of an individual's argument schema depends on the number, variety, and quality of prior encounters with argumentation. Prior experience with argumentation, such as engagement in dialogic interactions, is also hypothesized to affect the ability to access and apply the schema spontaneously, flexibly, and effectively.

To summarize, Argument Schema Theory provides an account of argumentative knowledge acquired through participation in dialogic interactions. Integrating schema theories with socio-cultural perspectives makes it possible to further specify *what* argumentative abilities are being developed and *how* social interaction affects their development.

AST is still in the early stages of its development. For now, it aims to provide a meaningful starting point for the development of a comprehensive theoretical structure. Importantly, the structure must be sufficiently detailed to offer a frame of reference that can assist empirical researchers in determining study design, identifying relevant variables, selecting data-analytic methods, and interpreting the findings. While many components of the current AST model need to be elaborated further through theoretical advances and empirical research, even in its current form, AST offers a useful guide to future empirical explorations in the area of argumentation development. In the next section, I will discuss P4C research related to argumentation and reasoning, calling attention to existing limitations and gaps in knowledge. I will also describe how AST, used in connection with several data-analytic strategies and technological tools, can help to advance empirical explorations in P4C to the new levels.

Enhancing the Empirical Foundations of P4C

Over the years P4C researchers have generated an impressive body of empirical literature describing and reflecting on the implementations of P4C in a variety of settings (for an annotated bibliography see http:// cehs.montclair.edu/academic/iapc/research.shtml). Yet, there are several areas that have been either overlooked or

SHE HAS AN IMAGINATION OUT OF HER HEAD

insufficiently addressed. Among the new directions related to the empirical investigations of P4C practice are 1) application of new research methodologies and technological tools for analyzing *the processes* of P4C discussions, 2) use of performance-based assessments to measure student *outcomes* following their participation in discussions, 3) systematic investigation of the *connection* between the processes of P4C interactions and related educational outcomes, and 4) modeling argumentation development with appropriate statistical procedures. I will now discuss each proposed direction in detail.

Examining the Processes of Discussions: Strategies and Tools

The renewed interest in socio-cultural theories of cognition resulted in impressive developments in methodology designed to examine and highlight the dialogic properties of classroom discourse (e.g., Anderson et al., 2001; Billings & Fitzgerald, 2002; Cavalli-Sforza, Gabrys, Lesgold, & Weiner, 1992; Cazden, 2001; Chinn, Anderson, & Waggoner, 2001; Keefer, Zeitz, & Resnick, 2000; Kumpulainen & Mutanen, 1999; Nystrand, Wu, Garmon, Zeiser, & Long, 2003). This research typically involves the fine-grained analysis and/or mapping of student and teacher contributions occurring during the discussions. For example, through constructing an argument network diagram of

student and teacher turns, Chinn and Anderson (1998) were able to evaluate the quality of interactions in terms of 1) the breadth of an argument developed by the participants, 2) the level of elaboration and explicitness, and 3) the amount of collaboration among participants in constructing individual arguments. Other researchers analyzed similar aspects of classroom interactions, including the assumed roles of participants (Billings & Fitzgerald, 2002), the structure of arguments during the discussion (Resnick, Salmon, Seitz, Wathen, & Holowchack, 1993), and the rate of acquisition of certain reasoning strategies (Anderson et al., 2001; Kumpulainen & Mutanen, 1999).

Studies of classroom discourse provide important information regarding the presence and development of dialogic exchanges in a classroom. They also describe specific methodological strategies and technological tools useful for capturing and representing important aspects of naturally-occurring discourse. For example, Anderson et al., (2001) sifted through 48 discussion transcripts, tracking the occurrence of various argument stratagems used by elementary school children during literature discussions. The stratagems served various rhetorical functions, including managing participation, positioning oneself in relation to a classmate's argument, acknowledging uncertainty, extending the story world, using story information as evidence, etc. The researchers were able to empirically demonstrate that the use of effective argument stratagems "snowballs." That is, once an effective stratagem is introduced by innovative group members, it tends to spread to other children and occur with increasing frequency.

In conducting their analysis, Anderson et al., (2001) utilized a software program called QSR Nvivo (1999). This program allows for a code-based analysis of qualitative-type data. One can assign a particular code to a selected text and perform various searches of coded text patterns. An identified and coded word string, such as a student turn in a discussion or a sentence in a persuasive essay, can be effortlessly placed back into its original context. The latter feature of QSR Nvivo allows for examination of not only the linguistic form of an utterance, but also its function, meaning, and condition of use. This, in turn, permits the required contextual sensitivity, which is often absent when natural discourse is fragmented into easily quantifiable segments.

QSR Nvivo enabled Anderson et al. (2001) to perform a fine-grained analysis of an enormous amount of data, consisting of 14,942 lines of discussion. Such expansive coverage of children's naturally occurring arguments made it possible to identify and document important trends in argumentation development. P4C researchers can greatly benefit from applying similar data-analytic methods and technological tools to the investigation of student-teacher interactions during P4C

sessions. Moreover, AST can provide a necessary "lens" for such analyses, drawing attention to those discourse characteristics that are theoretically interesting and important. By examining a hypothesized internal representation of argumentative knowledge, or an argument schema, one can postulate specific features of pedagogically effective social interactions. For example, an individual disposition to support his or her claims with reasons will come from participating in discussions where students are prompted to provide reasons for their positions or where they are able to appreciate the benefits of asking peers to justify their views. Thus, practices present in a dialogic discussion (e.g., presenting a position on an issue, generating and challenging each others' reasons, giving examples, and questioning assumptions) can serve as "psychological tools" (Vygotsky, 1981) that mediate the development of an individual argument schema. Using AST as a theoretical framework, a P4C discussion can be assessed in terms of exhibiting various elements of an argument schema, including, reasons, counterarguments, and rebuttals. This approach will help to make examinations of P4C discussions more systematic, methodologically sound, and grounded in an articulated theoretical orientation.

Analyzing Educational Outcomes: The Use of Performance-Based Assessments

While informative and valuable, research solely focusing on group interactions reveals nothing about the individual gains that may be occurring *after* students leave the social context. Yet, a fundamental concern of any educational initiative is transfer of learning. Have students participating in dialogic discussions internalized useful strategies that would allow them to perform better on reasoning tasks when social support is no longer available?

Several P4C researchers tried to address the above question by evaluating individual gains in argumentation and reasoning resulting from student engagement in group discussions (e.g., Burnes, 1981; Higa, 1980; Morehouse & Williams, 1998; Schleifer, Neveu, Mayer, & Poissant, 1999; Shipman, 1983). With a few exceptions (e.g., Morehouse &

Williams, 1998), these studies relied on objectively-scored tests utilizing select-type items, such as California Test of Mental Maturity, Progressive Matrices Test, or New Jersey Test of Reasoning Skills (e.g., Fields, 1995; Higa, 1980; Shipman, 1983; Simon, 1975).

These measurement instruments have several limitations. First, published measures, such as Progressive Matrices Test (Raven & Court, 1963), may not be sensitive enough to capture improvements in argumentative abilities, as they were not specifically designed for this purpose. Also, the general use of select-type tests of reasoning has been rightfully criticized for obscuring the thinking process that underlies the response (e.g., Chervin & Kyle, 1993; Norris, 1991). "Multiple-choice tests…provide only examinees' choices of answers to tasks, even though it is the reasoning that led to choices and not the choices themselves that are of greatest interest" (Norris, 1991, p. 459). Furthermore, poor psychometric properties of many objectively scored measures of reasoning (e.g., Bortner, 1965; Ellen, 1992; Hughes, 1992) make them inappropriate for a serious research endeavor.

Performance-based assessments are more suitable for measuring complex cognitive outcomes, such as the development of argumentation (Linn & Grounlund, 2000; Shepard, 2000). They allow for shifting the emphasis from test-takers' ability to supply the correct answer to their process of arriving at a conclusion. Several researchers

effectively utilized authentic open-ended tasks to measure student argumentative abilities (e.g., Crowhurst, 1987; Kuhn, Shaw, & Felton, 1997; Means & Voss, 1996; Reznitskaya, Anderson, & Kuo, in submission). For example, participants in Means and Voss study on argumentation (1996) were required to discuss an "open-ended problem with debatable solutions" in a face-to-face interview, reflecting on questions such as "If students misbehave in school, what should be done?" (Means & Voss, 1996, p. 144). In another study (Reznitskaya et al., in submission), students read a short, two-page story, and wrote a persuasive essay discussing the dilemma faced by the main character.

Importantly, the use of performance-based assessments does not preclude one from achieving high degree of precision and accuracy in measurement. In the Reznitskaya et al. study (in submission), researchers used a variety of strategies to enhance the reliability and validity of their data. Using AST in connection with QSR Nvivo software, researchers designed and applied an analytic scheme capable of discriminating the quality of student argumentative writing. Specifically, student compositions were first parsed out into the idea units. An idea unit, as defined by Mayer (1985), "expresses one action or event or state, and generally corresponds to a single verb clause" (p. 71). More detailed rules for chunking text into idea units are discussed elsewhere (Reznitskaya et al., 2001). Next, using QSR Nvivo, each idea unit was assigned a unique code, representing one of the elements of an argument schema, such as a position, a reason, a rebuttal, etc. The features of QSR Nvivo allowed researchers to engage in an on-going review of all instances assigned the same code (e.g., reason). The continuous review resulted in a more rigorous classification system and allowed to reduce subjectivity and personal judgment. The quality of the classification system was confirmed through high inter-rater reliability estimates, ranging from r=.87 to r=.92.

AST can assist not only with the evaluation of student responses on open-ended tasks, but also with the selection of the appropriate measurement instruments. For example, as successful transfer is generally influenced by the similarity between the learning context and the transfer task (Baldwin & Ford, 1988; Bassok & Holyoak, 1989; Brown, 1989; Gentner, 1989; Gick & Holyoak, 1987), AST would predict better transfer performance on argumentative tasks that are analogous to group oral discussions. In a study designed to explore the possibilities for and limitations of transfer performance in the domain of argumentation, Reznitskaya et. al. (in submission) designed post-intervention tasks to be increasingly different from the learning situation. The tasks included a face-to-face interview, a persuasive composition, and a recall of an argumentative text. Consistent with the AST model, students performed better on the tasks that were structurally and superficially similar to the learning context of group discussions.

When analyzing student performance on the outcome tasks, it is important to maintain the tension between the qualitative and the quantitative modes of scientific inquiry. A rigorous study of argumentation development requires rich, complex, and multidimensional sources of data. At the same time, it is often desirable to express these data in numbers, as opposed to words. The benefits of meaningful transformation of verbal information into numerical form include increased precision and differentiation, lack of disagreement, efficiency of analysis and communication, and the ability to use mathematical modeling for discovering useful generalizations. Using an explicit theoretical framework in combination with available data-analytic tools helps to preserve the authenticity and complexity of the data, while upholding the scientific principles of precision and generalizability.

Exploring the Connection between Processes and Outcomes

Research focusing exclusively on either the processes or the outcomes of P4C discussions, provides only a partial picture of the educational potential of dialogic interactions for the development of argumentation and reasoning. Yet, with rare exceptions (e.g., Sprod, 1998), processes and outcomes of P4C are not being examined concurrently. Thus,
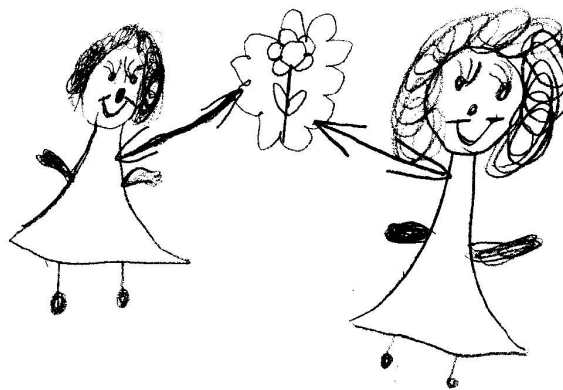
Also, outcome-type studies are typically designed to form an overall opinion regarding the success of a program, rather than to understand the underlying mechanisms of cognition. Such studies provide little information about the particular components of P4C practice and their relative contributions to the acquisition of intended skills. Thus, outcome-type research is ineffective for advancing our understanding of socio-cultural influences on learning and for providing P4C practitioners with specific, empirically-tested instructional strategies. In order to get a more accurate, thorough, and comprehensive picture of the educational potential of the P4C practice, there need to be more studies connecting the processes of the intervention to specific educational outcomes.

the relationship between the dialogic properties of interactions and the individual gains in reasoning is being assumed, rather than established.

Linking the processes of P4C intervention to specific educational outcomes may help to account for frequently documented mixed results in P4C research, when improvements are found on only some variables and for some comparison groups (Burnes, 1981; Higa, 1980; J. R. Meyer, 1988; Morehouse & Williams, 1998; Schleifer et al., 1999; Shipman, 1983). A plausible explanation for the inconsistent findings is the different quality of social interactions experienced by the students. Genuine dialogic discussions are difficult to achieve, making them a rare occurrence in a typical classroom (Almasi, O'Flahavan, & Arya, 2001; Alveraman, O'Brien, & Dillon, 1990; Billings & Fitzgerald, 2002; Nystrand, Wu, Gamoran, Zeiser, & Long, 2003). Teachers and students have various misconceptions about the nature and conduct of dialogic discussions and often confuse them with recitation-type exchanges (Nystrand, 1997; Wilen, 2004). Yet, the low dialogic quality of discussions does not become apparent in pre-post test design studies, focusing exclusively on outcome measures (e.g., Allen, 1998; Burnes, 1981; Camny & Iberer, 1988; Cummings, 1979; Iorio et al., 1984).

Statistical Modeling of Argumentation Development

There are several additional limitations in P4C research that relate to the application of statistical procedures. In their review of more than one hundred P4C investigations, García-Moriyón et al. (in press) state that "most reports do not include any statistical analysis, or what they do is poor, insufficient, or inappropriate." Common statistical problems in quantitative research of P4C include the use of gain scores, multiple t-tests, and inappropriate units of analysis (e.g., Burnes, 1981; Martin & Weinstein, 1985; Shipman, 1983).

Gain scores represent the difference between the pre and post test performance. The reliability of gain scores is inversely related to the correlation between the pre and post test scores (Crocker & Algina, 1986). Since tests used in P4C evaluation studies are likely to correlate with each other, gain scores ought to be unreliable measures of

student performance. Thus, the use of alternative statistical procedures that rely on individual test scores is more appropriate. For a single measure used as a pre and post test, such procedures include Analysis of Covariance or Repeated-Measures Analysis of Variance (Glass & Hopkins, 1996; Kirk, 1995; Pedhazur & Schmelkin, 1991).

Another common statistical procedure used by P4C researchers is to examine the differences between the treatment groups with a t-test (e.g., Higa, 1980; Morehouse & Williams, 1998; Simon, 1975). Since more than one comparison is often of interest, multiple t-tests are performed. In general, each test of statistical significance has an established probability of declaring a "false positive," called Type I error. Type I error occurs when a researcher finds differences based on the sample data, although these differences do not, in fact, exist in a population. With each additional t-test, the probability of committing Type I error goes up exponentially. An alternative statistical technique, Analysis of Variance, allows for comparison of several groups simultaneously, protecting against the increased probability of Type I error (Glass & Hopkins, 1996). When more than one student characteristic is being assessed, Multivariate Analysis of Variance should be conducted (Johnson & Wichern, 1998).

Considered last is the use of an individual student as an experimental unit of analysis. This approach ignores the fact that students in a particular classroom do not represent independent observations, as they share a unique collective context. Ignoring the contextual effects of belonging to a particular classroom or a discussion group, may bias statistical tests and lead to incorrect conclusions (Kirk, 1995; Kreft & Leeuw, 2002). The use of multilevel statistical procedures, such as Hierarchical Linear Modeling (Kreft & Leeuw, 2002; Raudenbush & Bryk, 2002), is advocated for the situations where participants are nested within groups. Unfortunately, multilevel modeling requires larger sample sizes, and may not be practically feasible for many P4C investigations.

In conclusion, this paper has presented several new directions for conducting empirical investigations of P4C pedagogy. It focused on a specific theoretical framework and described data-analytic strategies and software applications, capable of enhancing the rigor of P4C research. P4C represents a well-established pedagogical practice which embodies stimulating ideas of socio-cultural theorists. It deserves the construction of an empirical foundation, using theoretically-driven and methodologically sound approaches.

References

Allen, T. (1998). Doing philosophy with children. *Thinking, 7*(3).

Almasi, J. F., O'Flahavan, J. F., & Arya, P. (2001). A comparative analysis of student and teacher development in more or less proficient discussions of literature. *Reading Research Quarterly, 36*(2), 96-120.

Alveraman, D. E., O'Brien, D. G., & Dillon, D. R. (1990). What teachers do when they say they're having discussions of content area reading assignments: A qualitative analysis. *Reading Research Quarterly, 25*, 297-322.

Anderson, R. C. (1977). The notion of schemata and the educational enterprise. In R. C. Anderson, R. J. Spiro & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 415-431). Hillsdale, NJ: Erlbaum.

Anderson, R. C., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S., Reznitskaya, A., et al. (2001). The snowball phenomenon: Spread of ways of talking and ways of thinking across groups of children. *Cognition and instruction, 19*(1), 1-46.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson, R. Barr, M. L. Kamil & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 255-291). New York: Longman.

Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior, 17*, 1-12.

Bakhtin, M. M. (1981). *The dialogic imagination: Four essays by M. M. Bakhtin*. Austin, TX: University of Texas Press.

Bakhtin, M. M. (1986). *Speech genres and other late essays* (V. W. McGee, Trans.). Austin, TX: University of Texas Press.

Baldwin, T. T., & Ford, J. K. (1988). Transfer of learning: A review of future directions for future research. *Personnel Psychology, 41*, 63-105.

Bartlett, F. C. (1932). *Remembering*. Cambridge, UK: Cambridge University Press.

Bassok, M., & Holyoak, K. J. (1989). Interdomain transfer between isomorphic topics in algebra and physics. *Journal of experimental psychology: Learning, memory, and cognition, 15*, 153-166.

Berrian, A. (1984). Pixie: Classroom discussions following exercises in the Pixie manual. *Analytic teaching, 5*(1), 18-21.

Bigenho, F. W. (1992). Conceptual developments in schema theory (ERIC Document Reproduction Service No. ED 351 392). In. Nashville, TN: Peabody College of Vanderbilt University.

Billings, L., & Fitzgerald, J. (2002). Dialogic discussion and the Paideia Seminar. *American Educational Research Journal, 39*(4), 907-941.

Bortner, M. (1965). Review of the Progressive Matrices Test. In O. K. Buros (Ed.), *The sixth mental measurements yearbook* (pp. 489-491). Highland Park, NJ: The Gryphon Press.

Bransford, J. C., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*, 717-726.

Brewer, W. F. (2000). Bartlett's concept of schema and its impact on theories of knowledge representation in contemporary cognitive psychology. In A. Saito (Ed.), *Bartlett, culture, and cognition* (pp. 69-89). Hove, UK: Psychology Press.

Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology, 13*, 207-230.

Brown, A. L. (1989). Analogical learning and transfer: What develops? In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 369-421). Cambridge, UK: Cambridge University Press.

Burnes, B. (1981). Harry Stottlemeier's discovery: The Minnesota experience. *Thinking, 3*(1), 8-11.

Camny, D., & Iberer, G. (1988). Philosophy for Children: A Research Project for further mental and personality development of primary and secondary school pupils. *Thinking, 7*(4).

Cavalli-Sforza, V., Gabrys, G., Lesgold, A. M., & Weiner, A. W. (1992). Engaging students in scientific activity and scientific controversy. *Workshop on communicating scientific and technical knowledge, San Jose, CA*.

Cazden, C. B. (2001). *Classroom Discourse: The Language of Teaching and Learning*. Portsmouth, NH: Heinemann.

Chambliss, M. J. (1995). Text cues and strategies successful readers use to construct the gist of lengthy written arguments. *Reading Research Quarterly, 30*(4), 778-807.

Cheng, P., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive psychology, 17*, 391-416.

Chervin, M. I., & Kyle, J. A. (1993). Collaborative inquiry research into children's philosophical reasoning. *Analytic Teaching, 13*(2), 11-32.

Chinn, C. A., & Anderson, R. C. (1998). The structure of discussions that promote reasoning. *Teachers College Record, 100*(2), 315-368.

Chinn, C. A., Anderson, R. C., & Waggoner, M. A. (2001). Patterns of discourse in two kinds of literature discussion. *Reading Research Quarterly, 36*(4), 378-411.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Belmont, CA: Wadsworth.

Crowhurst, M. (1987). The effects of reading instruction and writing instruction on reading and writing persuasion. *Paper presented at the Annual Meeting of the American Educational Association, Washington, DC, (ERIC Document Reproduction Service No. ED 281 148)*.

Cummings, N. P. (1979). Improving the logical skills of fifth graders. *Thinking, 1*(3-4), 90-92.

Ellen, A. S. (1992). Review of the New Jersey Test of Reasoning Skills. In J. J. Kramer, J. C. Conoley & L. L. Murphy (Eds.), *The eleventh mental measurements yearbook* (pp. 606-608). Lincoln, NA: The Buros Institute of Mental Measurements.

Fields, J. I. (1995). Empirical data research into claims for using philosophy techniques with young children. *Early Childhood Development and Care, 107*, 115-128.

Fisher, R. (2001). Philosophy in primary schools: fostering thinking and literacy. *Reading Literacy and Language*, 67-73.

García-Moriyón, F., Rebollo, I., & Colom, R. (in press). Evaluating Philosophy for Children: A Meta-Analysis. *Thinking*.

Gentner, D. (1989). The mechanisms of analogical reasoning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 197-241). Cambridge, UK: Cambridge University Press.

Gick, M. L., & Holyoak, K. J. (1987). The cognitive basis of knowledge transfer. In S. M. Cormier (Ed.), *Transfer of Learning* (pp. 9-47). San Diego, CA: Academic Press.

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology*. Boston, MA: Allyn & Bacon.

Gordon, R. (1983). Second thoughts about *Harry*. *Analytic Teaching, 3*(2), 29-30.

Govier, T. (1985). *A practical study of argument*. Belmont, CA: Wadsworth Publishing Company.

Higa, W. R. (1980). Philosophy for children in Hawaii: A quantitative evaluation. *Thinking, 2*(1), 21-31.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140.

Hollihan, T. A., & Baaske, K. T. (1973). *Arguments and arguing: The products of human decision making*. Prospect Heights, IL: Waveland.

Hughes, J. N. (1992). Review of the Cornell Critical Thinking Tests. In J. J. Kramer, J. C. Conoley & L. L. Murphy (Eds.), *The eleventh mental measurements yearbook*. Lincoln, NA: The Buros Institute of Mental Measurements.

IAPC. (1982). Philosophy for Children: Where we are now? *Thinking, Supplement 2*.

IAPC. (1991). *Philosophy for Children: A report on achievement*. Montclair, NJ: Montclair State University.

Iorio, J., Weinstein, M., & Martin, J. (1984). A review of district 24's Philosophy for Children program. *Thinking, 5*(2).

Jenkins, J. (1986). Philosophy and religious studies: A report from Britain. *Analytic Teaching, 7*(1), 28-29.

Jensen, J. V. (1981). *Argumentation: Reasoning in communication*. New York: D. Van Nostrand Company.

Johnson, R. A., & Wichern, D. W. (1998). *Applied multivariate statistical analysis*. Upper Saddle River, NJ: Prentice Hall.

Keefer, M. W., Zeitz, C. M., & Resnick, L. B. (2000). Judging the quality of peer-led student dialogues. *Cognition and Instruction, 18*(1), 53-81.

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco, CA: Jossey-Bass.

Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences*. Boston, MA: Brooks/Cole Publishing Company.

Kreft, I., & Leeuw, J. D. (2002). *Introducing multi-level modeling*. London, UK: Sage Publications.

Kucan, L., & Beck, I. L. (1997). Thinking aloud and reading comprehension research: Inquiry, instruction, and social interaction. *Review of Educational Research, 67*(3), 271-299.

Kuhn, D. (1991). *The skill of argument*. Cambridge, UK: Cambridge University Press.

Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review, 62*(2), 155-177.

Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction, 15*(3), 287-315.

Kumpulainen, K., & Mutanen, M. (1999). The situated dynamics of peer interaction: An introduction to an analytic framework. *Learning and Instruction, 9*, 449-473.

Kyle, J. A. (1983). Thinking in writing. *Analytic Teaching, 4*(1), 5-9.

Kyle, J. A. (1987). Not a success story: Why P4C did not 'take' with gifted students in a summer school setting. *Analytic Teaching, 7*(2), 11-16.

Lave, J., & Wegner, E. (1991). *Situated learning: Legitimate peripheral participation. Cambridge, UK: Cambridge University Press.* Cambridge, UK: Cambridge University Press.

Leeuw, K. L., & Mostert, P. (1987). Learning to operate with philosophical concepts. *Analytic Teaching, 8*(1), 93-99.

Linn, R. L., & Grounlund, N. E. (2000). *Measurement and assessment in teaching*. Upper Saddle River, NJ: Merrill.

Lipman, M. (1981). *Pixie*. Montclair, NJ: IAPC.

Lipman, M. (1982). *Looking for meaning: Instructional manual to accompany Pixie*. Montclair, NJ: IAPC.

Lipman, M. (1988). *Philosophy goes to school*. Philadelphia: Temple University Press.

Lipman, M. (1991). *Thinking in education*. Cambridge: Cambridge University Press.

Lipman, M. (1996a). *Natasha: Vygotskian dialogues*. New York: Teachers College Press.

Lipman, M. (1996b). Pixie and the relationship between cognitive modeling and cognitive practice. In R. F. Reed & A. M. Sharp (Eds.), *Studies in Philosophy for Children: Pixie*. Madrid: Ediciones De La Torre.

Lipman, M., & Sharp, A. M. (Eds.). (1994). *Growing up with philosophy*. Dubuque, IA: Kendall/Hunt Publishing.

Lipman, M., Sharp, A. M., & Oscanyon, F. S. (1980). *Philosophy in the classroom*. Philadelphia, PA: Temple University Press.

Luria, A. R. (1981). *Language and cognition*. New York: Wiley.

Martin, J. F., & Weinstein, M. L. (1985). Thinking skills and philsophy for children: The Bethlehem Program, 1982-1983. *Analytic Teaching, 5*(2), 28-31.

Mayer, R. E. (1985). Structural analysis of science prose: Can we increase problem solving performance? In B. K. Britton & J. B.

Black (Eds.), *Understanding of Expository Text* (pp. 65-87). Hillsdale, NJ: Erlbaum.

Mead, G. H. (1962). *Mind, self & society from the standpoint of a social behaviorist*. Chicago, IL: The University of Chicago press.

Means, M. L., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability, and knowledge levels. *Cognition and Instruction, 14*(2), 139-178.

Meyer, B. J., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly, 1*, 72-103.

Meyer, J. R. (1988). A quest of the possible? Evaluation of the impact of the Pixie programme on 8-10 years old. *Analytic Teaching, 9*(2), 63-64.

Mishra, P., & Brewer, W. F. (2003). Theories as a form of mental representation and their role in the recall of text information. *Contemporary Educational Psychology, 28*, 277-303.

Morehouse, R., & Williams, M. (1998). Report on student use of argument skills. *Critical and Creative Thinking, 6*(1), 14-20.

Niklasson, J., Ohlsson, R., & Ringborg, M. (1996). Evaluating philosophy for children. *Thinking, 12*(4).

Norris, S. (1991). Informal reasoning assessment: Using verbal reports of thinking to improve multiple-choice validity. In F. J. Voss, D. N. Perkins & J. W. Segal (Eds.), *Informal reasoning and education*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Nystrand, M. (1997). *Opening dialogue: Understanding dynamics of language and learning in the English classroom*. New York: Teacher College Press.

Nystrand, M., Wu, L., Gamoran, A., Zeiser, S., & Long, D. A. (2003). Questions in time: Investigating the structure and dynamics of unfolding classroom discourse. *Discourse Processes, 35*(2), 135-198.

Nystrand, M., Wu, L., Garmon, A., Zeiser, S., & Long, D. (2003). Questions in time: Investigating the structure and dynamics of unfolding classroom discourse. *Discourse Processes, 35*(2), 135-200.

Oulette, G., Dagostino, L., & J., C. (1999). The effects of exposure to children's literature through read aloud and an inferencing strategy on low ability fifth graders' sense of story structure and reading comprehension. *Reading Improvement, 36*(2), 73-89.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis*. Hillsdale, NJ: Erlbaum Associates.

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart, & Winston.

QSR. (1999). QSR Nvivo [Computer software]. Victoria, Australia: Qualitative Solutions and Research.

Raudenbush, S., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (advanced quantitative techniques in the social sciences)*. Sage Publications.

Raven, J. C., & Court, J. H. (1963). *Progressive matrices*. Oxford, UK: Oxford Psychologists Press.

Reed, R. F., & Sharp, A. M. (Eds.). (1996). *Studies in Philosophy for Children: Pixie*. Madrid: Ediciones De La Torre.

Reed, S. K. (1993). A schema-based theory of transfer. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 39-67). Norwood, NJ: Ablex Publishing Corporation.

Resnick, L., Salmon, M., Seitz, C. N., Wathen, S. H., & Holowchack, M. (1993). Reasoning in conversation. *Cognition and Instruction, 11*, 347-364.

Reznitskaya, A., & Anderson, R. C. (2002). The argument schema and learning to reason. In C. C. Block & M. Pressley (Eds.), *Comprehension instruction* (pp. 319-334). New York: The Guilford Press.

Reznitskaya, A., Anderson, R. C., & Kuo, L. (in submission). Teaching and learning argumentation. *Research in the Teaching of English*.

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2&3), 155-175.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce & W. F. Brewer (Eds.), *Theoretical issues in reading and comprehension* (pp. 33-58). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 99-136). Hillsdale, NJ: Erlbaum.

Sadoski, M., Pavio, A., & Goetz, E. T. (1991). A critique of schema theory in reading and a dual coding alternative. *Reading Research Quarterly, 26*(4), 463-484.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.

Schleifer, M., Neveu, F., Mayer, M., & Poissant, H. (1999). Arguing with government. *Thinking, 14*(3), 33-38.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 1-14.

Shipman, V. C. (1983). Evaluation of the philosophy for children program in Bethlehem, Pennsylvania. *Thinking, 1*(1), 37-40.

Simon, C. (1975). Philosophy for students with learning disabilities. *Thinking, 1*(1), 21-27.

Splitter, L. J., & Sharp, A. M. (1996). The practice of philosophy in the classroom. In R. F. Reed & A. M. Sharp (Eds.), *Studies in Philosophy for Children: Pixie* (pp. 285-314). Madrid: Ediciones De La Torre.

Sprod, T. (1995). Cognitive development, philosophy and children's literature. *Early Child Development and Care, 107*, 23-33.

Sprod, T. (1998). I can change your opinion on that: Social constructivist whole class discussions and their effect on scientific reasoning. *Research in Science Education, 28*(4), 463-480.

Terry, A. (1988). I think, therefore I can: Attribution and Philosophy for Children. *Thinking, 8*(1).

Toulmin, S. E. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.

van Eemeren, F. H., Grootendorst, R., & Henkenmans, F. S. (1996). *Argumentation: Analysis, evaluation, presentation*. Hillsdale, NJ: Lawrence Erlbaum.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge: MIT Press.

Vygotsky, L. S. (1981). The genesis of higher order mental functions. In J. V. Wertsch (Ed.), *The concept of activity in Soviet psychology*. Armonk, NY: Sharpe.

Walton, D. (1996). *Argument structure: A pragmatic theory*. Toronto: University of Toronto Press.

Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 841-873). New York: Simon & Schuster Macmillan.

Wells, G. (1999). *Dialogic inquiry: Toward a sociocultural practice and theory of education*. Cambridge, UK: Cambridge University Press.

Wertsch, J. V. (1985). *Vygotsky and the social formation of mind*. Cambridge, MA: Harvard University Press.

Wertsch, J. V., & Bivens, J. A. (1992). The social origins of individual mental functioning: Alternatives and perspectives. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 14*(2), 35-44.

Wilen, W. W. (2004). Refuting misconceptions about classroom discussion. *Social Studies, 95*(1), 33-40.

Yeazell, M. (1981). What happens to teachers who teach philosophy to children. *Thinking, 2*, 86-88.

Evaluating Philosophy for Children: A Meta-Analysis

FELIX GARCÍA-MORIYÓN, IRENE REBOLLO, AND ROBERTO COLOM

Introduction

Philosophy for Children (P4C) is a program aimed at developing and fostering higher order thinking skills. It was implemented for the first time at the United States in 1970. From the beginning, there has been a deep interest in the impact of the implementation of the program on students' thinking (Weinstein, 1989). During the 70's, Matthew Lipman and his staff at the Institute for the Advancement of Philosophy for Children (IAPC) established a cooperative relationships with the New Jersey Department of Education and its Educational Testing Service. The first step was to design a test to evaluate thinking skills—the New Jersey Test of Reasoning Skills (NJTRS, Ellen, 1992)—focusing on reasoning skills such as they appeared in the curriculum for children eleven years old. Then they conducted a first field experiment in two schools in the Montclair district (Lipman & Bierman, 1975a) and, two years later, a wider study involving thousands of children in different schools in New Jersey (Shipman, 1983).

Since then, several investigations of the effectiveness of the program have been conducted in many countries. Much of that research uses a quantitative methodology with the NJTRS as the main evaluation tool. During recent years, some researchers, influenced by new trends in psychological and educational evaluation, have moved to more qualitative methods (Daniel, 2002; Pálsson, 1996; Santi, 1993). On the other hand, although the program is also intended to modify

other affective skills and personality traits, there are not many studies evaluating skills other than cognitive ones.

At present, most studies support the evidence of a positive impact on children's reasoning skills (I.A.P.C., 1982; I.A.P.C., 1991). However, discussions concerning the program evaluation have always been found to be controversial when the issue has been raised at international and national conferences. The main topics under discussion there are usually: a) the possibility of an evaluation of the skills fostered by the program; b) the skills fostered by the program that should be evaluated; c) the adequacy of quantitative or qualitative methods for a valid evaluation of those skills; and, d) the possible implications of any findings resulting from the program evaluation.

Despite all the work that has been done to evaluate the impact of P4C, there are as yet no clear conclusions. The IAPC published two reviews of a selection of findings (IAPC, 1982; IAPC, 1991), and Lipman offered a summary of those findings in his first book (Lipman et al., 1975b). However, most of those papers have been criticized for being mainly intended to advertise or convince readers of the worthiness of the program and its potential efficacy, instead of critically inquiring about its actual impact (Morehouse, 1995; Sigurdadottir, 2002; Slade, 1992).

In investigating what has been accomplished so far, the first problem encountered is *the wide variety of approaches and designs used*. Some scholars simply offer a short and simple description of their subjective positive (or negative) feelings after doing philosophy with children in their classrooms (Browning, 1988; Kyle, 1987; Schleifer & Poirier, 1996). Others are committed to a rigorous analysis and description of the categories that must be applied in qualitative research (Daniel, 2002; Echeverría, 2003). Since the middle of the 80's, many scholars have abandoned quantitative methodologies and moved to qualitative designs, but the lack of a specific and shared methodology makes it difficult to get a clear understanding of the findings of their

*Felix García-Moriyón teaches at Facultad de Formación del Profesorado, Universidad Autónoma de Madrid; Roberto Colom at the Facultad de Psicología, Universidad Autónoma de Madrid; and Irene Rebollo (*i.rebollo@psy.vu.nl*) at the Facutad de Psicología, Universidad Autonomo de Madrid and in the Department of Biological Psychology, Vrije Universiteit, Amsterdam (The Netherlands)*

In summary, there is a great deal of extant research on the implementation of the program. The results tend to offer some support for the positive effect of P4C, although there are also some evaluations that yield more skeptical results (GarcÌa-Moriyón, Colom, Lora, Rivas, & Traver, 2002; Meyer, 1988). It is difficult to reach a clear conclusion about the full implications of the positive impact or about its long lasting effect. In fact, we do not actually know if the program is working or what is the scope of its impact. The disagreement about the appropriate method of evaluation, or at least the variety of approaches, is a serious obstacle to comparing the accumulated evidence of the past 30 years. A similar problem arises from the difficulties involved in reaching an accepted definition of reasoning skills.

Interest in the rigor of empirical research has always been active among scholars within the area of P4C. Some papers have offered important contributions to clarifying theoretical and methodological problems of evaluation (Chervin & Kyle, 1993; Henderson, 1988; Santi, 1993). Two papers go further and offer a review of the research (Morehouse, 1995; Sigurdadottir, 2002), and both offer interesting suggestions as to how to move to more rigorous and well designed evaluations.

A meta-analysis such the one we are presenting in this paper can offer a better understanding of the effect of the program on children's cognitive development. The meta-analytic techniques offered are intended to revise the evidence in a field where it is difficult to reach a straightforward conclusion through a narrative review of the literature. By using both quantitative and objective criteria, we will be able to answer several important questions:

- Is there a relationship between the program's application and the factors that it is intended to influence (reasoning skills)?
- What is the extent of that observed relationship?
- Are the results obtained across studies homogeneous?
- If they are not, which characteristics of the studies could explain the variability of the results?

The answer to these questions presumably will help to define an agenda for future research in Philosophy for Children.

evaluations. Even if one focuses on the evaluations that use a classical quantitative methodology, one finds that several different designs are used, and it is difficult to reach a clear conclusion.

Secondly, most reports lack a complete and clear description of the methodology employed and the results obtained. Some give an incomplete amount of information, some do not present any data at all, and the majority do not follow the basic rules established by the scientific community for the presentation of research reports. This state of affairs could be explained by the philosophical background of the authors, and their lack of experience with the methodological requirements of educational and psychological research.

The third problem encountered is sample size. One can find well-designed studies with nicely presented results but, due to heterogeneous circumstances in the school or groups, the researcher worked with a very small group of children (Charlann, 1979; Lipman et al., 1975a; Slade, 1988; Strohecker, 1986). It is risky to generalize results reached with 7, 25, or even 35 children. And finally, most reports do not include any statistical analysis, or what they do offer is either insufficient or inappropriate.

Method

Selection of Studies and Inclusion Criteria
Four approaches were used to locate the sample of studies:
-   Computer searches of both PSYCINFO and ERIC databases were conducted using as the key term "Philosophy for Children," without any field specification. This search, conducted in May 2002, yielded 116 publications.
-   The main journals of the program—*Thinking, Analytic Teaching, Critical and Creative Thinking, and Aprender a Pensar*—were reviewed in a search for papers which fit the inclusion criteria.
-   Unpublished reports were requested through two mailing lists: P4C-list@belnet.be (English speaking) and filoninos@listserv.rediris.es (Spanish speaking).
-   A general request for unpublished or published reports was made at the NAACI Conference, in June 2002. Two papers were received as a result of this request.

Most of the collected papers were excluded from consideration because they did not fit the inclusion criteria, which were:

(1) To test the effectiveness of P4C in improving reasoning skills. Papers which included the keywords but were not related to the program, or those with only theoretical aims, were excluded.

(2) To measure reasoning skills or mental abilities as the dependent variable. Papers using personality variables or not using any measure at all were excluded.

(3) To include enough statistical information to calculate the magnitude of effects associated with the efficacy of the program in improving reasoning skills. The report had to include the sample sizes, means, and standard deviations, or the value of the *t*, or F tests, or the exact *p* value obtained from the previous tests. Studies which did not include any of those data, or only included graphs, frequencies, or verbal reports of the significance of the tests, were excluded.

Most of the collected papers (54) focused on the evaluation of reasoning skills, only two of them tested affective and cooperative skills in isolation, and some of them (15) evaluated both cognitive and affective characteristics. Finally, only twenty papers were found with theoretical aims without any data.

Coding Procedure
Sixteen publications fit the inclusion criteria. More than one study per publication was selected only if the samples among studies were independent. Two of the papers considered fit this criterion (Martin & Weistein, 1985; Slade, 1989). Otherwise, only one study by publication was analyzed. In all, eighteen studies were subjected to meta-analysis.

Each one of the 18 selected documents was independently coded by two expert P4C teachers. Initially, three kinds of moderator variables were coded (Lipsey, 1994):

(1) Substantive:

•   Measure: the instrument used to measure reasoning skills or mental abilities. Firstly, the specific name of the instrument was coded. Afterwards, this variable was recoded into two categories: (1) New Jersey Test of Reasoning Skills or Q4, and (2) others. The NJTRS was developed to measure the factors that the program is intended to change (the Q4 is an earlier version of the same test). If the program has real effects on reasoning abilities, those must be detected by measures not directly related to P4C, but also measuring reasoning skills. All the studies provided this information, and thus this variable was analyzed as a possible moderator, on the hypothesis that measures external to the program would lead to smaller effect sizes.

•   Mean age of the students: when the control and treatment groups differed in age, the mean between them was computed. Some papers reported the academic grade from which the mean age could be inferred. The mean age was 11.54 (SD = 1.97; range = 8_15.65). There was not enough variability to differentiate between children and adolescents and thus this variable was not analyzed as a moderator.

•   Book: The book used during the application of the program was coded: *Harry* (Lipman, 1982), *Pixie* (Lipman, 1981b), *Lisa* (Lipman, 1981a) or *Mark* (Lipman, 1980). Most of the studies followed *Harry*, which was designed to be used with children between 10 and 12 year olds. Thus, this variable was not considered as a moderator during the analyses.

•   Teacher training: the amount of training and experience of the teacher who applied P4C could influence its effectiveness. Unfortunately, most of the papers did not include this information or, when it was included, they were always considered experts. Thus, this variable was not analyzed as a moderator.

•   Administration time: given that the P4C program is thought to be a longitudinal treatment, it is reasonable to assume that the longer the application, the larger the effect. When the exact number of months was not reported, one academic year was coded as 9 months, and one term as 3 months. The mean duration of application was 7.33 months (SD

= 3.77; range = 2_18). There was not enough variability to consider time of application as a moderator variable, considering that there were 6 studies conducted over 9 months, and 4 studies over 8 months.

(2) Methodological:

- Research Design: there are two kinds of research designs that can be used in order to test the effectiveness of a treatment or program: Independent Groups and Repeated Measures. In the Independent groups design one group receives the treatment and the other group serves as a control (untreated). The difference between the groups on the outcome measure is used as an estimate of the program effect. In the repeated measures design a single group is used, and each individual is measured before and after the treatment (pre and post test). Then the difference between the individual scores before and after the treatment is used as an estimate of program effect. The independent groups post-test design gives a biased estimation of the effect of the program, given that the difference between the groups already present before the program implementation is unknown. The single-group pretest-postest design is also a biased because, without a control group, we do not know if the changes on the outcome measure are due to maturation or time. The most reliable and least biased design is a combination of the former two, testing an experimental and a control group

before and after the program application. In this way, the effects of maturation and prior differences between the groups are controlled. The studies submitted to meta-analysis were classified into these three categories in order to analyze the moderator effect of the applied research design. The hypothesis tested is directly related to the sources of bias present in each design. The independent groups + pretest-postest design will lead to the lowest effect sizes.

(3) Extrinsic:

- Year of publication: This factor was analyzed as a moderator under the hypothesis that older studies will lead to larger estimations of program effect, while more current studies will give lower estimations. This hypothesis is related, not only to the unconscious tendency of the researchers to find positive effects when a project is starting, but to the differences in sophistication among the research designs and statistical techniques applied as a product of time and experience.
- Source of publication: This factor was considered on the assumption that publications directly related to the program could bias the acceptance of papers towards those supporting the effectiveness of the program. If that were indeed the case, external journals or editorials would tend to publish papers with lower effect sizes. We did not find enough variability to test this hypothesis: 13 out of 18 studies belonged to journals related to the program

Table 1. **Moderator variables coding**

REFERENCE	MEASURE	DESIGN	YEAR
(Allen, 1988a)	NJTRS o Q4	Independent groups pretest-postest	1988
(Bierman, 1976)	Others	Independent groups posttest	1976
(Camhy et al., 1988)	NJTRS or Q4	Single group pretest-postest	1988
(García-Moriyón et al., 2002)	Others	Independent groups pretest-postest	2002
(Cummings, 1979)	Others	Independent groups pretest-postest	1979
(García-Moriyón et al., 1988)	Others	Independent groups posttest	1988
(García-Moriyón et al., 2000)	Others	Independent groups pretest-postest	2000
(Iorio et al., 1984)	NJTRS or Q4	Single group pretest-postest	1984
(Karras, 1979)	Others	Independent groups pretest-postest	1979
(Lipman et al., 1976)	Others	Independent groups posttest	1976
(Martin et al., 1985)(1)	NJTRS or Q4	Single group pretest-postest	1985
(Martin et al., 1985) (2)	NJTRS or Q4	Single group pretest-postest	1985
(Martin et al., 1985) (3)	NJTRS or Q4	Single group pretest-postest	1985
(Pálsson, 1996)	NJTRS or Q4	Single group pretest-postest	1996
(Reed et al., 1982)	NJTRS or Q4	Independent groups posttest	1982
(Slade, 1989) (1)	NJTRS or Q4	Independent groups pretest-postest	1989
(Slade, 1989) (2)	NJTRS or Q4	Independent groups pretest-postest	1989
(Sprod, 1997)	Others	Independent groups pretest-postest	1997

(*Thinking* and *Analytic Teaching*). Among the other 5 studies, only one was from a peer reviewed publication (*Psichotema*).

Table 1 contains information on the 3 variables finally considered for analyses (measure, design, and year) on the eighteen selected studies (Allen, 1988a; Allen, 1988b; Bierman, 1976; Camhy & Iberer, 1988; Cummings, 1979; García-Moriyón, Colom, Lora, Rivas, & Traver, 2000; García-Moriyón et al., 2002; García-Moriyón, Moreno, Pascual Díez, & Traver, 1988; Iorio, Weinstein, & Martin, 1984; Karras, 1979; Lipman & Bierman, 1976; Martin et al., 1985; Pálsson, 1996; Reed & Allen, 1982; Slade, 1989; Sprod, 1997).

Meta-Analytic Analyses

A meta-analysis or quantitative review must respond to three main questions:

(1) What is the global effect size estimated from the selected publications? Is it statistically significant?

(2) Are the effect sizes obtained across studies homogeneous? If they are not, is there enough variability among them to look for possible explanations?

(3) Is there any model based on the characteristics of the studies that could explain the observed heterogeneity?

To answer the first question it is necessary to translate the data of the primary research into a common metric, that is, the effect size (*d*). Given that there are three kinds of research designs that can be used to test the effectiveness of
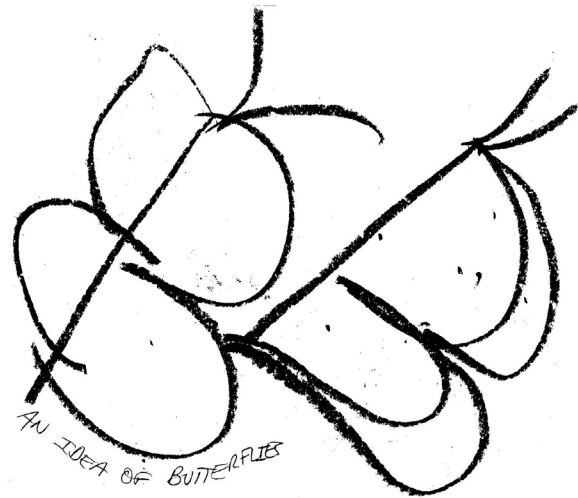
a given program, the data obtained from each one is not directly comparable. Different estimates of the effect size were applied to each design, according to Morris & DeShon (2002). To estimate the mean effect size, each study's effect size must be weighted by its sampling variance. The sampling variance is also affected by the research design, and Morris & DeShon's (2002) protocol is applied again.

Once the mean effect size is estimated, its heterogeneity must be tested by a significance test: the *Q* statistic. This statistic tests, against a chi-square

Table 2. **Sample sizes (N), Effect Sizes (*d*) and sampling variances (w$_i$) of the primary studies**

REFERENCE	N Control	N Experimental	*d*	95% Confidence Interval for the *d* Min	Max	w$_i$
(Allen, 1988a)	22	23	.6829	.6306	.7352	10.08
(Bierman, 1976)	14	14	.3944	.3421	.4467	6.32
(Camhy et al., 1988)		69	.6380	.5857	.6903	55.24
(García-Moriyón et al., 2002)	58	75	-.2284	-.2808	-.1761	32.00
(Cummings, 1979)	15	14	.7768	.7244	.8291	6.17
García-Moriyón, F. 1988	150	139	.3367	.2844	.3890	70.63
(García-Moriyón et al., 2000)	59	56	.2115	.1591	.2638	28.06
(Iorio et al., 1984)		336	.7873	.7350	.8396	254.45
(Karras, 1979)	64	64	.5394	.4871	.5917	30.35
(Lipman et al., 1976)	20	20	.8580	.8057	.9103	8.60
(Martin et al., 1985)(1)		287	.6933	.6410	.7457	229.32
(Martin et al., 1985) (2)		428	.5000	.4477	.5523	378.38
(Martin et al., 1985) (3)		249	.5097	.4573	.5620	218.31
(Pálsson, 1996)		62	1.3187	1.2664	1.3710	31.37
(Reed et al., 1982)	35	51	.7106	.6582	.7629	19.05
(Slade, 1989) (1)	15	15	.3218	.2694	.3741	6.86
(Slade, 1989) (2)	10	10	.1708	.1184	.2231	4.42
(Sprod, 1997)	29	28	.5469	.4946	.5992	13.20

distribution, the null hypothesis that there are not differences among the effect sizes of the primary studies and thus, all of them are estimations of the same parameter. If the Q statistic is statistically significant, it can be inferred that the variability across studies is not due to sampling or random effects, and different studies are estimating different parameters. Some other explanation must be found.

If the Q statistic yields a significant value, the third step is to find systematic sources of variation among effect sizes. Those possible sources are the so-called moderator variables. To estimate the effect of categorical moderators an ANOVA is computed. To estimate the effect of quantitative moderators, a regression analysis is applied.

Results

Conversion to a common metric

Table 2 shows the sample sizes and the Effect Sizes (*d*) estimated for each primary study selected for meta-analysis.

The effect size or *d* is the number of standard deviations that separates the compared scores. For the studies using a repeated measures design, *d* is the number of standard deviations that the participants have improved between the pre and post tests. For the studies using an independent groups design, *d* is the number of standard deviations that separate the scores of the control and the experimental group after the program implementation. A negative *d* means that the scores of the experimental group have worsened instead of improved.

All the effect sizes except one are positive. All of them are significantly different from zero, as their confidence intervals do not include a zero value. Apart from the negative one, the smallest value is .1708 and the largest one is 1.318. This indicates that the P4C program has a positive effect on students' reasoning skills.

Combination of the Effect sizes

The estimation of the mean effect size weighted by each study's sampling variance (also shown in Table 1) yielded a value of .5848 (*p* < .001; CI = .5325_.6372). What this indicates is that the implementation of P4C led to an improvement of students' reasoning skills of more than half a standard deviation. It must be noted that the majority of the studies implemented the program for a period of one academic year.

Heterogeneity test

The total variability estimated through the Q statistic is equal to 66.369 (*p* < .001). What this indicates is that there is some variation across studies in the sample sizes that must be explained by something different from sampling or random effects.

Moderator analyses

Table 3 shows the results obtained from the Analysis of Variance for the moderator "measure."

Table 3. **ANOVA for Measure**

Measure	*d*	95% C.I.		Q_{wi}	p (of Q_w)
		Min.	Max.		
NJTRS or Q4	.6272	.5325	.6372	13.4801	.0193
Others	.3106	.1703	.4508	47.1712	.0000
TOTALS	Q_B = 17.073 (p<.001)				
	Q_w = 49.296 (p<.001)				

Note. Q_{wi} is the variability existent into each category of the moderator; Q_b is the variability explained by the model; Q_w is te sum of the categories' variabilities and thus the remaining heterogeneity unexplained by the model.

The results indicate that there is a significant effect of the instrument used to measure reasoning abilities. Those studies using instruments directly related to the program obtained larger effect

sizes than those using measures external to the program. But it must be noted that the effect size estimated for studies using other measures (0.3106) is still significantly different from zero. The remaining heterogeneity unexplained by this moderator is still significant, which means that there must be other characteristics of the studies influencing the effect sizes. The type of instrument used explains 25% of the heterogeneity.

Table 4 shows the results obtained from the ANOVA for the moderator "design."

Table 4. **ANOVA for Design**

Design	*d*	95% C.I.		Q_{wi}	p (of Q_w)
		Min.	Max.		
Independent groups post-test	.4511	.2596	.6428	3.650	.3018
Single group pretest-postest	.6310	.5736	.6883	31.651	.0000
Independend groups pretest-postest	.2810	.1099	.4522	14.6092	.0413
TOTALS	Q_B = 16.459 (p<.001)		Q_w = 49.910 (p<.001)		

Note. Q_{wi} is the variability existent into each category of the moderator; Q_b is the variability explained by the model; Q_w is te sum of the categories' variabilities and thus the remaining heterogeneity unexplained by the model.

The results indicate that there is a significant design effect. The independent groups post-test design yields the larger estimations, while the independent groups with pre and post test design yields the smallest estimation. Given that the latter is the least biased and most reliable design, we

should take the effect size associated with it as the best estimation of the population parameter. Although it is lower than the mean estimation, .2810, it is still significant.

Another interesting result is that, only within the second category, there is still significant variability unexplained by this moderator. The type of design used explains 75.2% of the variability observed across studies.

Finally, a regression analysis was performed to study the possible effect of year of publication. The variance explained by the model is 9.2861 ($p <$ .001), and the regression coefficient was -0.019. The negative coefficient indicates that the more recent the research, the lower the effect size obtained. However, it must be noted that the year of publication is related to the type of design applied (eta = .461) so the more recent the study, the higher the probability to use a complete design (independent groups with pre and post test).

Discussion

There are several learning-to-think programs administered in many counties (Neisser et al., 1996; Nickerson, Perkins, & Smith, 1985). Virtually all of them are intended to foster high order cognitive abilities, such as deductive and inductive reasoning, language, or decision making (Feuerstein, Rand, Hoffman, & Miller, 1980; Herrnstein, Nickerson, Sanchez, & Sweets, 1986).The evidence concerning the positive impact of these programs is scarce and scattered, and very few efforts have been made to test it. However, such efforts should be seen as crucial, because applied educational psychologists must have some criteria in order to decide among possible alternative programs (Grotzer & Perkins, 2000).

A meta-analytic approach such as this one is especially useful for acquiring knowledge about the impact of these learning-to-think programs. Moreover, an analysis of published studies in terms of multiple criteria has the additional advantage of highlighting the inappropriate practices which may be present both in the research and in the reporting of results. In fact designs aimed at testing the effectiveness of a given program are usually far from ideal, basic data are unreported, or sample sizes are very small. For these reasons, we submitted several reports testing the effectiveness of P4C to meta-analysis. This meta-analysis has not been previously attempted. The results observed have several points of interest.

Before discussing the specific meta-analytic results we gathered, a few comments about the theoretical underpinnings of the program are especially germane. According to the assumptions of Lipman and his associates (Lipman, 1993), the success of the program requires that teachers emphasize the development of high order thinking skills, and that specific cognitive enrichment of children start as soon as possible. He claims that doing philosophy with children is a particularly suitable way of fostering those thinking skills, and that, as part of a curriculum adapted to the children's interests and level of personal growth, philosophy will empower children's cognitive dimension.

Scholars both in philosophy and psychology have typically received these claims with scepticism (Lipman & Sharp, 1978). Academic philosophers tend to consider philosophy much too abstract a subject matter for children, beyond both their abilities and their interests. Educational and developmental psychologists regard Lipman's approach with equal suspicion, in so far as he has challenged Piaget's widely accepted claims about the preoperational and concrete operational stages of cognitive development (Norris & Ennis, 1989; Swartz & Perkins, 1990).

The present meta-analysis offers valuable support for Philosophy for Children. In the studies reviewed, teachers who did philosophy in their classes, following the methodological
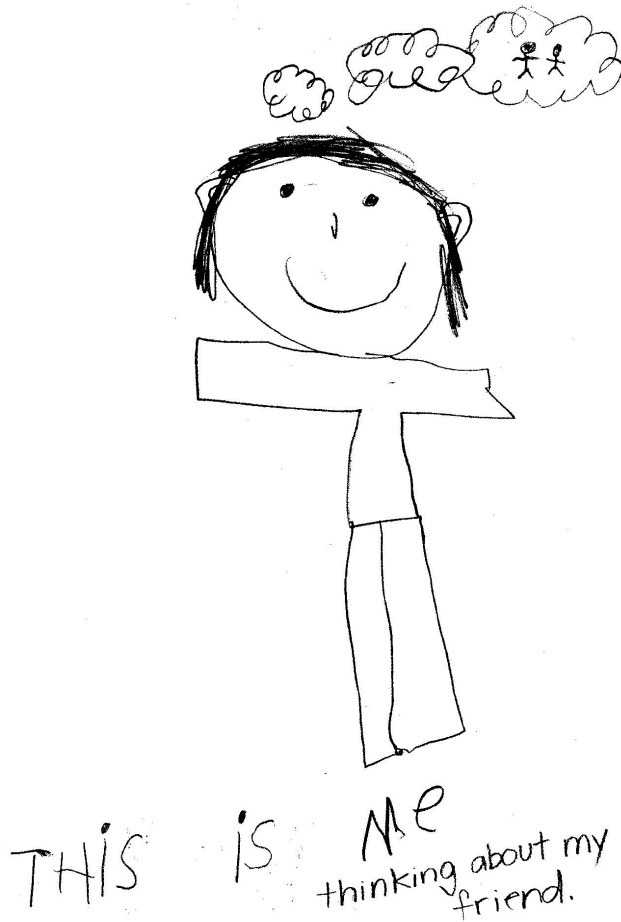
directions presented in the teachers' manuals and in the theoretical studies, and leading discussions which stressed the importance of rigorous thinking which is aware of its deductive and inductive tools, produced evident changes. This evaluation confirms the prediction stated by the theory: children do improve their cognitive skills through this methodology, and these results verify program effectiveness. Therefore, *children can do philosophy, and this practice helps them to develop higher order thinking skills*.

Returning to the most relevant meta-analytic findings, it is safe to state that P4C has a positive effect on the target thinking or reasoning skills. The average computed effect size is a noteworthy 0.58. This value implies that the gap between the treated (experimental) and untreated (control) groups is equivalent to more than half a standard deviation. For illustrative purposes, note that if the *z* score of the control group is 0 and the *z* score of the treated group is 0.5 (a difference equivalent to half a standard deviation), then the corresponding IQ scores will be 100 and 107, respectively. A *group difference* of 7 IQ points is usually considered to be a large difference (Hemphill, 2003; Nickerson et al., 1985).Thus, for instance, if we fix a cut-off IQ point of 120, then there will be 20% of people from the treated group falling beyond that point, while there will be 9% of people from the untreated group falling beyond this same cut-off point. This clearly makes a significant difference. The result is especially impressive if we note that P4C was never applied for more than one school year in all the studies reviewed. It is well known that the program was designed to be applied across a period of several years (Lipman, 1976), which allows us to predict that if the program were implemented for that duration, the effect would be both much higher and more lasting. However, it must be recognised that we do not have evidence to support that statement.

On the other hand, average effect size must be carefully interpreted, because significant variability across studies was found. This result indicates that different studies are estimating different population parameters for the effect size. The dependent measure applied in the study was one of the characteristics causing variations in the effect size. The tests employed to compare the treated and untreated groups do make a difference in the gap observed between them. The New Jersey Test of Reasoning Skills was designed as one P4C proxy measure. Unsurprisingly, greater effect sizes were observed for that test. Such practices must be avoided as far as possible. The selected dependent measure must be taken from the available measures that assess reasoning ability. Researchers can administer tests like the Culture Fair Intelligence Test, the Progressive Matrices Test, the Differential Aptitude Battery (DAT), or the Primary Mental

Abilities Battery (PMA). The composite measure that can be derived from those batteries assesses so-called general cognitive ability, which is very closely related to reasoning ability (Carrol, 1993; Jensen, 1998). Nevertheless, it is noteworthy that even using a dependent measure like the Culture Fair Intelligence Test, P4C still has a remarkable positive effect. Although it must be recognised that the effect is lower than the observed for the New Jersey Test of Reasoning Skills, it is still significant. Furthermore, the research design employed also has an effect on the observed results. We can take advantage of this result to urge researchers to use a standard design from now on. It is highly desirable to share a common design, not only for comparative purposes.

Morris and DeShon (2002) give a clear picture of the preferred design for these studies testing the effectiveness of a learning-to-think program. The most reliable and least biased design tests an experimental (treated) and a control (untreated) group before (test) and after (retest) program implementation. In this way, the effects of maturation and previous differences between the groups are controlled. We highly recommend that researchers follow this basic design. Moreover, researchers must provide the appropriate data: means, standard deviations, and number of participants, both for the pre-test and post-test measurements. These data are necessary in order to facilitate possible re-analyses like those performed in the present article.

In summary, the results of the reported meta-analysis reveal that P4C has a positive effect. However, our search for empirical reports indicates that research efforts are widespread, and we consider it clearly undesirable to compare studies performed in different countries, by different researchers, and with different samples. We recommend that researchers follow the guidelines proposed in the present paper. Moreover, P4C should be implemented according to its design—i.e. not only for one school year, but intensively across several. We expect that results from such studies would demonstrate dramatic gains in thinking skills, as opposed to the more moderate results which this study has documented.

References

Allen, T. (1988a). I think, therefore I can: Attribution and Philosophy for Children. *Thinking 8,* 14-18.

Allen, T. L. (1988b). Doing philosophy with children. *Thinking 7,* 23-28.

Bierman, M. L. (1976). A pilot study in the teaching of logic research conclusions. *Metaphilosophy, 7,* 35-39.

Browning, b. (1988). Harry in three classes. *Analytic Teaching 8,* 70-72.

Camhy, D. G. & Iberer, G. (1988). philosophy for children: a research project for further mental and personality development of primary and secondary school pupils. *Thinking 7,* 18-25.

Carrol, J. B. (1993). *Human Cognitive abilities. A survey of factor analytic studies*. Cambridge: Cambridge University Press.

Charlann, S. (1979). Philosophy for students with learning disabilities. *Thinking 1,* 21-34.

Chervin, M. I. & Kyle, J. A. (1993). Collaborative inquiry research into children's philosophical reasoning. *Analytic Teaching 13,* 11-32.

Cummings, N. P. (1979). Improving the logical skills of fifth graders. *Thinking 1,* 90-92.

Daniel, M. F. (2002). The development of dialogical critical thinking. Unpublished Manuscript.

Dewey, J. (1966). *Democracy and education*. New York: The Free Press.

Echeverría, E. (2003). El aprendizaje y la utilización del pensamiento crítico. Una investigación etnográfica. *Aprender a Pensar 5,* 60-69.

Ellen, A. S. (1992). Review of the New Jersey Test of Reasoning Skills. In J.J.Kramer & J. C. Conoley (Eds.), *Eleventh mental measurements yearbook* (pp. 606-608). Lincoln, Nebraska: Buros Institute of Mental Measurement.

Feuerstein, R., Rand, Y., Hoffman, M. B., & Miller, R. (1980). *Instrumental enrichment*. Baltimore, MD: University Park Press.

García-Moriyón, F. (2002). *Matthew Lipman: Filosofía y educación [Matthew Lipman: Philosophy and Education]*. Madrid: Ed. De la Torre.

García-Moriyón, F., Colom, R., Lora, S., Rivas, M., & Traver, V. (2000). Valoración de 'Filosofía para Niños': Un programa de enseñar a pensar. *Psicothema 12,* 207-211.

García-Moriyón, F., Colom, R., Lora, S., Rivas, M., & Traver, V. (2002). *La estimulación de la inteligencia*. Madrid: Ediciones de la Torre.

García-Moriyón, F., Moreno, A., Pascual Díez, F., & Traver, V. (1988). Evaluación de la aplicación del programa de Filosofía para Niños [Evaluation of the application of the program Philosophy for Children. Unpublished Manuscript.

Grotzer, T. A. & Perkins, D. N. (2000). Teaching Intelligence: A performance conception. In R.Stenberg (Ed.), *Handbook of intelligence* (pp. 492-515). Cambridge: Cambridge University Press.

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficient. *American Psychologist 58,* 78-80.

Henderson, A. (1988). Program evaluation issues and analytic teaching. *Analytic Teaching 8,* 43-55.

Herrnstein, R., Nickerson, R., Sanchez, M., & Sweets, J. A. (1986). Teaching thinking skills. *American Psychologist 41,* 1279-1289.

I.A.P.C. (1982). Philosophy for Children: Where we are now. *Thinking, Supplement 2*.

IAPC. (1991). *Philosophy for Children. A report on Achievement*. Montclair, NJ: Montclair State University.

Iorio, J., Weinstein, M., & Martin, J. (1984). A review of District 24's Philosophy for Children program. *Thinking 5,* 28-35.

Jensen, A. (1998). *The g factor*. London: Praeger.

Karras, R. W. (1979). Final evaluation of the pilot program in philosophical reasoning in Lexington elementary schools 1978-1979. *Thinking 1,* 26-32.

Kyle, J. (1987). Not a success story: Why P4C did not 'take' with gifted students in a summer school setting. *Analytic Teaching 7,11-16.*.

Lipman, M. (1976). Philosophy for Children. *Metaphilosophy 7*.

Lipman, M. (1980). *Mark*. Montclair: Institute for the Advancement of Philosophy for Children: University Press of America.

Lipman, M. (1981a). *Lisa*. Montclair: Institute for the Advancement of Philosophy for Children: University Press of America.

Lipman, M. (1981b). *Pixie*. Montclair: Institute for the Advancement of Philosophy for Children: University Press of America.

Lipman, M. (1982). *Harry Stottlemeier's*. Montclair: Institute for the Advancement of Philosophy for Children: University Press of America.

Lipman, M. (1993). *Thinking children and education*. Dubuque: Kendall/Hunt Publishing.

Lipman, M. & Bierman, M. L. (1975a). Field experiment in Montclair. In M.Lipman, A. Sharp, & F. Oscanyan (Eds.), *Philosophy goes to school* (Philadelphia: Temple University Press.

Lipman, M. & Bierman, M. L. (1976). Philosophy for Children; Appendix B: Experimental research in philosophy for children. *Metaphilosophy* 7, 217-224.

Lipman, M. & Sharp, A. (1978). *Growing up with philosophy*. Philadelphia: Temple University Press.

Lipsey, M. W. (1994). Identifying potentially interesting variables and analysis opportunities. In H.Cooper & L. V. Hedges (Eds.), *The Handbook of research synthesis* (New York: Russell Sage Foundation.

Martin, J. F. & Weistein, M. I. (1985). Thinking skills and philosophy for children: The Bethlehem Program, 1982-1983. *Analytic Teaching 5,* 28-31.

Meyer, J. R. (1988). A quest of the possible? Evaluation of the impact of the *Pixie* program on 8-10 years old. *Analytic Teaching 9,* 63-64.

Morehouse, R. (1995). Research in Philosophy for Children: An outline and an agenda. *Critical and Creative Thinking 3,* 74-82.

Morris, S. B. & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods 7,* 105-125.

Neisser, U., Boodoo, G., Bouchard, T., Boykin, A., Brody, N., Ceci, S. et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist 51,* 77-101.

Nickerson, R., Perkins, D. N., & Smith, E. (1985). *The teaching of thinking*. Hillsdale, NJ: LEA.

Norris, S. P. & Ennis, R. H. (1989). *Evaluating critical thinking*. Pacific Grove: Critical Thinking Press & Software.

Pálsson, H. (1996). We think more than before about others and their opinions. *Thinking 12,* 24-28.

Reed, R. & Allen, H. (1982). Analytic thinking for children in Fort Worth Elementary schools. Initial evaluation report. *Analytic Teaching* 5, 12.

Santi, M. (1993). Philosophizing and learning to think: Some proposals for a qualitative evaluation. *Thinking 10.*

Schleifer, M. & Poirier, G. (1996). The effect of philosophical discussion in the classroom on respect for others and no-stereotypic attitudes. *Thinking, 12.*

Shipman, V. C. (1983). Evaluation replication of the philosophy for children program-Final report. *Thinking, 5,* 45-57.

Sigurdadottir, B. (2002). Overarching statement: Research. Unpublished Manuscript.

Slade, C. (1988). Logic in the classroom. *Thinking 8*, 3.

Slade, C. (1989). Logic in the classroom. *Thinking 8*, 14-20.

Slade, C. (1992). Creative and critical thinking, an evaluation of Philosophy for Children. *Analytic Teaching, 13,* 25-36.

Sprod, T. (1997). Improving scientific reasoning through Philosophy for Children: An empirical study. *Thinking 13,* 11-16.

Strohecker, M. (1986). Results of the 1983-84 Philosophy for Children experiment in Lynbrook. *Thinking 6.*

Swartz, R. J. & Perkins, D. N. (1990). *Teaching thinking: Issues & approaches*. Pacific Grove: Critical Thinking Press & Software.

Weinstein, M. (1989). The philosophy of Philosophy for Children. *Analytic Teaching, 10.*

REFLECTIONS

An Interview with Matthew Lipman

SAEED NAJI

Can you tell us about the origins of the Philosophy for Children program?

In the late 1960's, I was a full professor of philosophy at Columbia University, in New York. I thought that my undergraduate students were lacking in reasoning and judgment, but that it was too late to improve their thinking significantly. I thought (and I was almost alone in this opinion at that time) that it needed to be done in childhood. There should be courses for children in Critical Thinking when the children were eleven or twelve years of age. But to make the subject "user-friendly," the text would have to be written in the form of a novel—a novel about children discovering logic. But this too seemed to me too narrow. The novel should be about children discovering philosophy. So I wrote

Harry Stottlemaier's Discovery (the title a pun on Aristotle). When I tried it out in an experiment (in 1970), it seemed to work very well, so I wrote (together with Ann Margaret Sharp) a teacher's manual containing hundreds of philosophical exercises. That worked well also. I left Columbia and set up The Institute for the Advancement of Philosophy for Children, as part of Montclair State University. After a few years, I wrote a book dealing solely with Ethics—a sequel to *Harry* called *Lisa*, for slightly older children. More and more books were written, each for a different age level, and with its own instructional manual. Also a variety of theoretical books were written and

published (by university presses like Temple, Cambridge and Teachers College Press).

In addition to requiring very unique textbooks (philosophical novels for children), Philosophy for Children has a unique pedagogy, in that students at every level begin by reading an episode aloud, raising questions about it, and then discussing the questions. It is this methodology, involving mutual criticism and scrupulously careful voicing of opinions and judgment, which educators recognize as an educational approach that prepares children to become citizens in a democracy.

How did it arise and how was it developed?

Philosophy for Children (P4C) didn't just emerge out of nowhere. It built upon the recommendations of John Dewey and the Russian educator, Lev Vygotsky, who emphasized the necessity to teach for *thinking*, not just for memorizing. It is not enough for children merely to remember what has been said to them: they must examine and analyze that material. Just as thinking is the processing of what children learn about the world through their senses, so they must think about what they learn in school. Memorizing is a relatively low-level thinking skill; children must be taught concept-formation, judgment, reasoning, etc.

After a small but intensive experiment with the program (which showed that children could be taught improved deductive reasoning without "teaching for the test"), a number of Philosophy PhD's were taught to train teachers throughout the United States, which they proceeded to do in the 1970's. By the end of the decade, some 5,000 classrooms in the country were using the program. (The experiment indicated that children of age 11 could be taught to reason 27 weeks better in mental age on formal reasoning problems, after only 9 weeks of exposure to the program.)

We proceeded, through the assistance of the New Jersey Department of Education and private foundation grants, to hold workshops for the training of teacher-trainers, who

Saeed Naji (saiid_n@yahoo.com) is a scholar at the Institute for Humanities and Cultural Science, Tehran, Iran. In Summer 2003, he addressed a series of questions to Matthew Lipman, founder of the Philosophy for Children program. This interview was originally published in the Iran News, *an English language daily, on August 24, 2003, and by Mehrnews Press, a Persian language electronic journal. It is also available on the IHCS website:* http://www.ihcs.ac.ir. *This is an edited version.*

would then turn around and train teachers, who would proceed to use the program with children. The trained teachers reported that the children responded to the program joyously, as it gave them an opportunity to talk openly in the classroom, and to discuss their ideas with one another and with the teacher.

Since Philosophy for Children is largely a language-based program, its success is related closely to its being accurately developed through a large number of languages. (Every country wants its own translation, and quite rightly so.)

How many approaches are there to Philosophy for Children? Would you please discuss them?

Just as there is only one discipline called "philosophy," so there is only one Philosophy for Children. But there are many different versions of each. For example, there are many different versions of the philosophy of science, the philosophy of psychology, the philosophy of art, and so on. Each "philosophy of" is an extensive critique of the discipline upon which it focuses. But the Philosophy for Children is not identical with Education, nor is it identical with Philosophy of Education. What is called "Philosophy for Children" represents an effort to develop philosophy so that it may function as a kind of education. It becomes education that employs philosophy to engage the mind of the child so as to try to satisfy the hunger of the child for meaning.

Philosophy *with* children has grown up as a small offshoot of Philosophy *for* Children, in the sense that philosophy with children utilizes discussion of philosophical ideas, but not through specially written children's stories. Philosophy *with* children aims to develop children as young philosophers. Philosophy *for* children aims to help children utilize philosophy so as to improve their learning of all the subjects in the curriculum.

When I first became interested in this field, I thought that children could do no better than "Critical Thinking"— that is, having their thinking trained to make it more rigorous, consistent and coherent. But critical thinking contains no concept formation, no formal logic, and no study of the works of traditional philosophy, all of which I have endeavored to supply in Philosophy for Children. Critical thinking does not lead children back into philosophy, and yet it is my contention that children will not settle for anything less. Nor should they have to. Critical thinking seeks to make the child's mind more precise; philosophy deepens it and makes it grow.

Which is the best and most influential approach to philosophy for children?

To me, the program that I have developed and that goes by the name of Philosophy for Children is identifiable as the best approach to the improvement of children's thinking. Here are some reasons:

1. *Interest*
   Children work best at whatever it is that most keenly interests them. This is P4C, first because it involves imaginative fiction, second because it is about children like themselves, and thirdly because it involves them in discussion of controversial issues (e. g. ethics). P4C goes beyond Critical thinking.

2. *Emotion*
   P4C is not limited to the improvement of critical thinking. It recognizes that thinking can be intensely exciting and emotional, and it provides ways in which children can talk about and analyze those emotions.

3. *Critical Thinking*
   P4C wholly embraces critical thinking, but it does so with greater breadth and depth. Critical thinking is generally only an "add-on" to the existing curriculum, but P4C recognizes the need children have to deal truthfully with what they find problematic or puzzling.

4. *Values*
   Children discover early on that our treatment of value issues tends to be ambiguous, vague and muddled. Consequently they welcome efforts to get them to think precisely and clearly. But this doesn't mean that their thinking should be dispassionate or lacking in feeling. Children can think better about issues that concern them, when their thinking, in addition to being critical, is caring, appreciative and compassionate.

5. *Creativity*
   Good thinking can be charged with imagination, as when we enter whole-heartedly into a story, or develop a hypothesis. P4C is therefore especially successful in the area of creativity.

6. *Communality*
   Philosophy is dialogical: it stresses the need to open the dialogue to all members of the community. In other words, it stresses shared inquiry. The world can think better about how to treat innocent victims when it feels compassion for them than when it does not.

What is your invented method in this area?

I've taken many familiar components and combined them in a new way, so as to devise a new form of education. Until Philosophy for Children came along, philosophy and education were considered quite alien to one another, not mixing the way oil and water don't mix. But a program like

P4C that urges children to think up questions, and try to answer one another's questions through open discussion is a program that combines learning and enthusiasm, feeling and thought, imagination and understanding. This is why one reviewer, speaking of my 2003 book *Thinking in Education*, says that I have "created the great maieutic epic. It is a passionate vision of what education can and should be."

I have tried to develop a new, reflective paradigm of education, whose regulative ideas are reasonableness (in personal character) and democracy (in social character). This paradigm emphasizes the importance not just of critical thinking, but of creative and caring thinking as well – all three varieties are necessary. It stresses making, saying, doing and feeling as the main arteries of judgment. Disciplined practice in these forms of understanding lies at the heart of the philosophical version of education. The pedagogy of such practice is what we call the "community of inquiry." This balanced, cooperative approach accords well with Rawls' "reflective equilibrium" and Dewey's experimentalism.

Children puzzle over many of the same concepts that philosophers puzzle over—concepts like rules, truth, goodness, justice etc. They have opinions on these matters, but they learn to develop these opinions into considered judgments. By giving them a great deal of practice in finding good reasons for their judgments, Philosophy for Children gives children an education of which they may well be proud. It teaches them how they *ought* to think.

What philosophers and psychologists greatly influenced your work in these areas?

John Dewey, for his intense sympathy for the child, his emphasis upon thinking in the classroom, and his seeing the importance of artistic creativity in getting the child to be emotionally expressive.

Justus Buchler, American philosopher in the 20th century, for his important studies in the nature of human judgment, and for his understanding of the role of judgment in the education of the child.

Lev Vygotsky, 20th century Russian psychologist, who recognized the connections between classroom discussion and children's thinking, between the child and the society by means of and through the teacher, and between the language of the adult world and the growing intelligence of the child.

Jean Piaget, 20th century psychologist and educator, whose work illuminated the relationships between thinking and behavior.

Gilbert Ryle, 20th century British philosopher, who analysed the connections between language, teaching and self-teaching.

George Herbert Mead, American philosopher and social psychologist, whose work dealt almost exclusively with the social nature of the self.

Ludwig Wittgenstein, 20th century Austrian-British philosopher, who explored with enormous sensitivity the complex social relationships that are expressed through the subtleties of language.

What are the advantages of your method?

I attempt to show those who prepare teachers for the classroom that education without philosophy in the elementary school is just as deficient as education without philosophy would be in the undergraduate and graduate areas of education. Children need to discover criteria for distinguishing between valid and invalid reasoning (logic), between supported and unsupported theories of knowledge (epistemology), and between acceptable and unacceptable forms of moral judgment (ethics), etc. The soundest way of doing this is to see education in all its vast complexity as a mode of *inquiry*, and to see

philosophy as a mode of inquiry into that mode of inquiry. Only in this way can students be encouraged to think for themselves about their own thinking. Another way of putting this is to say that inquiry is the genus of which the various forms of philosophy are species. Thus there is ethical inquiry, aesthetic inquiry, social inquiry, and so on.

Another advantage of this method is that it provides our various forms of knowledge and understanding with coherence and consistency. Children learn that most of the questions they ask can be dealt with in a reliable fashion by seeing them as occasions for philosophical inquiry. They also discover that this same method is in use now throughout the world, and that the time is not far off when this one basic method will enable them to communicate clearly with other children in similar programs in many countries and continents.

Is Philosophy for Children appropriate for all children?

Any child that is capable of using language intelligibly is capable of schooling and growth, and is therefore capable of the kind of discourse and conversation that philosophy involves. Philosophy begins when we can discuss the language we use to discuss the world. The aim is not to make children into little philosophers, but to help them think better

than they now think. Of course, the more accomplished children are with regard to listening and speaking, the more quickly they can adapt to philosophy, with its emphasis on mental acts, thinking skills, reasoning and judgment. But the program attempts to avoid any use of technical terminology.

In your book *Thinking in Education, 2nd Edition* (Cambridge, 2003) you contrast two paradigms of education, the traditional one and the reflective one. Can you explain this contrast?

Not without great oversimplification, I'm afraid. The traditional understanding people had of education, reaching all the way back to antiquity, was that it was a way of getting the younger generations to adopt the same knowledge—the same facts and relationships and values—as were considered true by the generations that had immediately preceded them. There was thought to be no need for students to reflect on or analyze what they were taught: such a critical examination had already taken place in the preceding generations.

The Enlightenment brought a call for reform. Montaigne demanded the teaching of judgment, but no one seemed to know how this was to be done, and in fact, there was little attempt to teach judgment until the end of the 20th century. Kant called for an education that would get children to think for themselves instead of one that parroted their elders; he even contributed to the design of the reflective model by referring to what he was advocating as "the practice of philosophical inquiry." But again, no one knew how this was to be done, so it was largely forgotten.

It was only with the beginning of the 20th century, particularly during the first half of that century, that educators like John Dewey began to erect the scaffolding of the new reflective paradigm. There were other educators, of course—like Durkheim and Goodman and Mead—educators who were inspired by the social sciences or by logic, enabling them to call for new principles in education, new criteria, new standards, and the best of these led to the formation of the reflective education paradigm. Except for Dewey, there was no call here for thoroughgoing pedagogical change. Dewey demanded over and over again that teachers must teach for *thinking*. It was not enough to teach merely for more up-to-date factual knowledge, just as it was not enough to teach just for reasoning or for truth. Children, like scientists, had to work together, because all inquiry was cooperative. All of it involved deliberation and participation. Thus the leaders of the reflective method of education in the 20th century saw that teaching for thinking had to be teaching for precise, open-minded, fair-minded thinking. Consequently in the latter half of the 20th century, the slogan of the more progressive educators was that the schools needed to teach for *critical thinking*—for thinking

that did not violate the principles of experimental science or of formal , or even of informal logic. But only Dewey went far beyond Ryle, Scheffler, Goodman, Nelson, Rawls and the like, to visualize education as the operative leading edge of an enormous social reform aimed at revising society into a world order in which people lived democratically as naturally as they walked upright. With the end of the 20th century came the end of the Deweyan phase of reflective education. With the start of the 21st century came the phase characterized by the introduction of such useful operatives as the community of inquiry, the reflective equilibrium, and the strengthening of judgment.

We can therefore distinguish of the earlier reflective model, shaped by the pedagogical philosophy of Dewey from the later reflective model, characterized by Philosophy for Children. The goal of thinking permeates both phases: both aim at producing thinking students, thinking teachers, thinking schools of education. Both have become sufficiently strong and enduring to see themselves planted in virtually every portion of the globe.

Why has the position you advance in the 2nd Edition of *Thinking in Education* been called the "great maieutic epic"?

The word "maieutic" is the Greek term for midwifery. It is usually taken to mean by Socrates, that there is an analogy between midwifery and teaching: the midwife delivers the pregnant mother of her child, as the philosopher delivers the ideationally pregnant student of his or her thinking. (A number of alternative interpretations have been advanced as to precisely how Socrates's pronouncement is to be understood. I think this one is as helpful as any and has the distinct merit of being applicable to the cornerstone of Philosophy for Children: helping children to think for themselves.)

And so the midwife helps the mother give birth not just to a child to a living thinker, indeed, to an entire world society of living thinkers. Socrates daringly invokes the maieutic paradigm: his doing so is a dialectical stroke of epic proportions. *Thinking in Education, 2nd Edition* is an effort to show that we are only at the beginning of the redesigning and refashioning of education. Our efforts can be considered heroically successful only when education as conceptual midwifery becomes the rule rather than the exception.

You say that to overcome the deficiency of education in elementary schools, we need to see education in all its vast complexity as a mode of inquiry, and to see philosophy as a mode of inquiry into that mode of inquiry. Furthermore, you see inquiry as "the genus of which the various forms of philosophy are species." This is a new conception of education and philosophy and

their relations. In what sense can education and philosophy be regarded as inquiry?

Inquiry is the investigative response to problematic aspects of human experience. It generally begins as questioning and moves from there into interpretation and hypothesis formation. Through discussion and deliberation, it seeks to transform the problematic into the controversial, the participatory and ultimately the reasonable.

Scientific inquiry is often invoked as alone embodying inquiry, but this is unjustified. All imaginative and creative thinking (hence all art) is inquiry, and all investigation of the ethically or valuationally problematic is inquiry. Aristotle was moving in this direction, I believe, when he asserted that all deliberation is inquiry.

But what of education and philosophy? To assert that education is ***not*** inquiry is to claim that it has to be identified with traditional rather than with reflective education, and this is unacceptable.

A person fingerpainting.

There is no justification for denying the status of inquiry to philosophy. All the humanities are forms of inquiry, and philosophy is one of the humanities. Philosophy can also be seen as a highly sophisticated form of education.

Generally speaking, why is a philosophical novel more effective in education than a merely philosophical text?

A textbook is an assemblage of a huge amount of information compiled by scholars and for scholars. But if the audience at which the textbook is aimed is made up not of

scholars, but of children, some device is needed to motivate the reading. A novel may provide a fictional, imaginative setting, an energetic dialogue, lively characters, a sprightly style, animation, humor, or all of these. In this manner, the author is able to pack the information to be communicated into the form of a ***story*** with which the reader identifies and which the reader is able to enjoy and understand. Those who write Philosophy for Children novels and workbooks can thus intersperse each page with lightly concealed philosophical meanings, problems and relationships. The children are likely to draw these out of the stories and bring them to the attention of their classmates. A sentence, a word—strikes them as ambiguous or vague or misleading. In this way, their inquiry begins, and continues until they are satisfied that they know the meaning of what was written or said.

In short, graduate students in philosophy may bring with them their own personal motivation for reading very abstract texts, but children need a motivation for doing so, and a story is often the best way of doing so. This is not a reason, however, to agree with Piaget's conception that young children cannot deal with abstractions. It would be better to say that children don't want to be made to deal with abstract, dry and technical vocabularies. They can do very well with short abstract words like ***good, bad, law, fair, hope, happy*** etc. when it happens that these are words whose usage they share with philosophers.

However, it is no simple matter to write a philosophical novel, for in addition to the criteria or considerations just cited, such a novel must challenge the readers to think independently and to discuss the embedded ideas with one another.

What characteristics, elements and components do the books concerning Philosophy for Children need to involve?

For any given novel, there is no set number of components, but a set of such novels (such as Philosophy for Children, which contains almost a dozen novels) will make demands on the author that are more specific:

(a)   Ideas must be drawn from a variety of sources in philosophy, such as epistemology, metaphysics, aesthetics, ethics, philosophy of education, logic etc. Some of these may be omitted in order to spend more time on others. For example, some areas of ethics must be represented in virtually every chapter of a given text. On the other hand, there could be a relatively small amount of ethics, but it is inadvisable to skip it completely.

(b)   There should be at least one program (a program is here a novel plus a manual) for each age level (a level usually comprises one or two ages.) Thus *Harry Stottlemeier's Discovery* is designed for children ages 9-11, and *Pixie* is for ages 7-9.

(c)   As much as possible, the language used by the speakers in the novels should correspond with the real-life language of the readers in the classrooms, their homes, and with one another.

(d)   Thinking exercises for each chapter of the novels should be provided in each separate manual. (For suggestions concerning the writing of philosophical exercises, see my article, "Philosophical discussion plans and exercises", *Critical and Creative Thinking*, 5 No.1, March, 1997.)

In addition to the humanistic contents which elementary school philosophy provides to children who study it, there are the cognitive structures and processes which it illuminates and thereby strengthens. (See Table 1.)

What are the differences between this kind of novel and other novels at the children's level?

Children's literature is a vast, complex and relatively uncharted field of writing and publishing. Much of it is directed to the home of the child, or to the school library, or to the individual to use for occasional purposes. On the other hand, P4C is specifically aimed at the classroom, where the teacher has been especially prepared for the teaching of philosophy with children.

Table 1.

Modes of Judgment	Types of Thinking	Systems Thinking Skills	Mental Act	Affective States
Making	Critical	Reasoning	Deciding	Hoping
Saying	Creative	Inquiry	Considering	Liking
Doing	Caring	Concept-Formation	Wondering	Honoring
Feeling		Translation etc.	Remembering	Encouraging
			Explaining	Prizing
			Understanding	Esteeming
			etc.	etc.

Another significant difference is that P4C aims at teaching children how to *do* philosophy—i.e. how to engage in philosophical practice. This is very different from fables or from proverbs, which aim to impart a small gem of wisdom, usually on the final page of the story.

The books in P4C are said to be novels about children "discovering philosophy." What do you mean by "children discovering philosophy?"

The stories are written in such a way as to scatter a number of quite diverse philosophical ideas at random on each page. Children, with their natural curiosity, cannot help being intrigued by these, and want the other members of the class to examine and discuss them. Such discussions should be encouraged by the teacher, and encourage the children to have faith in their own philosophical powers.

What are the differences among books concerning children of different ages, during which children may have special demands?

To some extent, I've already dealt with this in response to Question #5. I would add only that philosophy can be used to make children aware of how they are one with all people, and how, on the other hand, they are different from one another. These differences may involve differences of family traditions, manner of thinking, modes of artistic expression, language, skills etc.

What kinds of books are the best among all the books written in this field?

I find it difficult to answer this question, largely because I haven't read English translations of numerous books intended to be novels for teaching philosophy to children, nor have I been able to read those books that remain untranslated.

What methods are used in the books to teach reasoning and judgment? What is the difference between this method and the method for adults?

When preparing teachers to teach P4C, we use the same method as the method to be used by those teachers when they eventually teach P4C to children. That is, the children read excerpts from the novels dealing with reasoning, and they then discuss them, if they have any questions they want to ask. It is true that, with regards to teaching logic, the proportion of questions making use of didactic teaching is likely to be larger than those from other areas of philosophy. But there are a good many portions of logic that lend themselves to being taught by the discussion method: logical fallacies and informal reasoning are examples.

It is suggested that the books be translated, but there are difficulties in doing this. There are ethical values in the novels that the children of some countries would not have sympathy with. Also, there may be a cultural spirit that is inconsistent with the ethical values in different countries. How can this difficulty be overcome?

It is very difficult to try to teach Philosophy for Children in countries where the curriculum has not been translated into the language of those countries. The translation need not be literal, but it should be the same language that the people in that country speak.

Those who try using IAPC materials that touch on ethical or religious values should read *Philosophy in the Classroom* (Lipman, Sharp, & Oscanyan, Second Edition, Temple, 1978). They should also read the instructional manuals for the novels they are planning to teach. The situation is far from hopeless, even in countries where the influence of religion upon education is very strong.

There are some writers in Iran as well as in other countries who are working at writing thinking novels at the children's level. What rules should they follow so that their books conform to P4C standards?

I think you might address this question initially to the International Council for Philosophical Inquiry with Children. If, after reading the ICPIC response, you still have questions you would like me to deal with, I would be happy to give you further suggestions.

The Logic of Relations: Structures in Children's Literature as Channels for Teaching Philosophy to Children

JOHN ROEMISCHER

Most authors sit down to write with *no* fixed design, trusting to the inspiration of the moment; it is not, therefore, to be wondered at, that *most* books are valueless. Pen should never touch paper, until at least a well-digested *general* purpose be established. In fiction, the *denouement*— in all other composition the intended *effect*, should be definitely considered and arranged, before writing the first word; and *no* word should be then written which does not tend, or form a part of a sentence which tends, to the development of the *denouement*, or to the strengthening of the effect.

(Edgar Allan Poe)

Relationships between the characters in children's stories are most often reflections of social and moral problems that call out for clarification and resolution. In books written expressly for children, the pre-established design alluded to by Poe tends to be overtly indicated rather than subtly woven into the text. This one characteristic of children's literature makes it amply suitable for doing philosophy with a general audience, but especially for doing philosophy with children. Such eminent authors of children's books as Arnold Lobel and Leo Lionni have made this practice a benchmark of all their work. Replacing human characters with animal forms simply highlights the formal issue while avoiding the distraction of a recognizable content. Most often, children's books are simply benchmarks in the practice of moral edification.

Thinking is generated in such books through the underlying presence of character or event ratios and by the general presence of binary oppositions which compel relationships to move beyond static balance and into a developmental momentum.

The underlying logic of such ratios, which moves the story scale into and out of balance and thus provides "story-interest," can be expressed as ratios of similarity/difference on the one hand, and transference/resistance on the other. Understanding the underlying logic of relations in story provides the teacher of philosophy for children—a teacher in search of literary material for philosophic dialogue set within the context of a "community of inquiry"—with a structural basis for organizing concepts and building questions for generating philosophic dialogue. The apex of a story comes through as a moment of tension driving toward some denouement, and it is the question whether this tension can be defined and resolved that calls for a determination of the

John Roemischer (jroemischer@charter.net) is Adjunct Lecturer in the Department of Literacy Education at SUNY Plattsburgh. He taught at Brooklyn College in philosophy and history and philosophy of education from 1962-1991.

logic of the story's structure. It is through an attention to this structure as focal point of the story that specifically philosophic issues and skills can emerge. In the present discussion, attention will be paid mainly to philosophic issues; a full discussion of skills needs its own space and development.

Four Pre-established Designs

Relationships which are asymmetric/non-reciprocative. This is by far one of the most compelling principles of design for story construction. It creates a dynamic which moves very strong philosophic disputes (especially moral issues rooted in psychological tensions) to the forefront. Arnold Lobel's "Frog and Toad" stories are masterful examples of this genre. In one of these stories, "The Dream" (1971) the structure consists of a "tension of inverse division" which involves the growing separation of the characters for the benefit of one at the expense of the other. In this inverse growth ratio, one character (often the less altruistic character) prospers while the other character diminishes socially, existentially, even physically. The more the asymmetry expands, the less possible it is for reciprocation to manifest itself in the relationship. The message is clear: the asymmetry of egoism is morally destructive. For many story writers, like Lobel and Leo Lionni, the restoration of balance through the mechanism of reciprocation is a moral necessity. Stories which crash, which make restoration or balance impossible, represent a more postmodern expression of moral cynicism or non-certitude; that is, for a narrow postmodernist moral perspective, the logic of relations which governs human conduct can never get beyond the asymmetric/non-reciprocative condition (an ontological position avoided even by Existentialism).

Traditionally, however, stories which were truly asymmetric/non-reciprocative, unlike "The Dream," were stories which pitted irreconcilable forces against each other, often resulting in the death of the immoral or evil characters—the evil witches or step-mothers of traditional fairy tales; or, they represented a moral didacticism as

lessons intended for the construction of moral conformity in children through the example of punishing the story's character who breaks the rules. In Lobel's story, the imbalance is turned into a dream sequence which hurts no one and which waking reality resolves. A modern "humanization" of children seems to call for stories with less severe designs. But this design is not without its benefits. For it is possible, by way of intensifying the story's cognitive emotions, to begin a story with a resistant structure and resolve the resulting tension by shifting gears into a "softer" set of relations.

From a moral standpoint, the asymmetric/non-reciprocative structure often leaves the story's character in a perplexed or passive position; hence the story often depends on an outside source—an extraordinary force or authority figure—to resolve the tension implicit in this structure: the author as *deus ex machina*. On the one hand, such stories serve to challenge the sense of omnipotence and omniscience which makes children resistant to discipline—and dialogue; hence their traditional popularity. But on the other hand, the story structure can be transmuted so that the story's tension becomes the basis of an internal dialogue, a dialogue which gives the story's characters (and, hence, teachers and children) an opportunity for critical thinking and problem solving. In *A Color of His Own* (1975), Lionni takes as his subject matter a palpable asymmetric/non-reciprocative relationship (the ontic alienation of a chameleon due to its unique color adaptations), and turns thought into dialogue by allowing his characters (two chameleons) to resolve the story's problem conjointly and on their own. The chameleon whose color adaptations rob him of an identity finally achieves some identity in concert with another chameleon: by living together, they will simply change together. This transmutation of an asymmetric/non-reciprocative situation into an asymmetric/reciprocative one is simply the product of a transition made available through investing the story characters with rationality and a capacity for social bonding. The story structure opens the teacher and child to deliberation about reasonable choices in comparable situations, and to social solutions of problems created by the

Im ThINKINg
of a STEgOS

(a TAIL
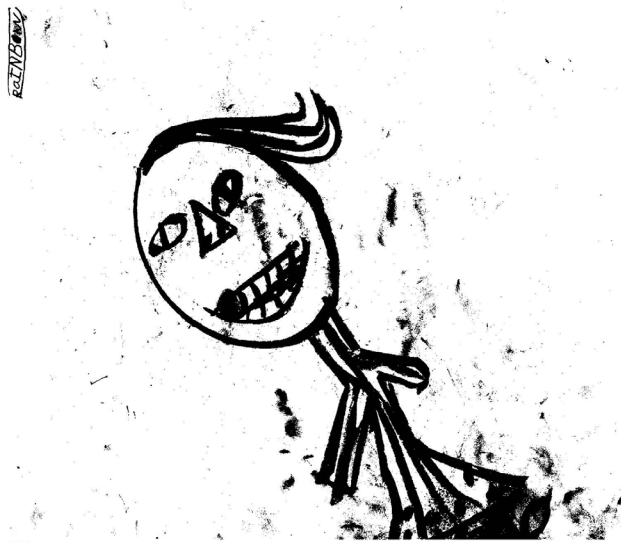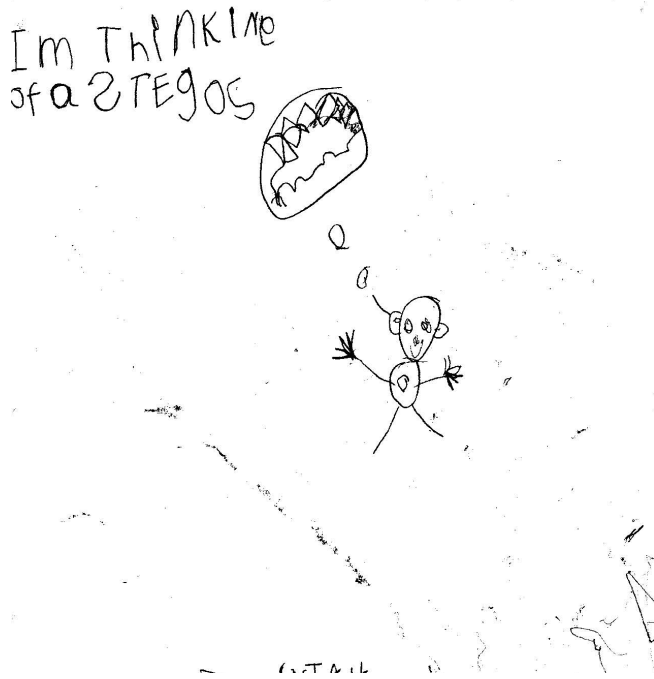
isolated individual. Seen from a theoretical perspective, an examination of different story structures can mitigate the more traditionally harsh moralistic story design used for disciplining children. Some of the most important metatheoretical work in the study of philosophy for children needs to be done in this area of "metareading." And this is a task which requires as much work in linguistics as it does in philosophy.

Relationships which are asymmetric/reciprocative. Many of Lobel's "Frog and Toad" stories are built on this structural base (for example, "A Swim," 1970); so also, such Lionni stories as *Swimmy* (1963). "In "The Dream," Toad dreams that he is growing in size and significance at the expense of Frog, whose size is diminishing into nothingness. Lobel was compelled to resolve Toad's asymmetric/non-reciprocative dream sequence, or the entire moral structure of the relationship between Frog and Toad would have dissolved: Lobel realized that "friendship" is not an unconditional relationship and that it requires equilibration when one partner has illusions of grandeur at the expense of the other. And so the talents of one friend can diminish the ego strength of the other. Stories with "happy endings", which achieve resolution from the inside deliberations of the characters themselves, achieve their dynamic development through the interplays governed by the asymmetric/reciprocative structural principle. Unlike asymmetric/non-reciprocative stories, in which the inner dialectic often demands a solution imposed from outside the characters' control, the structural element of reciprocity, often through

the intervention of reason, transforms serious existential irritations into accidentals in a broader harmonious ontology.

Lionni's story, *Frederick* (1967), reveals and develops the asymmetric distribution of human capacities. Frederick is a member of a group of mice who need to work hard to prepare for winter scarcity; but, notwithstanding the complaints of the others, his declared "work" is an imaginative exploration of the world through language. When winter comes and all the food supplies are depleted, Frederick demonstrates that though his "work" was asymmetric with that of his peers, the poetic products of his warm and nourishing imagination become reciprocative. This opens such larger issues as the difference between modes of production; between the arts and other fields of human endeavor; between education and the distribution of labor. This story structure is one of Lionni's favorites and can be found in *It's Mine* (1985), among others. In *Frederick*, altruism is a social necessity; in *It's Mine*, altruism is something each individual must discover for himself.

Relationships which are symmetric/non-reciprocative. Metaphoric, translational and analogical crossovers utilize this structural principle: they bring linguistic devices into play in order to resolve obdurate differences by imposing a semantic symmetry on what is initially, and ultimately, non-reciprocative. In Lionni's story *An Extraordinary Egg* (1994) three frogs discover an alligator egg; one of them decides to call it a "chicken egg," and when it hatches they call the baby alligator a "chicken." The alligator and one of the frogs become "inseparable friends," and from that point on the alligator is referred to as "chicken" both by the frogs and author. (Note that here the author does not provide a corrective voice from outside.) A bird soon informs this group that the "chicken's" mother is looking for it; one of the frogs and the "chicken" walk a long distance to find the mother. On contact, the mother alligator, recognizing her offspring, says "Come here, my sweet little alligator." But the frog is not deterred: she continues to refer to the baby alligator as "little chicken," and on returning to her friends she reports that the "mother chicken" insisted on referring to her baby as "my sweet little alligator"! The three frogs laugh when one of them exclaims: "Alligator! What a silly thing to say." (Are we, philosophically, back in the nominalism—realism controversy?)

The entire literature on the morality of teaching toleration is simply reducible to this: the historic distinction between "black" children and "white" children makes their condition as "children" symmetric, while their capacity to engage in play and story results in reciprocative relationships. On its own, the asymmetric color relationship drives out the possibility of broad social reciprocity: as in Lionni's *A Color of His Own* (1975), "black" children

transform an asymmetric/non-reciprocative relation into a reciprocative one by adopting the same solution as did Lionni's chameleons: their color difference drives them to "brotherhood."

Relationships which are symmetric/reciprocative. In a homogeneous social relationship, a problem which creates a dramatic event for an individual member or for the entire membership is resolved conjointly by those involved for the benefit of the troubled member or the entire group. From a philosophic standpoint, the issues of altruism and values development come to the fore in stories structured in this way, though these issues are not exclusively governed by this design. Stories governed by this structure involve the unfolding of reciprocal relationships.

Lobel's "The Letter" explores the sensitive reciprocation of friendship which allows small personal eruptions to be resolved by a generosity of spirit. Toad has not received a letter from anyone and is therefore depressed. Frog writes him a brief letter, telling him that they are friends, and, with Toad, awaits its delivery. It is a simple example of self-less altruism unfolded in a symmetric/reciprocative story medium. Though Toad and Frog can be categorized differently, their symmetric (relationally friendly) disposition allows them to share the news of the letter even before it arrives: the underlying reciprocity of spirit defines the meaning of the key term in the title of the book, "friends". Philosophically speaking, defining terms as expressions of conditional or unconditional relationships and reciprocal or non-reciprocal relations is possible here, as well as discussions concerning conceptual meanings vs. dictionary definitions. Is "friendship" a conditional or unconditional relationship? Does it demand "reciprocation"? And does quid pro quo apply to it? How is it different from "brotherhood" or "sisterhood"? Many social relationships can be explored within the symmetric/reciprocative story structure.

Philosophical applications

As already indicated, these story structures can be useful as ways of identifying and characterizing philosophic issues. In fact, it does not seem unreasonable to claim that most philosophic issues can be identified as matters falling into these structural patterns. Philosophers, historically, have seen the connection between story and philosophy and have resorted to story devices in order to philosophize—through allegory, parable, myth, among others. The patterns which create the dramatic interplay in story constitute the very basis of the dialectic tensions which have given philosophy its language: the one and the many, monism and pluralism, absolutism and relativism, theism and pantheism, appearance and reality, realism and idealism, good and evil, and so on.

Some substantive solutions are possible when distinctions are seen against the backdrop of these different structures. If it can be demonstrated that the asymmetric/non-reciprocative structure is the structural basis of both absolutism and relativism, then it is possible to argue that the argument between them is irresolvable. In social absolutism and relativism there is an ontological asymmetry or discontinuity between those governing and those governed; the result is an irresolvable non-reciprocity which makes totalitarianism or nihilism inevitable outcomes. In an analysis of the story use of the asymmetric/non-reciprocative structure, the story's characters and its readers soon discover the futility of assuming extremist positions. Here, the teacher can refer back to the history of ethics and the extreme differences between Hobbes, Rousseau, and Dewey: for example, the extreme positions which distinguish Hobbes and Rousseau from each other concerning the innate nature of the child. For Dewey, a nature which, in Hobbes's world, is asymmetric/non-reciprocative, can be philosophically resolved by moving to an asymmetric/reciprocative position. For Hobbes, the child is innately selfish; for Rousseau the child is innately generous. These positions are absolute and irreconcilable. Dewey, however, finds both characteristics important elements of childhood, and suggests that that is precisely why society intervenes to help the child make transition into a more balanced moral life. In this quasi-Hegelian view, what is asymmetric is the starting point of development; the social context provides the reconcilable medium for achieving resolutions when the asymmetry of human existence goes too far—what Dewey's instrumentalist evolutionism referred to as the resolution of "problematic situations."

Structural differences are central to philosophic disputes; an analysis of these differences can serve to clarify some of these issues. Moreover, they can throw light on the procedures which philosophers take to resolve important controversies. They can serve as a direct link between story and philosophic thought and provide the teacher a conceptual footing for teaching philosophy to children.

REFERENCES

Lionni, Leo (1963). *Swimmy.* New York: Pantheon Books.
(1967). *Frederick.* New York: A.A. Knopf.
(1975). *A color of his own.* New York: Scholastic.
(1985). *Its mine.* New York: A.A. Knopf,
(1994). *An extraordinary egg.* New York: Scholastic.
Lobel, Arnold (1970). A swim. *Frog and Toad are friends.* New York: Scholastic.
(1970). The letter. *Frog and Toad are friends.* New York: Scholastic.
(1971). The dream. *Frog and Toad together.* New York: Harper Collins.
Poe, Edgar Allen (1984). Chapter of suggestions. In *Essays and Reviews* p. 1293. New York: The Library of America.

Student Questions: Developing Critical and Creative Thinkers

ROSIE SCHOLL

Many teachers are aware of a growing demand to develop in students the ability to think both critically and creatively. In recent years the term "critical literacy" has come to the foreground of teacher education and inservice, as systems grapple with the idea of preparing students for active participation in a changing world. To develop in students the capacity to be critically literate or critically aware, teachers need a sound understanding of the process they are attempting to engage students in. Critical literacy requires that teachers and students be "open to the possibilities for reading to be defined as a mode of "second guessing" available texts, discourses and social formations" (Luke & Freebody, 2000, p.6).

This article will refer to "texts" broadly, as the interactions teachers and students may be involved in across texts, discourses and social formations. Discourses and social formations, that is the families, friendships, classrooms and community groups that students participate in, can be seen to supply teachers and students with "texts" to "second guess." Texts (in all forms) should be drawn from within and outside the classroom.

To be critically literate, or able to second guess texts, students need to be involved in comprehending texts (e.g. stories, communication within relationships, movies, artwork, theatre, television, playground antics etc), viewing texts from a variety of perspectives, understanding that authors are situated in social, cultural and political contexts,

and critiquing the meaning offered by the text. Engagement in these types of activities allows the student to develop a critical awareness of the world and how they are situated in the world they live in.

Questions can be used by students as a tool to "second guess" texts and as a way to employ their curiosity and sense of wonder. Questioning, as a method of "second guessing" texts, is then an important activity for both teachers and students to be exposed to and master. As teachers and students question texts they will uncover new perspectives and meanings, which may assist them in understanding the variety of texts and contexts they engage with and within. Questions will engender a sense of inquisitiveness, curiosity and wonder in the classroom. Questions can underpin all the teaching, curriculum and assessment activities that underpin classroom activities. The source of these questions needs to be identified.

Oliver (2002) has offered a number of strategies and ideas to guide the teacher in critically questioning texts. Oliver builds on the notion of critical literacy offered by Luke and Freebody by suggesting that teachers should ask questions to "give their [the students'] minds some critical directions to resist the designated reader position" (p. 37).

In this article I propose that teachers, including early childhood teachers, should build on Oliver's ideas by explicitly teaching students how to question texts, modeling examples of questions for students and providing them with opportunities to develop and enhance their own questioning techniques. Questioning or second guessing texts will then become a "habit of mind" (Costa, 2000a) that is developed in students to assist them in the process of becoming critically literate.

Developing a habit of mind means valuing "one pattern of intellectual behavior over another; [which] implies making choices about which patterns should be used at a certain time" (Costa & Kallick, 2000a, p. 8). Questioning is a habit of mind that students can learn, select and apply, to enhance their development in areas such as comprehension

*Rosie Scholl (*rosiescholl@yahoo.com.au) *holds a Masters in Educational Studies. She was most recently a teacher at Buranda State School in Brisbane, and has been involved in training other teachers across Queensland to teach Philosophy in Schools. She has also been involved in the New Basics Trial as a classroom teacher and critical friend at Buranda State School, and moderator for the Assessment and New Basics Branch.*

and research (Palinscar & Brown, 1984). If questioning is to be valued as a habit worth developing then it should be modeled and practised with students in the early childhood years, that is, preschool to Year 3, and throughout their schooling (Costa & Kallick, 2000b). Teaching young students to question facilitates the development of a habit which assists their literacy development and gives them a tool they can use to "second guess" texts.

If young students are taught to question texts, their education is more likely to be of the type intended by Freire (1972) as "problem posing" rather than "banking education." Friere (1972) proposed, "liberating [or problem posing] education consists in acts of cognition, not transferals of information" (p. 53). My experience is that students as early as Years 1, 2 and 3 are able to engage in a liberating form of education if they are presented with opportunities to engage in acts of cognition and develop "habits of mind," including posing their own questions. If students are explicitly taught how to question and are provided with the opportunities to experiment with and rehearse their developing questioning skills, they are able to participate as the problem posers in their own education.

My intention in this article is to share ideas from my experience of teaching students to question. In outlining my experience and reflections on teaching students to question, the following topics will be explored: explicit teaching, critical and creative thinking, transformation of pedagogy and practical approaches for teaching students to question.

Explicit Teaching

Students can be taught to question texts. To begin teachers must teach them to differentiate between a sentence and a question and to understand the purposes of both. The teacher may begin by asking the students "What is a question?" Students may give simple answers like, "You use a question when you want to ask something or when you need an answer; you use sentences to tell people things." Once students are able to make the distinction between sentences and questions they can then be explicitly taught that there are various types of questions with associated purposes. Some examples of question types are explored later in this article and include questions seeking comprehension, questions requiring factual research, questions to stimulate imagination, and inquiry questions that seek to explore concepts and ideas behind texts. Different questions then, lead to different thought processes.

Questioning texts has the fundamental purpose of engaging students in thinking about texts and issues within texts. To understand the text and associated issues students must have strategies for efficiently probing them.

To be efficient, students must be aware that questions have appropriate contexts for use, such as a search for information, reflection on practice or imagining new ideas. It is therefore necessary for the teacher to understand and make explicit to students that different types of questions facilitate different answers. Students need to understand where their questions will lead them—that is, what type of answer are they likely to receive. Students should be able to predict what the answers to their questions will offer them— for instance will the answers to their questions supply them with facts and information, critique of practice or new designs and creations? Subsequently students will be able to devise questions that enable them to engage with texts in both critical and creative ways.

Critical and Creative Thinking and Its Impact

Marzano et al (1988) suggest that critical thinking is primarily evaluative, while creative thinking is primarily generative. Students can devise questions to assist them and their classmates to think critically and to evaluate texts. For example, "Why are fairytales always about beautiful people?" might lead students to understand that beauty is not the sole characteristic for making a character's life worthwhile. Students can also employ their questioning skills to help them seek alternatives to the ideas supplied within the text and to generate new ideas. For example, "What would it be like to live in a world without beauty?" is a question that may lead students to create narratives or design pictures or models of the consequences of living in a world without beauty. When students know how to use their questions appropriately in various contexts, they can lead themselves and their classmates to thinking in both critical and creative ways.

To be able to think critically and creatively helps students analyze and create texts. Luke and Freebody (2000) have suggested "critical literacies - in all their forms – enable teachers, students and communities to explore alternative ways of structuring practices and texts to address new cultural and economic contexts" (p. 6). Students need to be given the opportunity to experiment with and apply questions that will help them to critique and create texts, as they address issues and inequities within their families, classrooms and communities.

Social action and change can occur when students are involved in questioning (critiquing and creating) practices in their classroom, school and community. An example of this would be students who question texts and community members about pollution problems in the streets and waterways around their school, in order to understand the problem and create a solution. Students can be involved in researching the problem, giving presentations to create a critical awareness in the

learning and thinking (Scholl, 2000). From the following list the Year 1 students chose the first five questions to focus on, the Year 2 students chose the first eight questions to focus on and the Year 3 students chose to use all the questions as stimulus for their reflection activities. The students generated all the following questions.

Table 1

Year 1, 2, 3   Questions about thinking and learning.

Year Level	Questions
Year 1/2/3	1) What do I have to do?
	2) Have I done any work like this before?
	3) Am I working quietly, and concentrating?
	4) Is my work neat?
	5) Am I working patiently?
Year 2/3 additional questions	6) Is this interesting or fun?
	7) Do I know what the answer is?
	8) Do I know how to do this work?
Year 3 additional questions	9) Do I need to ask for help?
	10) Am I going to get this finished on time?
	11) Does my work have mistakes that need fixing?
	12) Have I done my best?
	13) Is my work improving?
	14) What is my reward?

community, using various media to do this, and working with action groups to address such problems. Such activities are more relevant for the students if they have actually developed the questions to understand and analyze the contexts and to create solutions themselves. This type of activity is aimed at engendering in students the self and social empowerment (McLaren, 1995) required to co-construct appropriate learning experiences and lives for themselves and others in their circle of influence (Covey, 1989).

Problem posing or questioning can also be instrumental in the students' personal reflection. Teaching students to critique their own behavior and the social practices in the classroom, playground and community are aspects of the student's life, which they are able to reflect upon. For example, the use of self-questioning can help students monitor their behavior. Such questions can assist them in focusing on their own ability to work in groups and alone, to complete tasks, to work and play effectively, and to problem solve. Such self-monitoring is metacognitive in nature (Flavell, 1979). For example, students I have worked with in Years 1, 2 and 3 were able to construct the questions in Table 1, which were used in regular reflection about their own

Reflection may involve students in questioning their personal strengths and weaknesses and subsequently setting goals for improvement. Students' questions can be used to engage them in reflection on their own values and beliefs as they question texts that target issues, which may affect them personally and socially—for example trust, exclusion, racism or bullying. Instead of solely accepting the teacher's direction about these issues, students are able to start thinking critically and creatively about them. This thinking should be shared in dialogue with the whole class, in small groups and between the child and teacher. Using the students' questions offers an alternative to the practices of "banking education." Students are able to question and pose problems relevant to their own lives rather than trying to figure out the teacher's solution for a problem that the teacher has identified.

When students are involved in posing questions there is inevitably and alteration of the power structures in the classroom as the traditional teacher question/ child answer

routine (Cazden, 1988) is transformed. The teacher must be ready to be questioned, to not know the answer and to be surprised by the unseen alternatives that arise out of teacher-facilitated dialogue where students are able to construct their own meaning. For example, students in a multi-age Year 2/3/4 class I taught were seeking the answer to the question, "What is a lot?" posed by one of their classmates. The initial response in my mind was "one million." The students, in contrast, were able to create alternative responses like: "It's a bit more than a bit but not as much as a heap…" and "It depends on what you are talking about because a cell in your body is very small but if it has disease in it, that can be a lot." Their alternatives showed that the concept of "a lot" was problematic and that they were able to critically analyze the meaning of the words to gain a new depth of understanding about the concept. The teacher needs to support students as they seek to answer the student's questions and problems together. There will then be a greater possibility for students and teachers to be involved in "problem posing" education.

The Transformation of Pedagogy

Traditionally the power structures within the classroom position the teacher as questioner. These structures assume that the teacher has the knowledge and the questions to test the students' received knowledge. To position the student as questioner means the teacher will not have all the answers to the students' questions and must have the good grace to believe that sometimes the answers exist outside the teacher's mind and experience. The teacher must relinquish some power if the children are to genuinely be questioners and "problem posers" within the classroom. This shift in power may be destabilizing at first but with reinforcement of consistent, respectful boundaries both teacher and students will be positioned to engage in new relationships and new ways of learning.

The teacher should be open to an element of surprise as they learn that students are able to question, think and reasoning in complex and sophisticated ways. In addition this element of surprise may invoke a sense of wonder as the teacher, who so desires, learns from the students. For example, after reading a story titled "The Bird's Nest" (Cam, 1993), as the stimulus for a philosophy lesson (described further in this article), the students in a Year 2/3/4 class explored the concepts of discipline and punishment. They were eager to figure out why the girl in the story had not been disciplined or punished by her mother – a topic from their own experiences in relationships with their parents and guardians. After much dialogue, in which the words discipline and punishment were used interchangeably, the following ideas were articulated in a way that made me

reflect on how discipline and punishment had played out in my own life as a child and adult.

Student 1:	What is the difference between discipline and punishment?
Student 2:	Punishment has anger but discipline doesn't have to have anger.
Student 3:	Yeah – and guilt is the weapon of discipline and punishment.
Student 4:	Discipline is something you can do to yourself. Punishment is something they do to you.

Such interchanges provide the ideal stimulus for the teacher to reflect on the changing dynamics in his or her own pedagogy. Some teachers capture these reflections through a journal, anecdotes of classroom dialogue and annotated student work samples (Kemmis & McTaggart, 1988). The efforts of the teacher to capture his or her own personal reflections are often useful in tracking the shift in power, their personal transformative learning and the gains the students make in their ability to form and use appropriate questions and to think critically and creatively.

The following two excerpts from my own journal have allowed me the opportunity to reflect on and learn from student questions and subsequent dialogue of the students. The following examples were based on student questions and involved issues that were engaging for all students in the classroom. They uncovered deep levels of understanding.

The first excerpt comes from a Year 1/2/3 class. At the time I had been explicitly teaching the difference between open and closed questions in a series of lessons (Scholl, 2000). The students had been developing their own open and closed questions in response to texts they were reading and to a game that I had developed (Figure 4). This excerpt of dialogue taught me that open and closed questions were not discreet entities; rather they can be placed along a continuum depending on the context, on how the question is posed and on how the other person perceives the question and subsequently chooses to answer.

Student 1	Question: Are there any questions that are open and closed at the same time?
Student 2:	Yes. I think "Who am I?" would be an open and closed question at the same time. If it was a closed question I could tell you my name but if it was an open question I could tell you I am my mother's daughter but also a student and an artist and there might be many different answers for the question.

The second excerpt was taken from a Year 2/3 class during a Philosophy lesson. We were exploring the concept of trust through the dialogue. The stimulus was a student question, "What is trust?" We had come to a stalemate in the dialogue,

and seemed unable to come up with a deeper understanding of the concept of trust. I suggested to the students that we use analogy to deepen our understanding of trust. First, I explained that an analogy is a way of describing the common features of two items so we gain a more descriptive and deeper understanding of at least one of the items. I modeled an example of an analogy for trust saying, "Trust is like an egg. You hold it very gently but securely in your hand and carry it very carefully with you. One day, however, you trip and the egg falls and breaks. It seems impossible to put it back together." I asked if the students could develop their own analogies for trust. Their analogies provided new insights for all of us about the nature of trust (see Table 2).

Table 2

Analogies for Trust. Year 2/3

Student 1	Trust is like a cake and love and kindness are the ingredients.
Student 2	Trust is like a candle and if you blow it out you might not be able to light it again.
Student 3	Trust is like a block tower and if you knock it down it might take a long time to build again.
Student 4	(building on the analogy of Student 3) Yeh - and if you knock the tower down and then you rebuild it again you might have to wait for the cement to dry and that could take a long time.

These excerpts from my own journaling provide examples of a new depth of understanding for both the teacher and students. This depth of understanding can be created in the classroom when the students' knowledge and questions are used as a starting point for learning. The students need to know that what they say and ask will be valued and used in the learning experiences. The classroom then becomes characterized by communication rather than communiqué; problem posing and dialogue rather than teacher-driven curriculum delivery (Friere, 1972).

To be in a classroom where dialogue exists between teachers and students, and students and students, involves a transformation of pedagogy, which takes time, and is often a confusing process of trial and error. Once this journey of using students' questions to direct student learning has begun there is no turning back, as the process will alter the power structures within the classroom. Students who are explicitly taught to question, who have question types appropriately modeled for them followed by the opportunity to develop and use their questions in the classroom, will accept the challenge of being involved in the direction of their own education.

In what follows I share, from my own practice, some practical approaches for teaching young students to question,

including explicit teaching of questioning skills, using student questions in learning and assessment tasks, the uses of student questions within a Philosophy for Children class, the place of student questions within the Reciprocal Teaching strategy, and games which include the use of student questions. The following examples have all been implemented successfully with students in early childhood classes.

Practical Approaches for Teaching Young Students to Question

Explicitly Teaching Questioning

It is important to explicitly teach young students the nature of questions and how to develop them. In my research (Scholl, 2000) the students worked on understanding the difference between a question and a sentence. This involved looking explicitly at the punctuation, grammar and intonation we use in the creation of both. The students were then asked to suggest question words, that is, to suggest words that they regularly use to start questions. They were able to suggest the question words noted in Table 3. These question words were displayed on a poster in the classroom.

Table 3

Question Words

What	Who	Could	Do	Will	Didn't	Does
When	Which	Are	If	Isn't	Aren't	Shouldn't
Where	Should	How	Can	Don't	Did	Wouldn't
Why	Would	May	Is	Was	Shall	Couldn't

I modeled examples of questions using each of the suggested words. I then asked for further suggestions from the students. Each question, suggested by the students, was written on a poster and displayed in the classroom. Once I had established that the students were capable of developing questions the focus moved to discriminating between question types, namely closed and open questions.

Closed questions call for a correct answer. Open questions require further investigation and may have more than one answer that satisfies the question (Cam, 1995). To help students learn to differentiate between closed and open questions I wrote the closed questions that they generated on red cards and the open questions on green cards. The colored cards offered visual clues similar to traffic lights. Open questions do not have a single right answer. They are like a green light to thinking and allow alternative, analytic and imaginative answers. Closed questions, on the other hand, are the red light to the mind. They are comprehension questions, which require a search for facts. Generally the

I am thinking about snakes.

SOPHIE

answer to closed questions can be found in the text (e.g. the book, the play, the program, the experiment) or through research. Young students are able to give examples of both open and closed questions and recognize appropriate situations for their use if these types of questioning are explicitly taught and modeled (Scholl, 2000). Students should be given opportunities to experiment with their new questioning skills and receive feedback on their attempts. This can be done through individual, small group or whole class reading and questioning.

Once students have learned to establish the appropriate use for closed and open questions teachers can assist students in making further distinctions between these question types. Closed questions can be further classified into comprehension questions and research questions. The answers to comprehension questions will be found in the text, however the answers to the research questions may require the student to investigate other sources to find facts that exist outside the text.

Open questions can lead the students to either critical or creative outcomes. Open questions can take the form of inquiry questions, which require students to be involved in critique of ideas, analysis and reasoning. Quite often inquiry questions require students to reflect on personal experience or empathically engage in the practice of walking in another's shoes, in another time and space. This may help the student to understand others' perspectives and how the text is positioned in relation to them and their world. Creative thinking, in contrast, is a response to questions that

require students to employ their imagination to propose alternatives, think of new endings, answer "what ifs …" and to generate new ideas. These open questions may lead to amendments and changes to texts and their conventional or prescribed outcomes, to the adaptation of ideas, or to the creation of brand new ideas, designs and plots.

This understanding of question types is well illustrated in the diagrammatic model shown in Figure 1 (adapted from Cam, 2003). The x and y axes in the Question Quadrants represent two continua. The x-axis refers to continua on which the poles are closed and open questions and the y-axis refers to continua on which the poles are comprehension and intellectual questions. In Quadrant 1 the questions are closed comprehension questions. The answers to these questions can be found in the text. In Quadrant 2 the questions are open comprehension questions, they are questions based in the text that invite alternative responses and not one single right answer. In Quadrant 3 the questions are closed intellectual questions. Although these questions have a correct answer, answering them requires some form of research or analysis of data. In Quadrant 4 the questions are open intellectual questions. These questions are thought provoking. They are often the "big questions" in life and they may have many appropriate answers. They are the questions that stimulate philosophical dialogue (though questions from all the quadrants can be adapted move into the fourth quadrant). If students are asked to apply this model to the fairytale of Cinderella they may suggest questions such as:

Quadrant 1	How many stepsisters did Cinderella have?
Quadrant 2	What happened to Cinderella's stepsisters after she married the Prince?
Quadrant 3	How many real princesses lived in poverty in the last 100 years?
Quadrant 4	What is happiness?

Students can be asked to both devise questions for each quadrant or to place previously devised questions in the quadrants. This can be done in small groups on worksheets or as a whole class marking out the four quadrants on the floor of the classroom. Teachers need to help students understand that all questions have their purpose and should be used and applied appropriately. This may be best done if students are given various models for understanding question types and sorting, classifying and appropriately applying various types of questions.

Another way of differentiating between question types (Figure 2) is offered by Coley and DePinto (1989). I used this exemplar with students in a Year 1/2/3 class to classify and model the six different types of questions for students

(Scholl, 2000). After modeling the question type for the students they were given the opportunity to offer questions that could be placed in the same category. Posters were made using student examples of each question type and these posters were hung around the classroom. Once some student examples were available the students read texts together in pairs or small groups and questioned each other, focusing on the text at hand and a different question type each week. For six weeks we looked for emerging patterns by comparing and contrasting each question type with previously learned question types. Once learnt the question types were generalized and used in many classroom activities from class meetings, to research activities to establishing criteria for tasks that the students were to complete. Using this model the students felt that it was the judgment questions that would lead them to philosophical dialogue. They were able to determine this through comparison and contrast of the question types with the knowledge of the types of responses they were seeking in a philosophical dialogue.

In addition to providing models for asking and answering questions the teacher must provide real-life and life-like activities for the students to practice their new skills. Students' questions can be used during reading activities, for philosophical inquiry, as a stimulus to journal writing and homework activities, when drafting interview questions or as a springboard for planning classroom learning activities and/or assessment tasks. If students are involved in setting questions for the learning activities in the classroom they begin to be co-designers of the explicit criteria for judging or evaluating their performance throughout the learning and assessment tasks.

Learning and Assessment Tasks

Learning and assessment tasks can take many different forms and may be devised as one and the same item. Student questions can be used in the development of both learning and assessment activities. Students can provide questions that can be used to direct and review both the content and process of the student work. Their questions can focus on the "what" and/or the "how" of the learning and/or assessment activities. The students' questions can form all or part of the explicit criteria for the learning activities or assessment tasks to make them transparent and meaningful for students.

The students may, in partnership with the teacher, be included in developing the questions which seek the content, "the what" of their learning and assessment. The teacher will need to guide the students as they set up their questions, to ensure that there is sufficient depth and breadth of content knowledge (used and sought) and processes students will be involved in. Students may be involved in developing the questions for a research activity to learn about content areas such as plants, animals, communities or healthy lifestyles. Alternatively, they may ask questions that would form the steps of a design process for creative, artistic or scientific pursuits. Students can also be involved in questioning and seeking answers to the philosophical basis of the knowledge they are seeking.

The questions, however, do not always need to be focus on the content. The students' questions can facilitate student reflection on how effectively or efficiently they worked. These types of questions may account for student progress throughout the learning or assessment activities. They may be used in either peer or self-assessment. Answers to these "how" questions may be answered by using a form of Likert-

scale (see Figure 3) where smiley and sad faces represent the poles. The smiley faces indicate a very positive or positive response to the question and the sad faces indicate a negative or very negative response to the question. Questions that students may develop to accompany the Likert-scale might include: "How happy am I about the time I took to complete the task?" "How does my contribution to the group make me feel?" "How do I feel about the effort I put into the task?" The list of questions should help the student to reflect on how they worked. As the students become more proficient as writers they can add their reasons as to why they have chosen a particular smiley/sad face. The students who are less proficient as writers can explain verbally to a teacher, teacher aide or each other as to why particular smiley or sad faces were chosen. Further reflection can be stimulated by open questions suggested by the students, such as: "What could I do to help me finish this task?" "How could a better outcome for the group be achieved?" "What were the surprising parts of this work for me?"

Additionally, space for drawings that might answer the questions should be provided for young students as an alternative mode for answering questions and reflecting on how they have worked. This will mean that the process of self-review and peer-review does not become heavily weighted as a writing task, which can be laborious for young students or even deter them from engaging in the reflection activity. If students are involved at least as co-developers for both types of questions (the what and the how) for learning and assessment activities, they are more likely to have a coherent understanding of the questions and the purpose of the questions, which should positively alter the nature and quality of their response. Students are more likely to be engaged in seeking the answers to questions they have developed, thus responding to their own curiosities. Students can record the questions and answers in various ways (e.g. posters, cards, questionnaires, booklets, multimedia etc) to suit the specific learning activity or assessment task. They then need time to refer to and reflect on them during their learning.

Following are three practical ways to use student questions in the classroom. Student questions for the basis of philosophy lessons, they are one of the integral steps to the Reciprocal Teaching approach and are essential to a number of games that can be used in various ways in the classroom.

Philosophy for Children

Philosophy for Children is a literature-based approach to teaching children thinking and reasoning skills. Many stories have been purposefully written for teaching philosophy. Other texts may also be appropriate. The central tenets of Philosophy for Children are discussed by Lipman, Sharp and Oscanyan (1980). They include improvement in reasoning skills, encouraging children to think for themselves, development of creativity, personal and interpersonal growth, development of ethical understanding and development of the ability to seek meaning.

To engage in a philosophical dialogue students need to know that their questions have various purposes and that very good philosophical dialogue is facilitated by well thought out inquiry questions that pertain to Quadrant 4 (Figure 1). For example, following the reading of a familiar text such as "Beauty and the Beast," students with considerable knowledge of question types, may construct the "big questions" that are central to the text and to life such as:

- What is beauty?
- Is it right to judge people by how they look?
- Are all beautiful things good and ugly things bad?
- What is love?
- Are love and beauty connected? How?

The more students know about question types, the easier it is for them to shape questions that will lead them into philosophical dialogue. Students identify the "big questions" in life and through dialogue and concept development activities they are able to explore the answers to these "big questions." They then try to reach a collective understanding of or an answer to the them through the dialogical process of a community of inquiry (Cam, 1995). This dialogical process is student centered. The teacher's role in the community of inquiry is described by Cam, through the use of the analogy of teacher as conductor.

> You need to coordinate and enhance the performance. You may deed to be vigorous at one moment, but restrained at another. You need to make sure the children are listening to each other, and ensure that everyone has the chance of being heard (p. 41).

In philosophy lessons it is the students' questions that are used as the impetus for class dialogue. Depending on the stimulus used the students' questions will traverse across the philosophical domain which in fact is the connective tissue that underpins all the learning taking place in the classroom. The students and teacher may find themselves in dialogue about metaphysics, logic, aesthetics, ethics or epistemology. Students can, with practise, recognize which categories their questions are falling into, that is, which direction the philosophical dialogue will take them. They may realize in their philosophical dialogue that there are issues in their classrooms and local communities that can be addressed through positive action, similar to the creek example given earlier in this article. It is this philosophical questioning and dialogue that captures what is meant by communication as

compared to communiqué. The students are able, through their questioning, to locate and respond to issues in their families, classrooms and communities, in respectful and meaningful ways. Their knowledge and learning is student centered and co-constructed with other students and the teacher.

It is important to remember that it is the students' questions that drive the community of inquiry. Without the students' inquiry questions (Quadrant 4) the community of inquiry will lack depth and connectedness to the lives of the students. Thus explicitly teaching children to develop various types of questions is the key to facilitating a deep and engaging community of inquiry. Students need opportunities to identify question types and practise their new questioning skills so they can apply them appropriately. Practising question types can be facilitated through student participation in structured activities like Reciprocal Teaching and games.

Reciprocal Teaching

Reciprocal Teaching (Palincsar & Brown, 1984), has been shown to have efficacy in improving students' reading comprehension (Rosenshine & Meister, 1994), and can be used by students of varying reading ability with appropriate partners. Palincsar and Brown (1984) devised the Reciprocal Teaching strategy to improve students' reading comprehension. Students are required to work in pairs as they progress through the four explicit steps of summarizing, questioning, clarifying and predicting. A set of cards or some form of reference sheet on to which the steps are recorded may assist students to attend to the steps in the sequence.

Summarizing. This step requires students to identify, paraphrase, and integrate important information in the text. Following reading the students reiterate the main points within the text in order to check that the meaning of the text has been clearly understood. Once this is done the students can move onto the second step, which is questioning the text.

Questioning. Questioning is based on summarization of the text. Students cannot generate appropriate questions without knowledge of the sequence, content and meaning of the text. The types of questions asked are mainly closed comprehension questions, which will assist the students in uncovering meaning or knowing the text. Once these questions have been answered the students will have established a satisfactory understanding of the text, however, there may still be aspects of the text that require further clarification.

Clarifying. This step requires students to attend to the aspects of the text that are barriers to the student's access to the meaning of the text. These barriers may include new vocabulary, new referent words and unfamiliar or difficult concepts, pictures, maps or diagrams. Identifying these barriers to meaning, signals to the reader to reread, read ahead, or seek help outside the text. Once all the barriers have been addressed, and reintegrated into the meaning of the text, the student will have a deep understanding of the text and should feel confident to predict further parts of the text.

Predicting. The fourth step in Reciprocal Teaching requires the reader to hypothesize about what the author might discuss next in the text. This provides a purpose for reading on: to confirm or disapprove the hypotheses. Hence, an opportunity is created for the student to link new knowledge they have encountered in the text with the knowledge they already possess. It also facilitates the use of text structure as students learn that headings, subheadings, and questions imbedded in the text are useful means of anticipating what might occur next.

Questioning as an integral part of Reciprocal Teaching addresses the first quadrant in Figure 1. Other questions may specifically address different quadrants and they can be developed through explicit teaching and reinforcement through the use of games.
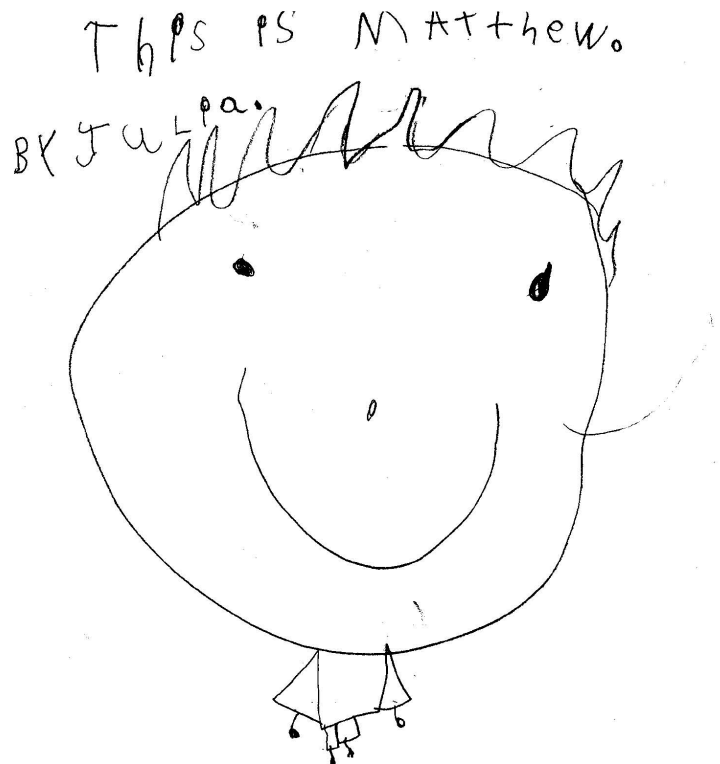
Games

Two such games are *Trivial Pursuit* (Horn Abbott International, 2003) and *Brain Quest* (University Games, 1993). These games provide questions for the players to answer. In the case of *Trivial Pursuit* the questions are closed questions. *Brain Quest* on the other hand involves students in answering questions that challenge the students to compare, contrast, visualize, interpret and solve problems – that is, to think. It is important that the students are cognizant of the type of question they are attempting to answer. Knowing the purpose of the question is essential to developing an appropriate answer.

Once students have mastered such games the next step is to involve them in games where their own questions are incorporated, for example *Celebrity Head* and *Guess Who* (Milton Bradley, 2001). In *Celebrity Head* three students sit at the front of the room. A headband or card with a celebrity's name written on it is placed on or above the student's head. The student does not know what name is written on the card. The student wearing the headband has to structure closed questions, which allow the audience to answer "yes" or "no" to give information about the writing on the headband. The questions generally start simply, such as, "Am I a girl?" "Am I on T.V.?" or "Am I a sports person?" If the student receives a positive answer from the audience they are allowed to ask another question. If not then the next contestant may ask their question. The first contestant to work out who or what they are, is the winner. Similarly, in *Guess Who* the questions are closed questions, which are asked in order to sequentially eliminate faces from the game board. For example the player will ask, "Does the person wear a hat?", "Have a beard?" or "Have blue eyes?" In both games the questions are closed, seeking correct information. Generally the games are very short which is useful maintaining the students' interest and motivation.

Figure 4 depicts a game I designed to encourage children to develop their own open and closed questions (Scholl, 2000). We called this game "The Open and Closed Question Game." This game is best played with about six to ten players so that every player has frequent opportunities to develop and answer questions. To play you will need: a game board, red (closed) and green (open) cards on which to write questions, a pencil, an eraser, enough counters for the number of players and a dice. The rules of the game are outlined in Figure 4.

Games such as those referred to in this section can be used to develop cognitive skills including questioning, in a similar way that ball games develop hand-eye coordination. The students generally enjoy the games and are mostly oblivious to the cognitive demands involved in constructing appropriate questions, as they seek to outwit one another or to develop questions about content they are passionately

connected to. Students should understand the processes they are involved in during these games. It is when they are metacognitive about the processes they are involved in that they are able to generalize these skills in order to be construct meaning in critical and creative ways.

The construction of meaning—be it in comprehending texts, philosophical dialogue or other classroom learning and assessment tasks—is essential to student learning. Using student questions, as well as, or instead of teacher questions may mean an alteration of power in the classroom; however, this is to be seen as positive where the student has the opportunity to become a self-directed, interdependent learner. Classrooms where student questions are central to learning are characterized by questions, dialogue, games and learning and assessment activities that satisfy the inquisitiveness and curiosity of the learner themselves.

Conclusion

Becoming literate is a great achievement for young students and it enables them to access texts that can facilitate their learning in many other areas of the curriculum. Students who are critically literate are able to gain more from their learning as they make connections between texts and contexts, explore alternatives and position themselves to critique and create ideas involved in positive social action and change throughout their lives.

Critical literacy demands that students are able to appropriately ask and respond to questions. Even young students can be taught to question, to use various types of questions in appropriate contexts and so enhance their ability to be critically literate. Students can also be taught appropriate ways to respond to questions. To facilitate these competencies teachers need to be explicit in the teaching and modeling of questions and responses, as well as provide the opportunities for the students to develop questioning as a "habit of mind."

If students are encouraged to question there will be a change in the classroom dynamics, in the positioning of power within the classroom and subsequently in the pedagogy. These changes are viewed as positive for a number of reasons. Firstly, allowing students to have more power in setting the direction of their learning promotes student engagement in learning. Secondly, when students are able to design questions for specific purposes they are able to act in a metacognitive manner about the type of activity and thinking that they are involved in, which may range from a search for facts to a philosophical dialogue. Student questions can also provide the stimulus for reflection on both student and teacher knowledge and practices. Questioning then can kindle a culture of learning, curiosity and reflection that positions the participants (students and teachers) to be

lifelong learners and active citizens both inside and outside their classrooms and workplaces.

Special thanks to Associate Professor Christa van Kraayenoord, of the School of Education at the University of Queensland for her many suggestions and editing of this article.

References

Cam. P. (2003). The question quadrant: Improving the quality of students' questions. *Critical and Creative Thinking, 11*(2), 61-64**.**

Cam, P. (1995). *Thinking together: Philosophical inquiry for the classroom*. Sydney: Hale & Iremonger.

Cam, P. (1993). The bird's nest. *Thinking stories 1: Philosophical inquiry for children.* Sydney: Hale & Iremonger.

Cazden, C.B. (1988). *Classroom discourse: The language of teaching and learning.* Portsmouth, NH: Heinemann.

Coley, J.D., & DePinto, T. (1989). *Reciprocal teaching: Theme and variations*. (Eric Document Reproduction Services No. ED 308 477).

Costa, A.L., & Kallick, B. (2000a). *Habits of mind: Discovering and exploring*. Alexandria, VA: Association for Supervision and Curriculum Development.

Costa, A.L. & Kallick, B. (2000b). *Habits of mind: Activating and engaging*. Alexandria, VA: Association for Supervision and Curriculum Development.

Covey, S.R. (1989). *The seven habits of highly effective people: Powerful lessons in personal change.* New York, New York: Fireside.

Flavell, J.H. (1979). Metacognition and cognitive monitoring. *American Psychologist 34*, (10) 906-911.

Friere, P. (1972). *Pedagogy of the oppressed*. Camberwell, VIC: Penguin.

Horn Abbott International Ltd. (2003). *Trivial Pursuit*. Eastwood, NSW: Hasbro International Inc

Kemmis, S., & McTaggart, R. (1988). *The action research planner*. Geelong, VIC: Deakin University Press.

Lipman, M. (1989). *Pixie*. Hawthorn, VIC: Australian Council for Educational Research.

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

Luke, A. & Freebody, P. (2002). A map of possible practices: Further notes on the Four Resources Model. *Practically Primary, 4*(2), 5-8.

Marzano, R.J., Brandt, R.S., Hughes, C.S., Jones, B.F., Presseisen, B.Z., Rankin, S.C., & Suhor, C. (1988). *Dimensions of thinking: A framework for curriculum and instruction*. Alexandria, VA: Association for Supervision and Curriculum Development.

McLaren, P. (1995). *Critical pedagogy and predatory culture: oppositional politics in a postmodern era.* London; New York: Routledge.

Milton Bradley, (2001). *Guess Who*. Newport, Gwent: Hasbro International Inc.

Oliver, N. (2002). Questioning reading practices. *Practically Primary, 7*(3), 36-38.

Palincsar, A.S., & Brown, A.L. (1984). Reciprocal teaching in comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*(2) 117-175.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of research. *Review of Educational Research, 64* (4) 479-530.

Scholl, R. (2000). *Young children's questions: Theory and applications*. Unpublished masters thesis, University of Queensland, Brisbane, Australia.

University Games (1993). *Brain Quest*. Merlo Park, CA: Workman Publishing Company.

Figure 1: The Question Quadrants

	Textual Questions	
Closed Questions	1. Reading ComprehensionThe answers to these questions are found in the text (stimulus material).They involve the learner in basic comprehension activities.	2. Literary Speculation The answers to these questions require creativity based on the text. These questions involve the learner in creative thinking, seeking alternatives and proposing solutions or new outcomes.
	3. Factual KnowledgeThe answers to these questions require a search for facts and knowledge related to the text. The facts may pertain to a specific discipline or across disciplines.	4. InquiryThe answers to these questions promote deep thought and integration of experience and knowledge. These questions are thought provoking and are often referred to as the "big questions."
	Intellectual Questions	**Open Questions**

(Adapted from Cam, 2003)

Figure 2: Thinking Through Question Types

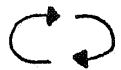# Thinking Through Question Types

R Recall   Similar   Difference

Facts   Likeness   Contrast
What Follows   Comparing   Comparing
Detail   Common Features   Differentiate
Summary

(Adapted from Coley & De Pinto, 1989)

Ex Examples   Evaluation/ Judgement

Cause - Effect

Cause   Categorise   Is it right?
Effect/Result Support   Judge
Predicting   Classify   Rate
               Generalise   Weigh up

Figure 3: Likert Scale

Very happy	Happy	Neither happy nor sad	Sad	Very sad

Figure 4: Open and Closed Question Game
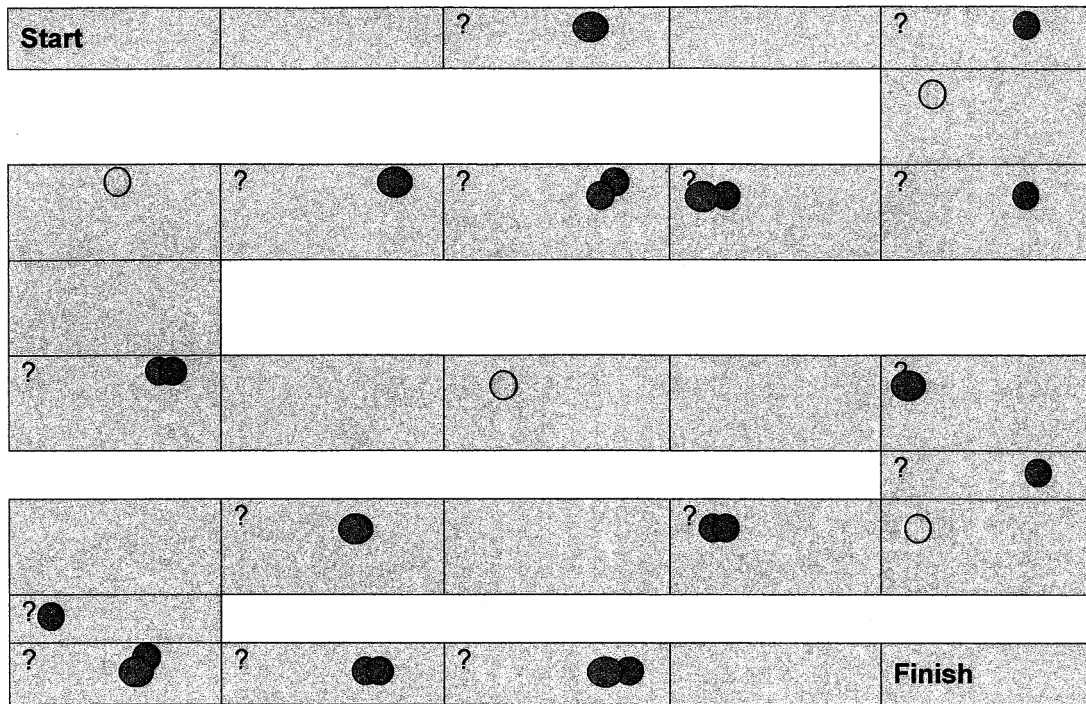
This game is best played with about six to ten players so that every player has frequent opportunities to develop and answer questions.

To play you will need: a game board, red and green cards on which questions can be written, a pencil, a rubber, enough counters for the number of players and a dice.

Game rules

If you land on a:

Object	Action
○	Move to the next yellow dot.
?●●	Make up a new open question for the game. Share it with other players to check that it is an open question.
?●●	Make up a new closed question for the game. Share it with other players to check that it is a closed question.
?●	Someone asks you an open question from the pack of cards.
?●	Someone asks you a closed question from the pack of cards.

The first person to the end of the game board is the winner.

(Scholl, 2003, p. 118-119.)

NOTES FROM THE FIELD

The following discussion took place in April 2004, in a fifth grade classroom at a public school in Montclair, New Jersey. It was facilitated by Dr. Nadia Stoyanova Kennedy (nadia.kennedy@verizon.net) then a doctoral student in mathematics education, and was one of 20 conversations she conducted with the same group over a period of nine months, all of them centered on logical and mathematical paradoxes. Also present as a discussant was Stacey Heuschkel (heuschkel@comcast.net), the classroom teacher, who regularly conducts Philosophy for Children discussions herself.

FIFTH GRADERS DISCUSS THE LIAR PARADOX

Facilitator:	Today we're going to discuss this question: A Montclairian is saying "All Montclairians are liars." Is he lying or is he telling the truth?
Sally:	O.K. Now, I think he's a liar, because my mom is a Montclairian and she's the most truthful person.
Voices:	Oh my god, oh, no!
Sally:	No, no. She's a Montclairian and I bet…O.K. Not all Montclairians are liars, so, he's a liar.
Nellie:	But, even he's a Montclairian and he said that all Montclairians are liars…Because he's from Montclair, so he's lying. So Montclairians are not all liars.
Facilitator:	Let's go slowly. O.K. Bill, do you want to respond to Nellie?
Bill:	Well, if he says… I agree with Nellie, but the question…She said that if all Montclairians are liars and he's a Montclairian, he's lying. But if he's lying and if he's telling the truth that means that he's lying about all that…and it's going on forever. Because you can't say that he's telling the truth….
Teacher:	Wait. He's a Montclairian and he said that all Montclairians are liars.
Bill:	Exactly.
Teacher:	And Nellie said he was lying and you're saying he's telling the truth.
Bill:	No, he's saying . . . .
Nellie:	He's lying and he's telling the truth.
Bill:	But, then it goes on forever.
Nellie:	Exactly.
[Multiple voices]	
Facilitator:	O.K. Let's ask somebody to summarize.
Victor:	O.K. Sally, Nellie, and Ben are trying to say that if he said that all Montclairians are liars he must be lying, since he's a Montclairian too.

Facilitator:	But, do we know whether "All Montclairians are liars" is a true or false statement. What if the statement "All Montclairians are liars" is true?
Bill:	Well, if it's true and he's a Montclairian, then he's lying , but he's telling the truth..
Facilitator:	So he's lying and he's telling the truth at the same time. Is it possible?
Voices:	No, but….
Facilitator:	Let's explore what happens if "All Montclairians are liars" is a false statement.
Victor:	If he's saying that "All Montclairians are liars" and it's not true, hmm, it means that he's lying.
Facilitator:	And what follows if "All Montclairians are liars" is false?
Victor:	Well, that would mean that not all Montclairians are liars, that some of them are not.
Facilitator:	Or, there is at least one Montclairian who isn't a liar. Could that be the speaker?
Bill:	Yes, but that would be the same circle as before, because he would be lying and not lying. But we aren't sure whether it's him.
Victor:	But what you guys got confused about is that it said " All Montclairians are liars," but does it say that all Montclairians *always* lie? Nooo, nooo. Montclairians may be liars, but they may not always lie. Not all the time. So he can't be telling the truth. It didn't say all the time.
Facilitator:	Let's see how we can interpret that. You remember when we were solving the problems with liars involved, and we were assuming that if someone was a liar, he or she was always lying. So, in this case let's see what will follow if we work with that assumption.
Chas:	O.K. If people are liars, it doesn't mean that they lie all the time. But in this case since it's such a simple problem, I think that the people would have suggested that if you're a liar, you always lie, like in the past problems that we had—there is a liar, a spy, and a truth teller, a liar always lies and the other tells always the truth…I think here they're trying to say that if you're a liar, you always lie, because they're *simple people*. So if he's saying that all Montclairians are liars, and if he's a Montclairian then if he's saying that they all are liars, then he's saying "Oh, I'm telling the truth." That means since he's a Montclairian,

he's a liar and then he can't be telling the truth if he *always* lies.

Facilitator: So, what's this?

Chas: What I'm saying is that if he's a Montclairian who always is lying, since…then if he's saying that he's saying the truth then he can't…since he's a Montclairian. Then he's basically saying "All Montclairians are liars. Oh, wait! I am Montclairian. So…."

Facilitator: So, If we conclude something, we'll say…

Chas: Because the people who make this, I think, since…last time they said "A liar always lies." So, I think they probably mean, a liar always lies.

Facilitator: So, if we assume that a liar always lies, in this case what happens?

Chas: I'm saying, since he's a liar, then what we say…. Well, basically, I agree with Nellie then.

Teacher: Is that disagreeing with Victor saying that just because liars… Let's not assume that he always lies, or that he's lying right now.

Chas: I mean… You might be right, but what I think what they're trying to say is that liars always lie.

Teacher: So, your statement then would be therefore liars always lie.

Chas: Since that's how they said in the past question….

Teacher: O.K. But Vicente disagreed saying "Just because he's a liar it doesn't mean that he lies now"? Yep?

Victor: Last time [in a previous problem] we said that the liar always lied, but it doesn't mean that this time this is the case.

Facilitator: So, we have this different definition for a liar: that a liar is someone who always lies and he might not always lie. Then what if we assume the first or the second definitions?

Darlene: When people say someone is a liar usually it tells you that they lie all the time and never tell the truth, because you wouldn't really call them a liar if they tell the truth. So I agree with Chas, that a liar always lies. They wouldn't be called liars if they told the truth sometimes, because they would be lying all the time.

Facilitator: So, you're taking the first definition that a liar always lies. If we go on with this definition then is the speaker lying or telling the truth?

Naomi: If we assume that liars always lie, then he's lying by saying that all Montclairians lie, but he would be also telling the truth if he had said that. If he says that "All Montclairians are liars" and he's telling the truth and he's a Montclairian then he wouldn't be lying.

Voice: No, he would.

Naomi: So, most likely he's lying by saying that… Because he can't tell the truth if he says that "All Montclairians are liars." Because he's a Montclairian, so it has to be that he's just lying.

Rush: Well, since he says that all Montclairians are liars and he's a Montclairian himself, if he says that… he's lying, but he's telling the truth.

Facilitator: So, what if we assume that liars always lie and if he's telling the truth he's lying as well. Is that possible? Is there another option? Can we think differently?

Samantha: It was already offered, but it's different. I agree with Vicente, because…it doesn't mean that he always lies, because if he always lies then it becomes impossible to answer the question. Because, if he said "All Montclairians are liars" and he's telling the truth, that means he's lying that all Montclairians are liars which means that he's telling the truth, so…. I agree with Vicente that liars don't always lie,…'cause if he does it wouldn't be possible to answer the question.

Facilitator: If we assume that a liar doesn't always lie?

Samantha: None of the others will work. Because, if he's telling the truth that all Montclairians are liars, since he's a Montclairian he would be a liar too. He would be telling the truth that would make him a liar and then the other way it goes back and forth too.

Facilitator: Isn't it possible that he's telling the truth at this particular moment assuming that other times he might be lying? Ben, what do you think?

Bill: I think it goes back and forth. I think that's the way this question is meant to go. How are we supposed to debate it if it keeps going on?

That's what I think the answer is: It keeps going on.

Victor: I just need a clarification. What makes you think that a liar always lies? Because I'm sure that you have called someone a liar before. And if you have called someone a liar before that doesn't mean that he always lies. Because they have to tell the truth [sometimes]. I don't think there is anybody that can possibly *always* be lying.

Facilitator: So, you're speaking for the second definition.

Victor: I'm just asking what makes you think that because a person is called a liar, he always has to lie.

Nellie: I'm saying that what makes us think that a liar always lies is because we're not talking exactly about real life, we're talking about the problem. The only reason we're saying that he's always lying is because—go back to the other problems with liars, we're bringing the liar back into this—because we don't know whether he always lies. We're working with what we have. So, we don't know and we don't have his life story or we can't predict the future.

Victor: I think we have to think like this is the first question we have ever done. Forget that in the other problem, the liar always lies, forget that. So?

Sally: I agree with Vicente, Because no one can lie all the time… 'cause, if that guy was saying everyone in Montclair is a liar, he would be telling the truth, if everyone is a liar in Montclair, but he would be from Montclair too. So, what am I trying to say?

Facilitator: Well, the question whether it's possible that he's lying and telling the truth at the same time is still open.

Sally: He really can't be telling the truth and lying at the same time. Because, I mean, if he's lying, well I'm sure he's lying, it's not that everyone in Montclair is a liar, but … Because he can say some people are liars in Montclair, then he would be telling the truth. So I don't think that would be possible.

Rush: I disagree with Sally this much. I agree with Nellie. This isn't a real life question.

Facilitator: Why not?

Rush: We don't have information. We don't know everybody in Montclair isn't a liar, we don't know that every person always lies.

Sally: No, no. Let's say everybody in Montclair is a liar. Then he would be saying "Everyone in

Montclair is a liar." He would be telling the truth, but he's lying. And he says "Everyone in Montclair is a liar," but there and then he would be telling a lie, but not the truth. It's kind of confusing [laughs]. But my point is that you can't tell a lie and tell the truth at the same time.

Facilitator: O.K. That's the conflict here. Because he seems to be lying and telling the truth at the same time. The question is whether it's possible? What can we make out of this?

Darlene: I disagree with Victor, because we don't have any proof that a liar doesn't lie all the time or that he does. So, you can't say that he doesn't lie all the time because we don't have proof and we don't know. I agree with Nellie. Ah….

Rose

Victor: O.K. I want to ask again what do we have here?

Rush: Eh?

Victor: I'm working with what I have here. So, if it was real life then I would be right, if it was not real life then it's a 50/50 chance [inaudible] because for the not real it could be either one.

Facilitator: And your idea is?

Victor: That liars don't always have to lie. I don't have any proof of who he is that I can say.

Facilitator: So, are you saying that he might be telling the truth this time?

Victor: Yes, liars don't always have to lie. Because if a liar always has to lie then I want to hear what person is a liar?

Jimmy: I agree with Victor. And everyone in Montclair is a liar [in that sense]. I am a liar, you're a liar, we're all liars… . The reason is because, I could say I want to give you 1000 bucks. I don't give you 1000 bucks and you tell me "You are a liar." Everyone has such moments when they're two years old and say "I can count to a million" and you can't, so you're a liar. Thinking like that everyone is eventually called a liar, so they eventually are simply liars.

Teacher: Are you saying that because there has been a chance that we told a lie at least once in our life that we all be considered liars?

Jimmy: We've all said a lie. I know for myself maybe not for the others, but I've told maybe a thousand lies in my entire life. I probably have told millions and there is no person who continuously lies and lies and lies. Because, if they say "I love your food," They're lying… And no one can live off lying, no one is… because they eventually can say something which can sound truthful. Because no one can always lie. Always lying can make people mad. They have to tell the truth at least once.

Nick: I think some people might always lie, and some people might tell the truth most of the time.

Facilitator: Last conclusion? We have to finish.

Nellie: We don't really know if someone [who is called a liar] always lies or if anybody always lies. Because we don't know everybody's life, so we don't know whether they always lie or just sometimes. But if they do always [lie] it creates a real problem and then…we can't answer the question.