



Paweł Walczak

ORCID: <https://orcid.org/0000-0002-8541-0396>

University of Zielona Gora, Poland; e-mail: [p.walczak@ifil.uz.zgora.pl](mailto:p.walczak@ifil.uz.zgora.pl)

## **Analysis of the State of Research on the Effectiveness of the Philosophical Inquiry Method**

<http://dx.doi.org/10.12775/PBE.2021.013>

### **Abstract**

The aim of the article is to present and analyze the state of research on the effectiveness of the method of philosophical inquiry with children. The method was developed by the American philosopher Matthew Lipman (1923–2010) in the 1970s and, over the last few decades, has found many supporters and enthusiasts around the world. The dynamic development of the method was related to the implementation of numerous evaluation projects of the *Philosophy for children* (P4C) program, the aim of which was to examine the effectiveness of this method in the context of the development of cognitive and social competences of students. The article analyzes reports and reviews of the most important research projects related to the P4C method, selected on the basis of the adopted methodology, scope, and subject of the research. The conducted analysis leads to the conclusion that the researches on the effectiveness of P4C done so far, confirming that there exists a relationship between the participation of students in classes conducted with the method of philosophical inquiry and the improvement of educational achievements of these students in the area of cognitive and social competence. As a result of the analysis, it was also found that the previous researches did not include the evaluation of the P4C method in the context of improving the competences of teachers working with this method. The potential positive importance of P4C for the development of teachers' communication and teaching skills is signalled by many theorists and practitioners dealing with philosophizing with children.

**Keywords:** philosophy for children, philosophical inquiry, Matthew Lipman, educational research.

## Introduction

The method of philosophical inquiry with children and youth was developed in the USA at the turn of the 1960s and 1970s. Creator of the method, Matthew Lipman (1923–2010), observed that the logical, critical, and creative thinking skills of most of his students were very low, which indicated that these competences were ignored in the earlier stages of education. He diagnosed that it was the result of the vision of education based on Piaget's concept of the gradual development of cognitive competence, according to which logical and critical thinking skills do not appear in a human before the age of 12; hence this stage of development was not at the center of pedagogical influence at early levels of education. The program developed by Lipman and his colleagues was based on the assumption that children from an early age have a natural tendency to problematize the experience, which makes them capable of conducting meaningful philosophical reflections (Lipman et al., 2008, pp. 83–85). The authors emphasize that the main goal of the method is to support children in the development of independent thinking, reasoning, and argumentation skills and to help improve their communication and social skills (Lipman et al., 2008, pp. 78–111). For Lipman, the evaluation of the program was very important, as evidenced by the fact that as early as in 1969 Lipman conducted his first experiments in two small groups of 20 students aged 11–12, with whom he conducted philosophical inquiries. The results of this trial showed that students in the inquiry group performed significantly better in logical thinking and reading skills (Lipman & Bierman, 1980).

In 1974, Lipman and his colleagues established a research institution at Montclair State College for the P4C program: the Institute for the Advancement of Philosophy for Children (IAPC). The purpose of the IAPC is primarily to develop and evaluate the program. The Institute has established a close working relationship with educational authorities (New Jersey Department of Education and the Educational Testing Service), which over time has resulted in the development of a special test examining skills related to the P4C program, the New Jersey Test of Reasoning Skills (NJTRS) (Slade, 1992, p. 26). In 1975, the IAPC conducted its first study in two schools in Montclair. Since then, several hundred studies and evaluation projects have been conducted worldwide to test the relationship between student participation in regular philosophical inquiry activities and improved school achievement.

The aim of this article is to present the current state of research on the effectiveness of the method of philosophical inquiry. Based on previous meta-analyses and reports from recent studies, the methodological assumptions of the described studies will be analysed, as well as their course and conclusions formulated in the presented papers. At the beginning of the 20th century, meta-analyses of research projects related to P4C evaluation were published: these are articles by Félix García-Moriyón, Irene Rebollo, Roberto Colom (2005) and Steven Tricky and Kieth Topping (2004). The authors of these studies collected and critically analysed data from past evaluations of the Philosophical Inquiry programme. Up-to-date information on research projects and evaluation of the P4C method were collected and published on the IAPC<sup>1</sup> and SAPERE (Society for the Advancement of Philosophical Enquiry and Reflection in Education)<sup>2</sup> websites. The most methodologically and substantively advanced research project on the effectiveness of the method of inquiry is the research of the team of researchers led by Professor Stephen Gorard of Durham University in cooperation with the British organisation SAPERE (Gorard et al., 2015). The presentation of the state of the research is a pretext for formulating questions about further prospects for the development of the research as well as the method itself and an attempt to answer the question about the place and importance of philosophical inquiry in contemporary education.

### **Analysis of methodological assumptions of selected research projects**

The research conducted so far on the effectiveness of the philosophical inquiry method has focused on two areas: cognitive competence and noncognitive competence. Cognitive competences are defined in this research as skills related to efficient and creative reasoning and, to some communication skills. The most frequently mentioned skills are logical and critical thinking, correct reasoning, formulating questions, correct argumentation, drawing conclusions, defining concepts, reading with understanding, expressing and specifying one's own position, reconstructing other people's views. Skills related to creativity and flexibility of thinking are often highlighted: finding new solutions, non-schematic thinking, or nonstandard applications, as well as the ability to self-assess one's own knowledge and self-awareness of learning.

---

<sup>1</sup> [www.montclair.edu/cehs/academics/centers-and-institutes/iapc](http://www.montclair.edu/cehs/academics/centers-and-institutes/iapc).

<sup>2</sup> [www.sapere.org.uk](http://www.sapere.org.uk).

The category of noncognitive competence includes social and emotional skills and often emphasises attitudes related to dialogue, cooperation, mutual respect, and tolerance. To date, there have been few studies focusing exclusively on noncognitive competence. Cebas and Moriyon, in their analysis, identified only two such studies (Cebas & Moriyon, 2015, p. 11). The largest study dedicated to this category of skills was that conducted by researchers at Durham University as part of the *Non-cognitive Impacts of Philosophy for Children* project carried out between 2014 and 2016 (Siddiqui et al., 2017).

Along with the publications announcing the outcome of research projects related to the effectiveness of the P4C programme, there were numerous critical voices that pointed to the problem of a proper definition of the subject of evaluation because what skills fall into the category of “philosophical”? Lipman himself emphasises above all logical, critical, and creative thinking skills, but also emotional, moral, and social competences. The study of the quality and size of the possible impact of the programme on the development of the student requires the specification of the qualities to be measured. This is because an adequate selection of research tools depends on it. The review of reports shows how the manner of defining the subject of research has changed under the influence of critical reflection on the methodology of evaluation. In older projects, the authors formulated research objectives and hypotheses using very general categories. In a study by Austrian researchers Daniela Camhy and Gunter Iberer carried out between 1984 and 1986, the aim of the research was formulated very generally: “the main task of the research (...) is to test the hypothesis that the proper teaching of philosophy, by stimulating the philosophical thinking of students, fosters their further mental and personal development” (Camhy & Iberer, 1990, p. 188). Researchers with the Clakmannashire Project (2001–2004) aimed to see if regular philosophical inquiry could lead to the development of cognitive abilities, the development of critical thinking skills, and the deepening of classroom dialogue, and the emotional and social development of students (Trickey, 2005, p. 2). A recent 2014 study by a team of Durham University researchers focused on several specific competencies as standard criteria for assessing a child’s level of development in schooling. The researchers stated the aim as “to investigate whether annual P4C classes can lead to better school achievement in mathematics, reading and writing” (Gorard et al., 2017, p. 3).

The P4C evaluation study consisted of measuring the impact of a pedagogical intervention on the development of specific skills and characteristics of the students. The pedagogical intervention, in this case, was the regular participation of the students under the study in the philosophical inquiry method classes

for a specified period of time. Two basic research models were used to investigate the effectiveness of the programme: independent groups and dual measurement. The model based on independent groups consists of conducting the study in two groups, one of which participates in the classes (experimental group) and the other follows the usual school programme without P4C classes (control group). The differences in measurement results between the experimental group and the control group make it possible to estimate the size of the impact, i.e., the numerically expressed effectiveness of the programme. In dual-measurement studies, tests are conducted on a group of programme participants before the intervention (pre-test) and after the completion of the class cycle (posttest). In this example, the size of the impact is estimated based on the differences between the measurements. The procedure based on independent groups with a post-intervention test (post-test) does not give a fully reliable result due to the fact that the difference between the control group and the experimental group before the intervention is unknown here. In a study based on a double measurement in one group, without reference to a control group, we cannot answer the question whether the result obtained is related to the intervention or to other factors. Therefore, the most reliable results are obtained from studies whose methodology combines these two models. In some projects, an extended procedure is used, which involves the use of intermediate measurements or measurements examining the maintenance of the intervention effect over time (follow-up test). An example of such an extended study is the Clamannashire Project. In addition to pre- and post-intervention measurements, the same groups were surveyed 16 months after the end of the intervention to see if, and to what extent, the measured effect persists after some time (Topping & Trickey, 2007a, p. 787).

Most participating students took part in the study for one school year at a rate of one hour per week. In the study analysed by Professor Moriyon's team, the average length of the intervention was 7.33 months (Garcia-Moriyon et al., 2005, p. 8). In this respect, the most interesting project is the only long-term study so far carried out in Madrid by the European School of Madrid, where researchers are trying to answer the question whether it is possible to sustain long-term positive effects in the development of cognitive and emotional competences in students participating in philosophical inquiry (Colom et al., 2014, pp. 50–56). Regular weekly P4C classes of 1 hour per week have been run in this school since 2002. These classes are integrated into the school curriculum at all levels of education from the first grade of elementary school (6 years) to the last grade of secondary school (18 years), each student going through a 12-year cycle of P4C classes. The study is planned for 20 years. During the

12 years of the experiment, 3 intermediate measurements were carried out, from which preliminary results were developed in 2014 when the first group completed the 12-year cycle of study.

It is worth noting that the evaluation research concerns the method of philosophical inquiry in its various variants, which arose as the method was popularised. The P4C methodology is based on two pillars: the structure of the classes and the idea of a community of inquiry. The differences concern mainly the use of different materials as a stimulus for discussion, while the basic structure of the classes and the principles of turning the class into an inquiry community remain the same.

In 1983, researchers at the IAPC in collaboration with the Educational Testing Service of New Jersey developed a testing tool, the New Jersey Test of Reasoning Skills (NJTRS), a test that examines skills in those competencies that are specifically shaped during philosophical inquiry. The NJTRS test is a choice test, which involves indicating one answer to a question from several proposed options. This test targets different aspects of logical thinking: finding the right premise and equivalent thinking by analogy, syllogistic inference, relations of containment and disjunction, etc. (Camhy & Iberer, 1990, p. 201).

In studies conducted in the UK, standardised tests for the evaluation of cognitive abilities and logical skills (CAT) and the *Myself as Learner (MALS)* test were used as the basis for measurement. The *Cognitive Abilities Test (CAT)* is a standardised test used in British education to assess the development of a student's cognitive skills. It measures the level of development in terms of verbal, non-verbal, and mathematical skills. The test assesses skills such as reasoning, identification, matching, determining connections, classification, recognition, memorization, and analysis. The MALS test is used to evaluate students' perceptions of themselves as learners and active problem solvers in the learning process. The test consists of a series of statements describing the learner's characteristics, and respondents must indicate on a five-point scale how much the statement applies to themselves (Trickey & Topping, 2006, pp. 6–7).

As reflections on the methodology of the social sciences developed, other research techniques and tools emerged. Qualitative methods such as observation, interviews, questionnaires, and surveys, as well as psychological and personality tests, were increasingly incorporated into research. Already in the early 1980s, the Camhy and Iberer project planned to use the observation of conducted classes using video recording analysis. Unfortunately, this was abandoned at the time due to the very poor quality of the recordings made (Camhy & Iberer, 1990, pp. 190–191). This type of tool was successfully applied in

2001 by the Clamnnashire Project. The video recording analysis was based on six recordings of group activities (four experimental groups and two control groups) before the introduction of the P4C classes and six recordings after the six-month period of teaching. In November 2001, classes were conducted in all six groups on the basis of a common scenario, where teachers were expected to follow strictly the instructions in the class description. The *Thinking through Philosophy* activity programme was then introduced to the experimental groups. After six months, in both the experimental and control groups, classes were conducted and recorded according to the same scenario. The analysis of these records consisted of assessing certain observable and measurable behaviours of students and teachers according to a specific observation schedule. The schedule detailed behaviours such as students' spontaneous expression of their own opinions, students' expression of agreement or disagreement with the opinions and arguments of other students or the teacher, the total time of students' and teacher's own statements, the number of questions asked, etc. (Topping & Trickey, 2007b, pp. 73–84).

The reliability of this type of research is largely dependent on the size of the sample. Most research projects on philosophical inquiry have unfortunately been conducted with a small sample of students. The first large study, and the largest to date, was designed and implemented by IAPAC with the cooperation of the Educational Testing Service (ETS) in 1981. A total of 4,500 students from various schools in the state of New Jersey participated in the year-long experiment. Durham University researchers surveyed 3159 students from 48 schools in England (2013–14) and 2722 students from 42 schools in 2015–16. One hundred eighty-six pupils took part in the Clamannashier Project. In summary, the evaluations to date have included groups ranging from a few dozen pupils to several thousand. The average age of participating students is 11–12 years (García-Moriyón et al., 2005, p. 8).

## **Review of the results of selected research projects**

In the early 2000s, Prof. Felix Garcia-Moriyón and his team from the University of Madrid undertook a data collection of existing studies on the effectiveness of P4C. The research was based on published reports, articles, and unpublished data. Their search resulted in a list, published in 2004, of over 100 studies conducted in different parts of the world (Cebas et al., 2015). The authors concluded that despite such a large number of studies conducted so far, their results can still be questionable and the question whether (and possibly how) the P4C

programme works is still difficult to answer. One of the more common objections is to highlight the fact that most of the projects were carried out by supporters and promoters of the programme, and the studies were meant to confirm and convince of the effectiveness of P4C. Hence, the results of these studies are seen as biased rather than the result of a critical examination of the real impact of the programme. Furthermore, the authors highlight the wide variety of approaches to study design and the varying methodological level. Most of the study reports do not provide full information about the methodology of the measurements carried out, which leads one to believe that they were conducted without following the basic standards of scientific research. The reason for this approach was probably the fact that many of the researchers who undertook the evaluation of the P4C programme were teachers, enthusiasts, or academics who do not deal with social research daily; hence the research was often of a semi-amateur character.

In 2005, the authors published an article presenting the results of their analysis (García-Moriyón et al., 2005). The first stage of the meta-analysis by researchers from Madrid consisted in selecting and rejecting those reports that did not meet several basic criteria. It was assumed that the studies should be designed in such a way that their explicit object was to investigate the impact of the P4C programme on the development of cognitive competence. Due to the very small number of studies on noncognitive competences, the authors decided to analyse only studies measuring skills related to reasoning. Furthermore, studies that did not contain complete data on measurement methodology were discarded. After the selection, it turned out that only 18 research reports met these requirements, and the results of these studies were analysed.

The analysis consisted of estimating, based on the compiled results, the average impact size from the studies. In addition, the authors tried to identify factors that could influence the results of the measurement and identified three such factors: the type of research tool, the research model, and the time of publication of the report. All studies by type of research tool were divided into two categories: studies using NJTRS and others. The analysis showed that studies using the NJTRS gave better results than studies where other tests were used. The NJTRS, as a test dedicated to the evaluation of the P4C programme, contains tasks that are too closely correlated with the content and methodology of the programme, so it may give a biased picture of the real impact. Significant for the results of the measurement was also the model of research conducted. Studies combining the independent group approach and dual measurement showed a smaller effect than studies without a control group or with a single measure-



ment. The year of publication, in turn, was significant because of the marked change in the approach to research methodology that occurred over time, particularly in the 1990s. A comparison of reports shows that the later the study, the smaller the effect size, while the higher the methodological level of the study.

This meta-analysis resulted in an average effect size of 0.5848. The authors illustrate this result by relating it to the intelligence quotient: if we assume that the IQ of the control group is 100, then the IQ of the experimental group is 107. If in the control group, 9% are in the range above 120 IQ, then in the experimental group, 20% will be in this range. This shows that the study provides evidence that participation in P4C activities has a positive impact on the development of cognitive competence in children. On the other hand, the analysis showed that the way in which the study is conducted influences the outcome. Nevertheless, even the lowest results are statistically significant and show the positive impact of P4C.

In 2004, British researchers Steven Trickey and Keith Topping published an article (Trickey & Topping, 2004) in which they reviewed existing research on the effectiveness of the P4C programme. They assumed that studies using pre- and post-intervention testing in experimental and control groups are the most reliable in measuring the impact of the programme on skill development, as free as possible from the influence of confounders interfering with this measurement. Of the ten studies selected, two reports did not contain sufficient data to estimate the magnitude of impact. The studies analysed by Trickey and Topping were conducted over a 30-year period, between 1970 and 2002, in the USA, Canada, and the UK. They estimated an average impact size of 0.425 with a standard deviation of 0.09 (Trickey & Topping, 2004, p. 375).

The authors reflect on the sources of fear, resistance and some distance to the P4C programme among teachers and educational authorities. According to them, the biggest barrier may be the deeply entrenched traditional teaching culture. They point out that 'teachers' interest in achieving specific changes in pupils' behaviour and their perception of their role as behaviour managers seems to be in conflict with listening to and following what children are saying' (Trickey & Topping, 2004). An important conclusion of the analysis is the authors' indications for future research. They emphasise the need to improve methodological standards in the evaluation studies of the P4C programme and to include more pupils from many schools in the study, and to operate the programme over a longer period of time (Trickey & Topping, 2004, p. 377).

Ten years after the publication of the analyses by researchers from the UK and Spain, the results of a study carried out by a team of researchers led by Pro-

fessor Stephen Gorard from Durham University in collaboration with the UK organisation SAPERE were announced (Gorard et al., 2015). The project, carried out in 2013–2014, aimed to test whether one year of regular P4C instruction could lead, in pupils at Year4 (8–9 years) and Year5 (9–10 years) learning levels, to improve achievement in mathematical skills, reading and writing, and in cognitive skills as measured by the *Cognitive Abilities Test* (CAT).

Over 3000 pupils from 48 schools across England took part in the study (22 schools formed the experimental group, and 26 schools formed the control group). For 12 months, 1 hour per week of regular P4C activities was conducted in the selected schools according to the SAPERE programme, the structure of which is modelled on Lipmann's programme. The research methodology was based on both quantitative and qualitative tools. For quantitative measurements, a standard Key Stage 2 test was used to determine students' progress in mathematical skills, reading, and writing. In addition, the participating students were surveyed using the CAT test in two measurements: a pre-test before the start of P4C classes (December 2012) and a post-test after the end of classes (January 2014).

The qualitative research consisted of interviews with pupils and teachers and observations made by observers according to a predesigned questionnaire. The researchers were interested in changes in pupils' behaviour (interaction between pupils, pupils' relation to the teacher and to the school, activities in lessons) and in teachers' behaviour (encouraging pupils to discuss, dominating discussions, listening to pupils, asking questions, etc.). Observers conducted surveys at the beginning and end of the project in selected schools.

The results of the study provide evidence of the positive impact of P4C on pupils' improved achievement in reading and writing. Key Stage 2 tests showed that pupils in the experimental group were ahead of pupils in the control group in their progress in reading and math by an average of 2 months. However, the tests did not show greater progress for pupils in the experimental groups in writing.

In the CAT study, the skills studied are captured in 4 categories: verbal, non-verbal, quantitative, and spatial thinking skills. The experiment revealed a positive impact on verbal skills, with no significant positive changes in the other categories. Students who started the programme at Year 5 level (9–10 years old) showed a positive effect of P4C, but younger students who started the programme at Year 4 level had no better results than the control group, which shows the lack of benefit of the programme in this group.

In the qualitative research, it was shown that the overall positive effects of P4C activities are significantly influenced by the way in which these activi-

ties are integrated into the timetable according to which the school works. For the programme to be more effective, it is important for these activities to be a regular part of the weekly timetable, rather than an irregular activity added to the regular timetable. Surveys and observations have shown that P4C has had a beneficial effect on improving many of the pupils' social and emotional skills, such as increased pupil attention, more frequent self-disclosure, greater patience when listening to others, and balanced self-esteem. Some teachers also highlighted that P4C had a positive impact on students' overall engagement in class and a tendency to ask more questions in other lessons as well.

## **Summary**

The analysis of previous research projects on the programme of philosophical inquiry with children and young people leads to the conclusion that the unequivocally positive assessment of the relationship between students' participation in P4C classes and the development of their competences may be questioned due to the insufficient substantive and methodological level of most evaluation projects. The authors of the above presented reviews and meta-analyses have pointed out the problems with maintaining methodological standards in many studies. On the other hand, it is worth emphasising that the meta-analyses cited in this article was based on selected studies with no serious reservations, and their results led to positive conclusions. Moreover, the 2015 study by Prof. S. Gorard's team from Durham University was conducted in compliance with all methodological standards required for this type of research, which allows us to consider its results as reliable. The results of Prof. S. Gorard's study are generally consistent with the conclusions of the meta-analyses discussed here. Taking into account any criticism of the methodology of older studies, however, it seems reasonable to conclude that the evaluations so far indicate that there is a relationship between participation in regular classes conducted using the P4C method and better academic performance, which in turn seems to confirm the positive significance of philosophical inquiry in education. The above analysis also shows that there is still a need for extended professional evaluation research that would provide the data necessary to improve the method and develop an effective application of P4C in the education system.

Research to date has focused on student achievement. The interest of the researchers has been to show the relationship between participation in classes with the philosophical inquiry method and achievement in cognitive, emotional, and social skills. However, other participants in this process – teachers – have

not been studied yet. An interesting question is whether conducting classes using this method may be a factor shaping the new communicative and didactic skills of the teacher, reorienting his/her attitudes and beliefs, and thus indirectly influencing the process of transformation in the broadly understood school culture? As I mentioned above, one of the pillars of the method is the rule of turning the class into an inquiring community. This process involves a complete reversal of the traditional communication paradigm in the school classroom and a change in the role of the teacher. In a community of inquiry, the teacher, together with the students, make a joint effort of inquiry with the aim of finding an answer to a question. The teacher becomes an equal member of the community, and his voice is in no way privileged over the pupils. This organisation of communication in the classroom conflicts with the traditional model of teaching based on the formal authority of the teacher. Turning the classroom into an inquiring community requires the teacher to change his or her view of his or her role in the classroom and enter into a different kind of relationship with the children. What is at stake is an authentic community of children, the aim of which is a joint search for answers to children's questions. As Anna Łagodzka rightly notes, "the model of inquiry (...) has been designed in such a way that students not only undertake a philosophical reflection and communicate it, but that children explore the questions and philosophical insights created by children, rather than thinking about the assertions or questions suggested by the teacher. These are dialogues centred on issues raised by pupils, their questions, and their answers, conducted in the forum that the school classroom becomes" (Łagodzka, 2014, p. 108). A teacher's practice of collaborative philosophical inquiry in the school classroom can influence his or her perception of school reality and the way he or she functions in other subjects as well. Investigating the nature and scale of this influence should become a subject of reflection and research for future evaluation projects.

## References

- Camhy, D., & Iberer, G. (1990). Nauczanie „Filozofii dla dzieci”. Badania nad umysłowym i osobowościowym rozwojem uczniów [Teaching 'Philosophy for Children'. Research on the Mental and Personality Student's Development]. *Prakseologia*, 1–2 (106–107), 183–212.
- Cebas, E., & Garcia-Moriyón, F. (2015). What we Know About Research in Philosophy for Children? Retrieved 15 January 2019 from <https://philoenfant.org/2015/10/30/resume-de-103-recherches-en-philosophie-pour-les-enfants/>.

- Coleghorn, P. (2002). *Thinking Through Philosophy*. Blackburn: Educational Printing Services.
- Colom, R., Garcia-Moriyón, F., Magro, C., & Morilla, E. (2014). The Long-Term Impact of Philosophy for Children: A Longitudinal Study (Preliminary Results). *Analytic Teaching And Philosophical Praxis*, 35, 1, 50–56.
- García-Moriyón, F., Rebollo, I., & Colom, R. (2005). Evaluating Philosophy with Children: A Meta-Analysis. *Thinking. The Journal of Philosophy for Children*, 17(4), 14–22, doi: 10.5840/thinking20051743.
- Gorard, S., Siddiqui, N., & Huat See, B. (2015). *Philosophy for Children. Evaluation Report and Executive Summary*. Durham: School of Education, Durham University. Retrieved 15 January 2019 =[https://v1.educationendowmentfoundation.org.uk/uploads/pdf/Philosophy\\_for\\_Children.pdf](https://v1.educationendowmentfoundation.org.uk/uploads/pdf/Philosophy_for_Children.pdf).
- Juuso, H. (2007). *Child, Philosophy and Education. Discussing the Intellectual Sources of Philosophy for Children*. Oulu: Oulun Yliopisto, Acta Universitatis Ouluensis E 91.
- Lipman, M., & Bierman, J. (1980). Appendix B: Experimental Research in Philosophy for Children. In: M. Lipman, A.M. Sharp, & F. Oscanyon (Eds.), *Philosophy in the Classroom* (pp. 219–224). Philadelphia: Temple University Press.
- Lipman, M., Sharp, A.M., & Oscanyon, F. (Eds.) (2008). *Philosophy in the Classroom*. Philadelphia: Temple University Press.
- Łagodzka, A. (2014). Dyskusja dialogiczna: filozofia dzieciństwa i filozofia dorosłych [Dialogic Discussion: Philosophy of Childhood and Philosophy of Adult]. *Analiza i Egzystencja*, 25, 99–122.
- Pobojewska, A. (2019). *Edukacja do samodzielności. Warsztaty z dociekań filozoficznych. Teoria i metodyka* [Education for Self-Reliance. Workshop on Philosophical Inquiries. Theory and Methodology]. Łódź: Wydawnictwo UŁ.
- Siddiqui, N., Gorard, S., & Huat See, B. (2017). *Non-Cognitive Impacts of Philosophy for Children. Project Report*. Durham: School of Education, Durham University. Retrieved 15 January 2019 from <http://dro.dur.ac.uk/20880/1/20880.pdf?DDD34+DDD29+czwc58+d700tmt>.
- Slade, Ch. (1992). Creative and Critical Thinking. An Evaluation of Philosophy for Children. *Analytic Teaching*, 13(1), 25–36.
- Topping, K.J., & Trickey, S. (2007a). Collaborative Philosophical Enquiry for Schoolchildren: Cognitive Gains at 2-Year Follow-Up. *British Journal of Educational Psychology*, 77, 787–796, doi: 10.1348/000709907X193032.
- Topping, K.J., & Trickey, S. (2007b). Impact of Philosophical Enquiry on School Students'

Interactive Behavior. *Thinking Skills and Creativity*, 2, 73–84, doi: DOI:10.1016/J.TSC.2007.03.001.

Trickey, S., & Topping, K.J. (2006). Collaborative Philosophical Enquiry for School Children. Socio-Emotional Effects at Eleven to Twelve Years. *School Psychology International*, 27, 1–16; doi: 10.1177/0143034306073417.

Trickey, S., & Topping, K. J. (2004). “Philosophy for Children”: A Systematic Review. *Research Papers in Education*, 75(3), 365–380, doi: 10.1080/0267152042000248016.

Trickey, S. (2005). *Promoting Social and Cognitive Development Through Collaborative Enquiry: An Evaluation of the “Thinking Through Philosophy” Programme*. Retrieved 15 January 2019 from <http://www.clacksweb.org.uk/document/228.pdf>.