Ricerche di Pedagogia e Didattica – Journal of Theories and Research in Education 15, 2 (2020). ISSN 1970-2221.

Computational Thinking, Spatial and Logical Skills.
An investigation at primary school

Marialaura Moschella
Free University of Bozen-Bolzano

Demis Basso
Free University of Bozen-Bolzano
Universidad Católica del Maule, Chile

Abstract

The pedagogical outcomes of the implementation of Computational Thinking (CT) have been recently defined, while different approaches into the didactic dimensions have been proposed. From an epistemological point of view, some psychological aspects and their relationship with the general construct deserve further investigation. The present study aimed to evaluate the relationship between logical thinking, divergent thinking and spatial skills with CT. Elithorn test, Raven's Progressive Matrices, Creative Thinking test were evaluated on students from 3rd to 5th grade, a sample of 51 pupils for the first training, and a sample of 43 for the second one. The tests were administered before and after the coding activities. Teachers and researchers collaborated on the planning and implementation of the training. The activities consisted of eight workshops, based on several CT-related processes (planning, problem-solving tasks using building bricks, storyboards, coding with LegoWeDo and Scratch). Qualitative data were collected in the form of a focus group and an interview to investigate teachers' perception at the end of the experiment. Results showed a relevant relationship of CT with spatial planning, and a moderate one with logical thinking, while no impact was due to creative thinking. These results further elucidate the recent CT theoretical framework, in which spatial skills play a meaningful role.

L'implementazione del Pensiero Computazionale (PC) sin dalla scuola primaria propone un cambio di paradigma nei diversi approcci pedagogico-didattici. Da un punto di vista epistemologico, alcuni aspetti psicologici e il loro rapporto con il costrutto generale meritano un approfondimento. Il presente studio propone di valutare il rapporto tra pensiero logico, pensiero divergente e competenze spaziali con il PC. Il test di Elithorn, le matrici progressive di Raven, il test del Pensiero Creativo sono stati proposti ad alunni (N=94) frequentanti le classi terze, quarte e quinte della scuola primaria durante due anni scolastici. I test sono stati somministrati prima e dopo le attività di programmazione. Insegnanti e ricercatori hanno collaborato alla progettazione e alla

DOI: https://doi.org/10.6092/issn.1970-2221/11583

realizzazione dell'intervento. Le attività comprendevano quindici ore di workshop, basati su diversi processi relativi al PC (pianificazione, problem-solving con Lego, storyboard, codifica con LegoWeDo e Scratch). Un focus group e un'intervista condotte con le insegnanti hanno completato la rilevazione dal punto di vista qualitativo. I risultati hanno mostrato una correlazione significativa tra le attività di coding con le abilità visuo-spaziali; una moderata correlazione con il pensiero logico; mentre nessun impatto è stato dovuto al pensiero creativo. Questi risultati chiariscono ulteriormente il recente quadro teorico del PC, in cui le competenze spaziali giocano un ruolo significativo.

Keywords: computational thinking; cognitive skills; computer science education; K-12; teacher training

Parole-chiave: pensiero computazionale; processi cognitivi; didattica dell'informatica; scuola primaria; formazione insegnanti

1. Theoretical background

Computational Thinking (CT) is worldwide indicated as a set of essential skills for the digital age. The vast amount of definitions published in the scientific literature (Lowe et al., 2019) is an indicator of the presence of several approaches to the construct. For instance, the didactic approach (Guzdial, 2015), the evaluation issues (Kong & Abelson, 2019), the epistemological and ontological aspects on which the debate is still open (Guzdial, 2015; Tedre & Denning, 2016). However, the psychological implications (Pea & Kurland, 1984; Mendelsohn et al., 1990) of introducing CT already at the kindergarten level were not developed yet as done in the previous areas and require further investigation.

The need for creating digital culture was claimed by Papert (1993) when it was just hypothesized that every family was about to dispose of a personal computer. The constructivist theory of learning supported the dimension of being "active user" and having an "object-to-think-with". According to the constructivism hypotheses, children could achieve a different level of abilities at a different stage during their growth. Although relevant, the constructivist hypotheses were criticised, given that such a sequence of stages does not correspond to actual cognitive development (Siegler, 2016). On the other hand, the approaches of Fröbel (1826/1903) and Montessori (1995) insist on the meaningful role of processes rather than on the result. These approaches are thought to have highlighted the relevance of developing autonomy, body experience and tinkering, and in turn, fostered research about their application in the field of Computer Science (CS) education (Resnick & Robinson, 2017; Eriksson, 2019). Notwithstanding, the need to appropriately develop digital literacy and CT skills is urgent while it was proposed that learning in the digital age should be based on three essential skills, among others: creativity, flexibility and learning to learn (Trilling & Fadel, 2009). The abstraction ability to process information along with creativity and flexibility is particularly notable, given that it allows being able to transfer procedures and algorithms from a problem to another, from field to field. The skills, as mentioned above, might be considered as the most interdisciplinary ones. Independently from the field, it might be possible to claim that they are advisable for life and work in the digital age, as postulated by Wing (2008). For its intrinsic characteristics, learning to think computationally could be essential but also feasible for all, respecting the principles of inclusion.

Recently, researchers (Guzdial et al., 2019; Nardelli, 2019) suggested revising the approach to computing education: it is argued that it would be better to let everyone be a "good thinker". Notwithstanding, it was also mentioned that CT skills cannot be taught since CT skills can be several and very different ones. Design elements and tools should be implemented and improved so that students could better understand and think. This way, computing education could be considered "as a discipline of transversal value" (Nardelli, 2019). The arguments about the disciplinary aspect of Digital Literacy (what to teach - what to learn) are as ambitious as the didactic aspects (how to teach). Some of these dimensions have been recently considered, in regard to:

1. the STEAM education, for which the tinkering, or "lifelong kindergarten" approach seems to be successful (Resnick & Robinson, 2017);
2. the evaluation of the appropriateness of the agile methodology and its implications for lower grades and not just for CS-related topics.

DOI: https://doi.org/10.6092/issn.1970-2221/11583

3. whether teachers without previous experience in the CT field can integrate and implement CS principles effectively and efficiently.

Even though the theoretical debate has not produced definitive results, the application of CT in education still generates much significant research. Several studies on the inclusion of concepts such as CT and programming in K-12 education had a substantial impact on the school curricula in many countries (Szabo et al., 2019). Institutional curricula have been developed and implemented in parallel to an increasing interest in the application of CT in informal contexts, like museums and libraries (Eriksson, 2019). The exposure to CS principles since early childhood has generally shown a significant impact on related mathematical skills (Baytak & Land, 2011). Besides, it was shown that students might learn specific mathematical concepts while they develop coding abilities (Sengupta et al., 2013), probably because CS and mathematics share similar concepts (e.g., variables, loops, conditions). Recently, studies focused on the correlation between coding and problem-solving skills (Çiftci & Bildiren, 2019) as well as visuospatial reasoning and different aspects of numeracy (Tsarava et al., 2019).

Besides cognition, CT may take effects on other human abilities. Ciancarini et al. (2019) showed the effects of pair programming on the soft skills in the high schools and hypothesized the approach of "Cooperative Thinking" as very effective on both social and disciplinary achievements. The trend of including parents and families has been recently documented in the literature, with positive effects on the learning performances (Sheehan et al., 2019). Basso et al. (2018) proposed a method to evaluate CT by integrating observation on the coding performance with a) learning and b) social and cognitive skills. They argued that the simple evaluation of the coding performance could not be considered sufficient to estimate CT, and other skills (problem solving, motivation and attention, among others) play a relevant role in the evaluation of scholar children performance. However, their proposal still needs to be corroborated with data and was not linked to a concrete example of teaching intervention.

Teacher training is an argument of great interest, but the bridge between CT and education was beyond the aim of many proposals, as Guzdial (2019) has clearly shown. In his work, he remarked how focusing on teachers and their methods is a crucial goal in the research on CT, which has been so far underestimated: «If we want better thinking and problem-solving, we have to improve the computing and use that to change our teaching» (p. 28). The relevance of training is high while referring to the initial grades, in which the preparation and training process do not foresee specific courses on CS. When present, programs are generally not planned with teachers, and teachers are mainly asked to execute tasks, regardless of their attitude and experience (Armoni, 2018). The consequences are that these interventions, without having a deep-enough disciplinary knowledge, often produce low self-confidence and motivation in teachers, even though they are often motivated and willing to introduce the subject. Among the few studies that tried to bridge the gap, Armoni (2018) posed some questions about this topic and proposed some suggestions, whereas the work of Falkner and colleagues (2019) described the alignments between intended and enacted curricula, suggesting relevant reflection on the field.

The current state of research in the field depicts an unbalanced figure: while many papers have discussed the theoretical basis of CT, the data collected to verify the theories raise some concerns. On the one hand, data were collected following stringent protocols based on computer science; only a few accounted for the contribution of learning and cognitive factors; on the other hand, many training sessions involved trainers as executors instead of co-designers.

DOI: https://doi.org/10.6092/issn.1970-2221/11583

The present study aims to collect further evidence filling these two methodological gaps. The effects of a coding course on the non-verbal cognitive abilities and logical reasoning of pupils are examined while evaluating a training created in collaboration with teachers at their first experience with coding activities. The focus was mainly on the psychological aspects underlying Computational Thinking Education (CTE) as a way to explore the generalization of learning onto three cognitive skills. Two explorative research questions were defined: 1) whether a correlation between CT activities, logical and divergent thinking, as well as spatial skills, may emerge due to the training; 2) whether the didactic strategies of inserting CT in the second language classes, rather than on mathematics can be meaningful. The choice of embedding coding in training into a different subject than mathematics was necessary to evaluate training feasibility and its adaptation to teaching practice, which more often than mathematics is present among the everyday lessons.
The following sections are organized as follows: research methods (participants, materials, data collection and analysis) are described in Section 2, Section 3 contains the results, which are discussed in the fourth sections.

2. Methods

The study was conducted according to a mixed-method design, in which qualitative and quantitative tools are employed. It was based on the Embedded design (Creswell & Plano Clark, 2007) that can consist of one or two phases (according to the data collection time). Data collection took place for ten weeks for a total of 13-15 hours per each experimental group. This project embedded qualitative data collection (interview, observation, focus group) with quantitative data (pre- and post-test). Together with a pre-test/post-test model for the experimental group, the data collection included a focus group (for the first training) and an interview (in the second training) with teachers. The Focus Group (Morgan, 1996) was used as a method for gathering qualitative data after the completion of the training. This tool, combined in a mixed-methods design, could contribute to a deep understanding of several aspects of the study, which could not emerge if based on quantitative tests only.

2.1 Participants

Given that the school director office already defined the classes, the sample was not randomly assigned, and it was composed as follows: In the first intervention a total of 51 children (26 girls), were split into three classes, attending the third, fourth and fifth grade of three primary schools, placed in a rural area of northern Italy, in the school year 2018/19. The average age of the children in the experiment group was 8.45 (SD=0.65). Also, their teachers (four) took part in the research: two of them were second language teachers and were supported by two class teachers. Their ages were 28, 36, 38 and 42 with respectively 4, 6, 10 and 15 years of teaching experience. None of the participants had any neurological, psychological or behavioural problems. In the second intervention, the sample was composed of 43 pupils, divided into two classes and two teachers. Their ages were 36 and 52, with respectively 10 and 25 years of experience. Teachers, children and their legal caretakers provided either verbal or written consent for participation. The experimental procedure adheres to the Declaration of Helsinki and has been checked and approved by the principal of the school.

DOI: https://doi.org/10.6092/issn.1970-2221/11583

2.2 Materials

2.2.1 Classroom Intervention

Researchers and teachers carried out the intervention, which was divided into three main phases. In the first phase, children chose stories that could be adapted and transformed using coding activities; they tried to develop algorithms, managing construction of bricks and blocks, through trials and errors strategies.

The second phase included the training using coding tools, which was divided into the following sub-phases:

1. The introduction to the new topic;
2. Free discovery of the programming language;
3. Tinkering and "debugging" (that is, self-evaluation and self-correction of the products).

Afterwards, creative activities were devised, in which children had the chance to think about a storyboard and to build the sprite/characteristics they need for the story (this phase foresees the use of Artec Blocks and Lego WeDo for the first implementation; the essential Scratch (Resnick et al., 2009) features for the second one).

The third part involved code writing, in order to program the robots and make the objects or characters "alive". In this session, children experienced the fundamental principles of CT, such as designing and managing procedures, writing algorithms, reformulating problems, restructuring processes, solving problems. In the end, a specific time was proposed to summarize and reflect upon the process. This part allowed children to observe the evolution of thoughts and behavior, and to stimulate creativity through reiterate exchanges between class members. Contextual observations and remarks were carried out throughout the training implementation.

2.2.2 The tests

The Creative thinking test. Among the several tests to study creative thinking, the present study has adopted the one developed by Williams (1994). It is based on the description of four cognitive factors of divergent thinking: fluency, flexibility, originality, and elaboration. Fluency can be defined as the generation of a wide number of ideas and meaningful responses; flexibility is linked to changing ideas passing from one category to another; originality is defined as the capacity to produce rare and infrequent ideas; elaboration is considered as the capacity to develop, embellish, and enrich ideas with details. The Italian version of the Test of Creative Thinking is a protocol with 12 frames, containing incomplete graphic stimuli. Children are asked to finish drawing in the bow and to write a title, in order to complete each frame according to their idea. The "fluency" score is attributed according to the total number of meaningful pictures (range: 1-12 points). The "flexibility" score is computed with one point every time in which one frame is different from one category to a different one (range: 1-11 points). The "originality" score is given by the total number of drawings inside or outside each stimulus (range: 1-36 points): one point is assigned to drawings outside stimuli, two points if it was inside the stimuli, and three points to each drawing made both inside and outside stimuli. The "elaboration" point is gained by drawing asymmetric pictures (range: 1-36 points): zero points are assigned to perfectly symmetrical pictures, one point to asymmetric pictures drawn outside stimuli, two points to asymmetric pictures inside the stimuli, and three points to asymmetric drawings both inside and outside stimuli (ibidem).

DOI: https://doi.org/10.6092/issn.1970-2221/11583

The Raven (2000) Standard Progressive Matrices. In the Raven Standard Progressive Matrices (SPM) each item presents a pattern made of 8 parts out of nine, organized in a 3 by 3 schema. Since one piece has been previously excluded, the gap should be filled by the examinee choosing the appropriate piece from a set of six different possibilities (just one answer is correct). This test aims at determining reasoning and analytical capability with a non-verbal test so that any cultural or age-related difference is limited (Giles, 1964). In our experiment, the brief version was administered, consisting of 12 tables.

The Elithorn Perceptual Maze Test. The test has been firstly developed for investigations on brain deterioration or brain injuries. Later on, this test was considered for the assessment of the Executive Functions, in particular related to spatial and planning skills. During the test, the instructor invites the student to draw on the example figures a path from his side of the sheet to the opposite side. He is then given the 3 rules, practically exemplifying them on the figure:

(1) at each intersection, it is possible to go to the right or left, but not to go back;
(2) it is not allowed cutting the mesh of the net;
(3) the path to be found is the one that crosses as many black points as possible, which is different for each track, and it is indicated below the track (as well as read aloud by the instructor).

The first item is done by the instructor commenting on the task and repeating the rules. The following items are made by the students, helping him and correcting him in case of difficulty, commenting on the errors. If the task has been understood, they move on to the real test, reminding the student that it is important to perform the task as quickly as possible. The maximum time allowed for each track is 2 minutes (Davies & Davies, 1965). The score consists of the total number of correct paths.

2.3 Procedure

The first training was planned in the first semester of school, from October 2018 to January 2019, over ten weeks. The researchers provided students with several sets of Artec Blocks (School Mathematics set) as well as the sets of Lego Wedo. Lego Wedo was also the software installed on computers. The training took place in the computer lab, and at least one computer was provided to each pair of students. Teachers were asked to use books and comics taken from the curricular activities, from which they could extract the stories to be proposed to children. Teachers were also suggested to base their pre-selection on the preferences of the pupils, to increase the level of pupils' engagement. A storyboard frame was provided, together with a list of functions that are similar in stories and in coding (Sequences, Loops, Algorithms; Debugging and Reformulating; Questioning: Brennan & Resnick, 2012).

Children were then asked to choose from the different stories and legends. After having read and understood the stories, with the help of the storyboard, children were asked to draw the main sequences and write the appropriate comics. Once the story was completed, they had to build the characters with the bricks provided. Pupils were also asked to record the dialogues and to find the right code strings to make characters alive, according to the story, so that it could work without any further intervention. Finally, the projects were presented to the other groups, and pupils had the chance to share and reflect on the process they had followed. After the end of this task, post-training tests were administered, and a focus group with teachers was recorded, in which they could share their point of view and criticism through a SWOT (Strengths, Weaknesses, Opportunity, Threats) analysis (Evans & Wright, 2009).

The second training was pursued in the period between October 2019 and February 2020. 43 pupils performed the pre-tests at the beginning of October. After that, the intervention started and lasted until January 2020, for a total of 10 weeks, 3 hours per week, split into two 3rd grade classes. Teacher and researcher shared the steps of the intervention, according to the curricular needs and the classes' characteristics. Pupils were divided into groups of 2 or 3 students, which remained the same for the whole investigation. Due to the lockdown for the pandemic emergency, the final step of data collection (teachers' interview) was completed only in June 2020. However, the researchers, together with the teachers, had the chance to mention and analyze essential dimensions of the project, as emerged during the online lessons. For each class, the linguistic part (storytelling, storyboard creation, sequence analysis and description, dialogues invention) was adapted to the curricular contents. The workshops for students were pursued as follows: in the first two workshops an introduction to Scratch and the basic functions was provided; then, as for the first project, a storyboard frame and the explanation of the possible actions that could enrich the story tales (sequences, reformulations, questioning, loops) was given. This process took three hours. Two workshops were dedicated to the tales so that each group ideated the stories and created sequences and dialogues. The remaining part of the intervention was dedicated to the coding activities, in which children transferred their stories on Scratch. The project ended with a presentation of the digital stories to the other classes and parents. In June, the interview with the two teachers took place and following the structure of a SWOT analysis (Evans & Wright, 2009).

2.4 Data Analysis

The dependent variables in the model are the creative thinking skills and non-verbal cognitive abilities of children. Data from the three tests were collected before and after the training. A series of three t-tests (one for each measure), was run to evaluate whether a difference between the pre- and post-training evaluation exists. Then, the correlations between pre- and post-training were computed. Besides, a graph describing individual measures between the three tests in their final assessment was proposed to evaluate commonalities and differences in participants' performances. Analyses and figures were obtained using the statistical software R (version 4.0.3).

The Focus groups were taped and recorded, then transcribed into a written text. The coding phase was run according to the procedure described by Schreier (2012). The coding process led to the construction of coding frames that are summarised in the result section.

For the second intervention, the interviews took place on Google Meet and followed the structure of a SWOT analysis.

3. Results

3.1 First training

The t-test analysis did not show differences between the pre- and post-evaluation, in any of the three measures (Creative Thinking: $t(97)=1.207$, $p.=0.231$; Raven: $t(99)=-1.377$, $p.= 0.172$; Elithorn: $t(99)=-1.588$, $p.=0.116$).

DOI: https://doi.org/10.6092/issn.1970-2221/11583

The correlations between pre- and post-evaluation are depicted in Figure 1. The black line separates participants who improved (their post-evaluation is higher than the pre-evaluation, half above the black line) with respect to those who did not improve (below the black line).

Visual inspection of the graphs showed that, on each of the three tests, participants in the lower tail of the distribution generally improved their performance from the pre- to the post-evaluation more than what happened for the higher tail. In particular, in the Creative Thinking task the "threshold of change" (that is, the point in which participants obtained the same evaluation in the two assessments) is below the 50th percentile, indicating that for more than half of the participants the value in the post-evaluation was lower than that in the pre-evaluation. In the Raven and the Elithorn tasks, the threshold of change was roughly at the 75th and 95th percentile, respectively. In these two tests, the majority of participants improved from the pre- to the post-evaluation.

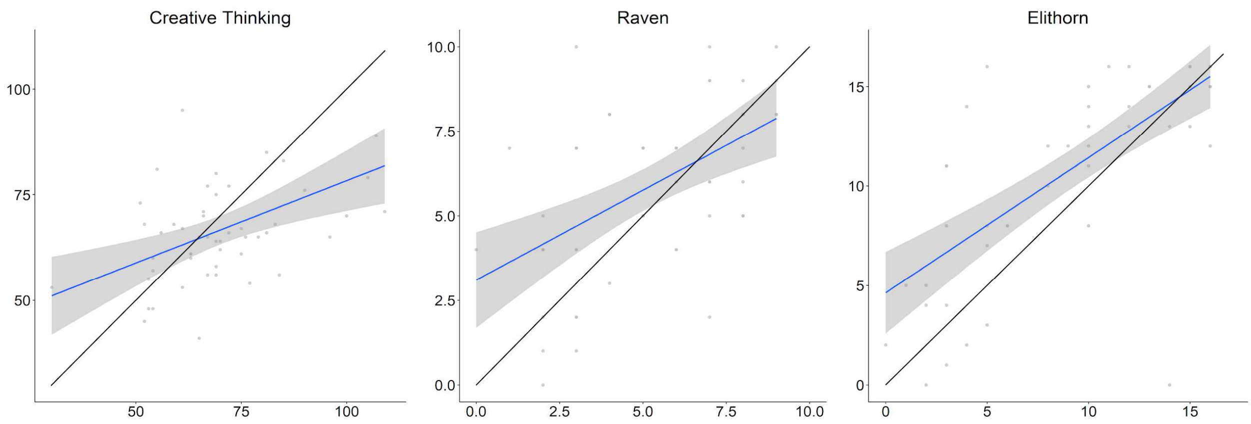


Figure 1: The three graphs represent on the x-axis the performance of the test in the first evaluation, while the post evaluation is represented on the y-axis (Creative thinking on the left, Raven in the center, Elithorn on the right). The blue lines represent the estimated correlation between the pre- and post-values (the grey ribbons represent the standard error of the estimation). The black lines represent the equivalent performance, that is, the same value in the two evaluations. Each grey dot indicates the performance of a participant in that test.

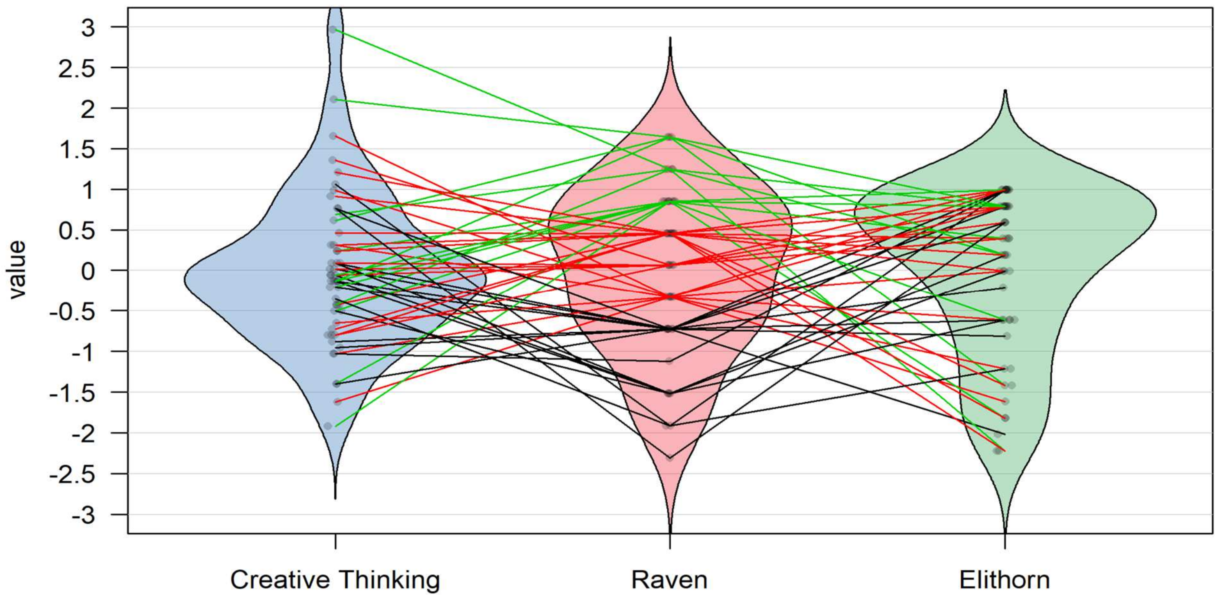


Figure 2: The violin plots represent the distributions of the participants (grey dots) on the three tests in the post-evaluation (data are normalized within each test). The lines connect the evaluations in the three tests of the same participants. Line colors (green, red and black) divided participants in three groups, depending on their performance in the Raven test.

The correlations within the tests ranged from moderate (Creative Thinking: r^2=0.47; Raven: r^2=0.53) to good (Elithorn: r^2=0.72). The correlations between the three tests, in the post-evaluations, were low to absent (Creative Thinking-Raven: r^2=0.13; Creative Thinking-Elithorn: r^2=0.20; Raven-Elithorn: r^2=0.01). A further analysis was made in order to further explore the results in the post-evaluation by analysing individual behaviour with respect to the three tests. Figure 2 visually represents the three distributions of participants' scores, one for each test. Lines were coloured based on the rank in the Raven test, separating for participants' result in that test: green for the highest third of participants, red for the second third, black for the lowest third. Roughly, it could be seen that a) the majority of the participants in the higher third scored high also in the other two tests, b) those scores which were the worst in the Raven test were generally not the worst in any of the other two tests, c) those scoring in the second third of the Raven were unpredictable in the other two tests: that is, they could obtain whatever score.

With regard to the qualitative data, Table 1 proposes the main categories emerged from the content analysis of the first Focus Group.

STRENGTHS	WEAKNESSES
• Teachers expressed self-confidence and motivation about the prospect of replicating the project with other classes. • The coding software was more straightforward with respect to their expectations.	• The initial phase, in which not every step was planned, was considered a weakness (it would likely depend on the beginning moments, when they could not imagine what to expect, mainly in terms of results).

• The preference of implementing the training during Maths or Science, even though they considered it as an activity transversal with respect to several subjects, in which storytelling is used (e.g., Music and Art). • The choice of combining coding and linguistics features. • Children were highly motivated not just to speak and write stories in the second language, but also to step-by-step organise themselves, plan and negotiate the work. • The opportunity for younger students to teach the older ones emerged two times, being probably due to the presence of two little children, but it could be a fictitious result: it could depend on when children get in contact with functions and algorithms, independently from their age. • Social skills were also mentioned by teachers, who stated that there was an improvement in the classroom atmosphere. • The tangibility and body experience features were precious, according to them. On this topic, they all agreed about the necessity of having some other materials to touch and feel, to smell, beyond to the computer desktop.	• It emerged that the study of an argument from scratch using Lego Wedo would probably be too superficial, if not previously contextualized.
OPPORTUNITIES • Teachers reported the possibility of involving not just students, but also families, young centers and libraries. Due to the flexibility of the tools, particularly of the Artec Blocks, and the potential applications in different environments, this would be a relevant improvement.	**THREATS** • The availability of tools represented a threat, probably because of the high cost, and this was connected to Lego Wedo; • The possibility of occurrence of technical problems like loss of bricks, lack of parts, software availability and compatibility, above all referred to the Lego set.

Table 1: Results (SWOT Analysis) of the Focus Group conducted with the teachers involved into the project.

4.2 Second training

As in the first training, the t-test analysis of the Creative Thinking test did not show differences between the pre- and post-evaluation ($t(82)=0.850$, $p.=0.398$). Conversely, both Raven and Elithorn tests showed a significant improvement from the pre- to the post-evaluation (respectively: $t(84)=2.458$, $p.=0.016$; Elithorn: $t(42)=8.425$, $p.<0.001$).

As in the first training, a correlation between pre- and post-test evaluation was computed (depicted in Figure 3). Even in this case, the black line separates participants who improved (their post-evaluation is higher than the pre-evaluation, half above the black line) with respect to those who did not improve (below the black line). In

DOI: https://doi.org/10.6092/issn.1970-2221/11583

this second training, the analysis showed that in the Creative Thinking test, participants in the lower tail of the distribution generally improved their performance from the pre- to the post-evaluation more than what happened for the higher tail. In the Raven's Progressive Matrices and, mostly in the Elithorn, the great majority of participants improved from the pre- to the post-evaluation. The correlations within the tests are reported as follows: Creative Thinking: r^2=0.75; Raven: r^2=0.65; Elithorn: r^2=0.42.

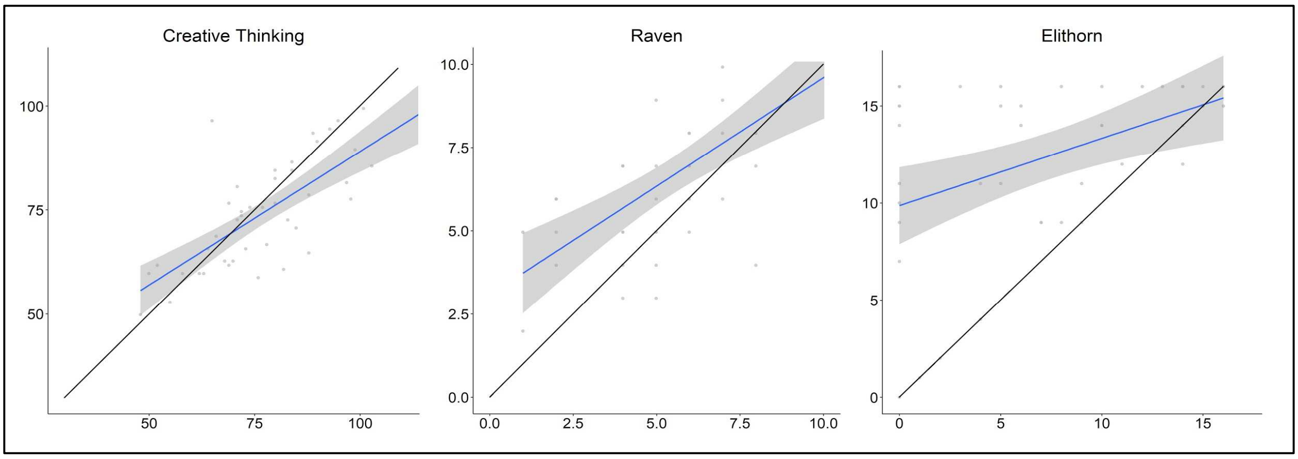
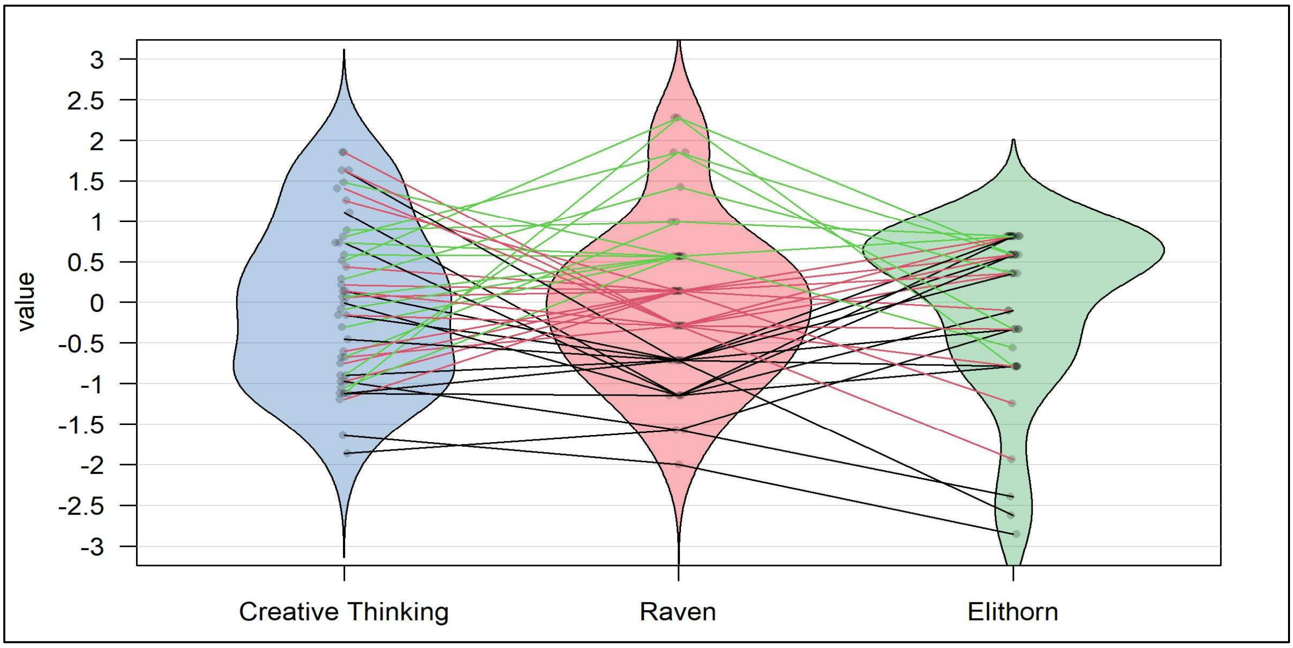


Figure 3: Likewise Figure 1, the three graphs represent on the x-axis the performance of the test in the first evaluation of the second training, while the post evaluation is represented on the y-axis (Creative thinking on the left, Raven in the center, Elithorn on the right). The blue lines represent the estimated correlation between the pre- and post-values (the grey ribbons represent the standard error of the estimation). The black lines represent the equivalent performance, that is, the same value in the two evaluations. Each grey dot indicates the performance of a participant in that test.



DOI: https://doi.org/10.6092/issn.1970-2221/11583

Figure 4: The violin plots represent the distributions of the participants (grey dots) on the three tests in the post-evaluation of the second training (data are normalized within each test). The lines connect the evaluations in the three tests of the same participants. Line colors (green red and black) divided participants in three groups, depending on their performance in the Raven test.

Even in this case, it was possible to evaluate the individual behaviour for the three tests. Figure 4 visually represents the three distributions of participants' scores, one for each test, with the same features described for Figure 2. Even in this second training, it could be observed that a) the majority of the participants in the higher third scored high also in the other two tests, c) those scoring in the second third of the Raven were unpredictable in the other two tests, likewise the first training.

The ANOVA analysis in Table 2 reports the values obtained by merging the two samples in the Creative Thinking test. In this first measure, the two samples had different starting points, whereas in the other two measures (Raven Progressive Matrices and Elithorn maze test) the starting points were similar. As a consequence, researchers considered presenting the results shifted for the two groups, also considering external factors (e.g., instructors could have been more skilled in the second training).

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
time	1	0.00	0.000	0.00	1.000000
training	1	12.96	12.956	13.79	0.000271 *
time: training	1	0.21	0.207	0.22	0.639529
Residuals	184	172.84	0.939		

Table 2: The ANOVA Analysis reports the values obtained among the two samples (N=51 and N=43) in the Creative Thinking test.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
time	1	0.00	0.000	0.00	1.000000
training	1	2.42	2.4173	2.43	0.121
time: training	1	0.56	0.5570	0.56	0.455
Residuals	184	183.03	0.9947		

Table 3: The ANOVA Analysis reports the values obtained among the two samples (N=51 and N=43) in the Raven Progressive Matrices.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)

DOI: https://doi.org/10.6092/issn.1970-2221/11583

time	1	0.00	0.000	0.00	1.000000
training	1	0.06	0.063	0.064	0.8001
time: training	1	5.05	5.049	5.136	0.0246 *
Residuals	184	180.89	0.983		

Table 4: The ANOVA Analysis reports the values obtained among the two samples (N=51 and N=43) in the Elithorn test.

In the line time:training, the interaction between the factors time (pre- and post-test) and training (the two samples :N=51 and N=43) is reported . The p-value related to the Elithorn test might suggest a summative effect between the two factors (time and training). The content analysis emerged from the interview related to the second training is reported in Table 5.

Strengths	Weaknesses
• Eye-motor skills - linked to the mouse use. • Developing creativity. • Teamwork. • Motivation to use the computer to create – invent. • Awareness of proactive opportunities in the pc use. • Critical pc use has been understood (many have grasped the weakness of using it passively). • Lack of the object-to-think-with was considered not entirely harmful (They already have had experience with B-boot in the previous year). • Pair/interactive work - even more effective than other subjects or fields (everyone could field their agent skills, and cope with difficulties). • Coping with difficulties in pairs (Some children could not have been able to code alone from the beginning); • Greater respect for turn compared to other activities, including playful activities.	• The functional part has been a bit long (picture finding, characters drawing) - image search has been time-consuming. • The division into groups is necessary but working with small and few groups is recommended. • Work in pairs would have been more efficient (pairs management to be organized). • Resources to improve the context condition - more teachers - more laptops.
Opportunities	**Threats**
• The heuristic path of each group or child could be improved if more resources were available. • Small groups would be ideal (max. 8 children). • Improve the significance of initial and final moments of reflection. • Positive and negative outcomes not emerged adequately due to the lockdown of schools.	• Overexposure to the screen - children should be educated to the risk of overexposure. • Effects of prolonged use of a computer could be overestimated. • A systematic introduction of the project could cause a monodisciplinary approach, with activities strictly related to the CS ("it would be a pity to waste such a project considering it an end in itself").

DOI: https://doi.org/10.6092/issn.1970-2221/11583

• Assigning homework could create a link between home and school - kids, parents. • Digital divide (having the computer - during lockdown was not taken for granted). • Fostering the relationship and the skill divide (some children explained to their parents some processes). • Gap between who was "scaffolded" at home and who not. • Storytelling option to improve scientific texts exposure - geometry – mathematics. • Scientific part and the humanistic part (stories - tales) were both fostered. • Challenging the prejudices of some teachers who believe that they cannot implement it in the language classes. • Dealing with issues of digital civic education - in retrospect. • Creating conditions to address hot topics (identity protection - cyberbullying) in a protected environment. • Defending themselves against fake news - raise awareness about the need to prove the sources.	• Collaboration with the postal police to raise awareness of the real risks - especially for times when the young student will be alone in front of the computer ("It's good that the search for taboo words takes place at school, in a protected environment"). • Teachers could be not well prepared for some demanding moments - how to handle an embarrassing situation - how to respond - how to clarify with the class (cyberbullying, hate speech). • Lack of awareness that machines do not solve but help. • Concern about spending too much time in front of the screen.

Table 5: Results (SWOT Analysis) of the interview conducted via Google Meets on June 19th, 2020 with the teachers involved into the project.

5. Discussion

This study aimed to evaluate the effectiveness of a Computational Thinking training, planned and executed by both researchers and teachers, onto a set of cognitive skills. Results obtained from the tests and the SWOT analysis were partially overlapping, but both showed how the proposed intervention was generally successful.

The first research question obtained a mixed answer. While the comparison of the pre- and post-training evaluations through the t-test did not achieve the significant level, the direct comparison of individual performances showed that spatial and logical skills were likely to improve due to the training. The Elithorn perceptual maze test, which is used to measure the planning skill, showed that the post-training evaluation was more likely to be higher than the first evaluation in the large majority of students. Almost the same could be said about the Raven Standard Progressive Matrices test, used to study logical thinking, limiting the contribution of cultural and linguistic components. A moderately high positive correlation between performances in Elithorn and Raven post-evaluations could probably depend on the nonverbal-visuospatial reasoning activity prompted during the coding, which could be a shared component in the two tests. The improvement of scores in the first two tests corroborates the recent findings of Tsarava (2019) and Parkinson (2018). If CT is linked to visuospatial and logical skills, these skills are likely to underlie the process of coding and, in turn, they would help to plan algorithms and to imagine their outcomes. The technical process could be facilitated by these cognitive skills, while it seems

DOI: https://doi.org/10.6092/issn.1970-2221/11583

that creativity does not benefit from it in the same fashion. The Creative Thinking test did not show substantial changes, probably because there was high variability between the first and the last evaluations.
It could be argued that the adopted software have not properly stimulated creativity because of some intrinsic characteristics, for example, few programming blocks, fundamental user-friendly aspects, a scarce opportunity of interaction with the community of users, scarce personalisation options. An alternative explanation could be based on the method: either the duration of the training could be not enough to produce changes in such a higher cognitive process, or the selected tests were not sensible enough to detect changes.
The second research question was focused on the didactic aspects related to the peculiarities of the training, conducted in the second language classes. The qualitative data analysis showed that teachers considered the storytelling aspect of creating animations as interdisciplinary and transversal to many subjects. Furthermore, the aspect of having, in the first training, a tangible object to build was very positively rated as essential, both for the plays and video recordings of the stories and for the possibility to touch and feel what students were building. The initial phase generated uncertainty and anxiety in teachers, probably because they did not attend any preparation course. In this regard, it could be suggested to either provide teachers with a draft program for the several meetings of the training or to plan with them the details of the first meetings at least. Recent literature has proposed that a lack of training created with teachers, in general, may have compromised the validity of the results obtained by the application of CT training in schools (Armoni, 2018; Guzdial et al., 2019). Furthermore, research on cognitive planning (Basso & Saracini, 2020) has shown that the planning at the beginning of a task is likely to pursue two goals: definition of the general plan to achieve the final state, and outline of the first actions. Given that this involuntary behaviour is commonly implemented, it follows that teachers, to face anxiety, may need to know in advance the outline of the task and the details of the first actions to be taken.
Then, the implications on many different subjects and the versatility resulted in being an essential aspect for both teachers and students. The storytelling approach has shown to be very helpful (Wright, 1995; Papadimitriou, 2003) while the decision not to apply the training linked to "scientific" subjects may have convinced teachers that coding should not be confined to mathematics, but it could be rewarding for other subjects as well. Computational thinking as a set of highly demanded skills in the digital society may strongly contribute to the building of general and specific knowledge of the students, who represent the future human capital. The tendency of introducing it in the school curriculum from kindergarten to high school has led to several contributions, from different fields, and the debate is still open for example, on teacher training, as Armoni (2018) pointed out, and on assessment issues (Fraillon et al., 2018). Among other aspects, the related pedagogical and psychological aspects may be worth investigating because the mainstream in the study of CT is focused on defining the best way in which students could learn to code. Primary school children are, in general, motivated to use computers and other digital tools, but even if they are considered "digital natives" (Palfrey & Gasser, 2011), they need to learn CS likewise any other disciplines. The present study has evidenced the relevance of visuospatial and logical skills, while coding activities could promote their development and, in turn, assist the teaching of CT. Besides, students would be facilitated by high order thinking skills like critical thinking and metacognition. The former was thought to play a central role in programming, given that it entails the abilities to analyse pros and cons about choices and to think forward (Imperio et al., 2020), while the latter was defined as "higher-

order thinking that involves active control over the cognitive processes engaged in learning" (Livingston, 2003; p. 2).

Another relevant aspect of the present study consisted of the inclusion of teachers from the planning phase of the activities. The increase of motivation and devotion to the activity when people are asked to take an active role in all the phases is straightforward. Although the results on cognitive processes would likely be due to this factor, this remains speculation, which deserves further support. Notwithstanding, it is still surprising that only recently, teachers and their methods were considered central factors in the evaluation of whatsoever training in school (Guzdial, 2019). Based on the results of the focus group and the interview, it is recommended to adopt the practice of including teachers in the research, at least for the validation of the proposed training.

6. Limitations

The current work contains limitations due to the sample, which was not randomly assigned and self-selection bias. The researcher chose the groups according to the availability of the teachers. Therefore, our results could be generalised to a similar population of students and teachers. Moreover, notwithstanding the training was the same, the different teaching styles could have affected the internal validity of the results. Also, the difference between the results of the t-tests and the results depicted in Figure 1 and Figure 2, indicates that the sample size was probably not big enough.

Further limitations are related to the materials. The different conditions of schooling between the two training sessions need to be further investigated, particularly when home-schooling is becoming more than an option.

7. Future work and conclusion

Although it is not possible to consider the current study as a milestone, the evaluation of individual performance was still helpful and allowed development of knowledge related to the contribution of some cognitive skills and the relevance of teachers' evaluation. Researchers are encouraged to replicate the present study while coping with the limitations stated above. In addition, it is likely that the study would improve by adopting other software (e.g., Alice, Tynker) and including a control group in the design. Another line of research would focus on the different executive functions, which could allow exploring the skills underlying the construct. For example, among the set of cognitive skills, visuospatial planning emerged as a promising skill to be investigated. The Maps task (Basso et al., 2001) is likely to be the best candidate for this purpose.

The main contribution of the current study is likely to consists in having defined how collaboration between researchers and teachers is relevant. Therefore, this line of research seems really promising and deserves attention. Future interventions may focus on the pedagogical aspects emerged in the current study, posing teachers and teacher training as focal points and introducing new approaches like the "Lifelong Kindergarten", as proposed by Resnick and Robinson (2017).

DOI: https://doi.org/10.6092/issn.1970-2221/11583

Bibliography

Armoni, M. (2018). Computing in schools. Training teachers for K-6 computing education. *ACM Inroads*, *9*(3), 19-21.

Basso, D., Bisiacchi, P.S., Cotelli, M., & Farinello, C. (2001). Planning times during Travelling Salesman's problem: differences between closed head injury and normal subjects. *Brain and Cognition*, *46*(1-2), 38-42.

Basso, D., Fronza, I., Colombi, A., & Pahl, C. (2018). Improving assessment of computational thinking through a comprehensive framework. In: *ACM Proceedings of the 18th Koli Calling International Conference on Computing Education Research*, a15.

Basso, D., & Saracini, C. (2020). Differential involvement of left and right frontoparietal areas in visuospatial planning: An rTMS study. *Neuropsychologia*, *136*, 107260.

Baytak, A., & Land, S. M. (2011). An investigation of the artifacts and process of constructing computers games about environmental science in a fifth-grade classroom. *Educational Technology Research and Development*, *59*(6), 765-782.

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In: *Proceedings of the 2012 annual meeting of the American educational research association*, Vancouver, Canada.

Ciancarini, P., Missiroli, M., & Russo, D. (2019). Cooperative thinking: Analyzing a new framework for software engineering education. *Journal of Systems and Software*, *157*, 110401.

Çiftci, S., & Bildiren, A. (2019). The effect of coding courses on the cognitive abilities and problem-solving skills of preschool children. *Computer Science Education*, *30*(1), 1-19.

Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Los Angeles: Sage.

Davies, A. D., & Davies, M. G. (1965). The difficulty and graded scoring of Elithorn's perceptual maze test. *British Journal of Psychology*, *56*(2-3), 295-302.

Eriksson, E., Iversen, O. S., Baykal, G. E., Van Mechelen, M., Smith, R., Wagner, M., & Hjorth, A. (2019). Widening the scope of fablearn research: Integrating computational thinking, design and making. In: *ACM Proceedings of the FabLearn Europe 2019 Conference*, a15

Evans, C., & Wright, A. (2009). How to conduct a SWOT analysis. *The British Journal of Administrative Management*, *24*(65), 10-34.

Falkner, K., Sentance, S., Vivian, R., Barksdale, S., Busuttil, L., Cole, E., Liebe, C., Majorana, F., McGill, M.M., & Quille, K. (2019). An international comparison of K-12 computer science education intended and enacted curricula. In: *ACM Proceedings of the 19th Koli Calling International Conference on Computing Education Research*, a4.

Fraillon, J., Ainley, J., Schulz, W., Duckworth, D., & Friedman, T. (2019). *IEA international computer and information literacy study 2018 assessment framework*. Springer International Publishing.

Froebel, F. (1903). *The Education of Man*. New York, London: D. Appleton Company. (Original work published in 1826).

Giles, G. C. (1964). Predictive validity of progressive matrices and two other nonlanguage tests of mental ability. *Journal of Educational Measurement*, 1(1), 65-67.

Guzdial, M. (2015). *Learner-centered design of computing education: Research on computing for everyone*. San Rafael: Morgan & Claypool Publishers.

Guzdial, M., Kay, A., Norris, C., & Soloway, E. (2019). Computational thinking should just be good thinking. *Communications of the ACM*, *62*(11), 28-30.

Imperio, A., Kleine Staarman, J., & Basso, D. (2020). Relevance of the socio-cultural perspective in the discussion about critical thinking. *Ricerche di Pedagogia e Didattica*, *15*(1), 1-19.

Kong, S., & Abelson, H. (Eds.). (2019). *Computational thinking Education*. Singapore: Springer Nature.

Livingston, J. A. (2003) *Metacognition: an overview*. Retrieved September 27, 2020 from: https://files.eric.ed.gov/fulltext/ ED474273.pdf.

Lowe, T., Brophy, S., & Cardella, M. (2019). Exploring the definition of computational thinking in research and the classroom. *Paper presented at the Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, Minneapolis, Minnesota.

Mendelsohn, P., Green, T. R. G., & Brna, P. (1990). Programming languages in education: The search for an easy start. In: J. Hoc, T. R. G. Green, R. Samurçay & D. J. Gilmore (Eds.), *Psychology of programming* (pp. 175-200). London: Academic Press.

Montessori, M. (1995). *The absorbent mind*. New York: Macmillan.

Morgan, D. L. (1996). Focus groups. *Annual Review of Sociology*, *22*(1), 129-152.

Nardelli, E. (2019). Do we really need computational thinking? *Communications of the ACM*, *62*(2), 32-35.

Palfrey, J. G., & Gasser, U. (2011). *Born digital: Understanding the first generation of digital natives*. New York: Basic Books.

Papadimitriou, C. H. (2003). MythematiCS: In praise of storytelling in the teaching of computer science and math. *SIGCSE Bulletin*, *35*(4), 7-9.

Papert, S. (1993). *Mindstorms: Children, computers, and powerful ideas*. New York: Basic Books.

Parkinson, J., & Cutts, Q. (2018). Investigating the relationship between spatial skills and computer science. In: *Proceedings of the 2018 ACM Conference on International Computing Education Research, Espoo, Finland* (pp. 106-114).

Pea, R. D., & Kurland, D. M. (1984). On the cognitive effects of learning computer programming. *New Ideas in Psychology*, *2*(2), 137-168.

Raven, J. (2000). The raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology*, *41*(1), 1-48.

DOI: https://doi.org/10.6092/issn.1970-2221/11583

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., et al. (2009). Scratch: Programming for all. *Communications of the ACM*, *52*(11), 60-67.

Resnick, M., & Robinson, K. (2017). *Lifelong kindergarten: Cultivating creativity through projects, passion, peers, and play*. Boston: MIT Press.

Schreier, M. (2012). Qualitative content analysis in practice. In: U. Flick (Ed), *The Sage handbook of qualitative data analysis*. London: Sage Publications (pp. 170–183).

Sengupta, P., Kinnebrew, J. S., Basu, S., Biswas, G., & Clark, D. (2013). Integrating computational thinking with K-12 science education using agent-based computation: A theoretical framework. *Education and Information Technologies*, *18*(2), 351-380.

Sheehan, K. J., Pila, S. C., Lauricella, A. R., & Wartella, E. A. (2019). Parent-child interaction and children's learning from a coding application. *Computers & Education, 140*, 103601.

Siegler, R. S. (2016). Continuity and change in the field of cognitive development and in the perspectives of one cognitive developmentalist. *Child Development Perspectives*, *10*(2), 128-133.

Szabo, C., Sheard, J., Luxton-Reilly, A., Becker, B. A., & Ott, L. (2019). Fifteen years of introductory programming in schools: A global overview of K-12 initiatives. In: *ACM Proceedings of the 19th Koli Calling International Conference on Computing Education Research*, a8.

Tedre, M., & Denning, P. J. (2016). The long quest for computational thinking. In: *ACM Proceedings of the 19th Koli Calling International Conference on Computing Education Research*. (120–129).

Trilling, B., & Fadel, C. (2009). *21st century skills.: Learning for life in our times*. San Francisco: John Wiley & Sons.

Tsarava, K., Leifheit, L., Ninaus, M., González, R. M., Butz, M. V., Golle, J., & Moeller, K. (2019). Cognitive correlates of Computational Thinking: Evaluation of a blended unplugged/plugged-in course. In: *ACM Proceedings of the 14th Workshop in Primary and Secondary Computing Education, Glasgow, Scotland*, a24.

Williams, F. (1994). *TCD. test della creatività e del pensiero divergente.* Trento: Erickson.

Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, *366*(1881), 3717-3725.

Wright, A. (1995). *Storytelling with children*. Oxford: Oxford University Press.

Marialaura Moschella is PhD Student in General Pedagogy, Social Pedagogy, General Didactics, and Disciplinary Didactics, Faculty of Education, Free University of Bolzano. Her research interests lie at the intersection of Computer Science (particularly Computational Thinking) and Education (Didactics and Curriculum Development). Research contributions have been to investigate the skills underlying Computational Thinking, and to consequently improve the K-12 Computer Science Curriculum, leveraging on the collaboration between researchers and teachers.
Contact: mmoschella@unibz.it

Demis Basso is Associate Professor of *General and Cognitive Psychology* at CESLab, Faculty of Education, Free University of Bozen-Bolzano, and visiting researcher at the Research Center in *Neuropsychology and Cognitive Neuroscience* (CINPSI Neurocog) of Universidad Católica del Maule (UCM), Talca, Chile. His research interests are Visuo-spatial Planning, Prospective Memory and Life Skills, with relevance on education.
Contact: demis.basso@unibz.it