# Can Computational Thinking help me? A quantitative study of its effects on education

Rivanilson S. Rodrigues
Federal University of Campina Grande
Software Practices Laboratory – SPLab
Paraíba, Brazil 58429–900
rivanilson@copin.ufcg.edu.br

Wilkerson L. Andrade
Federal University of Campina Grande
Software Practices Laboratory – SPLab
Paraíba, Brazil 58429–900
wilkerson@computacao.ufcg.edu.br

Livia M. R. Sampaio Campos
Federal University of Campina Grande
Distributed Systems Laboratory – LSD
Paraíba, Brazil 58429–900
livia@computacao.ufcg.edu.br

***Abstract*—Some researchers consider that the benefits of Computational Thinking (CT) in education are unclear, and therefore a quantitative study to understand the effects of CT skills in the basic levels of formal education is needed. The present study aims to analyze the effects of CT on the academic performance of students in high school. We seek to answer the following research questions: (RQ1) Can CT increase the problem-solving skills of students in high school? (RQ2) Is CT related to the performance of students in high school? (RQ3) Can CT improve the performance of students in high school? We conducted an analysis between two groups of students with similar profiles belonging to the same population, where the factor CT proficiency is the only feature that distinguishes both groups. The results show that there is a moderate correlation between the performance of evaluated students and their skills of CT developed by computer programming. In addition, we found that students proficient in computer programming have superior performance up to 10.2% in ENEM (National Exam of High School) assessment compared to those with little or no experience level and even better performance 21% in the WASI test problem-solving in students on attending high school.**

## I. Introduction

A quality education is vital to the socioeconomic development of a population. However, for many countries, providing quality education is still a challenge [1]. This problem is partly demonstrated by the Organization for Economic Cooperation and Development – OECD through the Programme for International Student Assessment – PISA, that assesses student proficiency level on basic education.

According to OECD 65 countries (representing more than 80% of the world economy) participated in PISA 2012, involving about 510.000 students aged 15 to 16 years (representing 28 million students). Students were evaluated in areas of math, reading and problem solving, and 21.4% of which did not reach the baseline Level 2 proficiency in problem solving. This level is considered too low by OECD for a student to continue in adult life and to exercise their citizen rights. [2].

The OECD report clearly indicates the need to study methods to enhance learning in problem solving by students on basic education. In this sense, some researches suggest that Computational Thinking skills can be used to improve problem solving skills, systemic and algorithmic thinking, problems decomposition and abstraction.

Computational Thinking – CT can be defined as a set of methods, techniques and skills for problem solving based on Computer Science. CT skills are not limited to the field of computing, but are extensible to mathematics, biology, science, economics, reading and other areas [3], [4], [5], [6].

On the other hand, there are few quantitative studies published in the literature that have demonstrated evidence of the effects of CT in basic education [7]. Thus, many researchers pointed out that, the benefits of CT in this context are unclear. According to the authors [8], [7], [6] it is necessary to acquire theoretical and practical knowledge on how to assess CT skills and understand their effects on formal learning process to make clear CT skills for students on basic education.

Considering this scenario, the objective of this study is to analyze the effect of CT on the performance of high school students. We focus on answering the following research questions: (RQ1) Can CT increase the problem-solving skills of students in high school? (RQ2) Is CT related to the performance of students in high school? (RQ3) Can CT improve the performance of students in high school?

In order to explore the above mentioned research questions we conducted an empirical study with ex-post facto design [9], using two groups of students (experimental and control) where one group is mainly differentiated from the other by the proficiency in CT.

For this work, we considered 149 volunteer students from Brazilian schools where 36 of them were in high school, and 113 had just completed high school. These two groups of students were evaluated by the WASI – Whimbey Analytical Skills Inventory [10] and National High School Exam – ENEM [11], respectively assess performance in problem-solving and student proficiency in high school.

The results show that students learning programming topics during high school have a performance of 21% in the WASI problem-solving exam. Moreover, we observed a moderate correlation between general performance of students in ENEM and their skills in computer programming. We also found that the performance of students who learned programming in high school is statistically better in the areas of Mathematics (10.2%) and Natural Sciences (4.4%) as compared to students not been submitted to CT skills (control group).

This paper is organized as follows: in Section 2, we briefly describe key concepts of CT, Basic Education in Brazil and

National High School Exam – ENEM. Section 3 is dedicated to related works. In Section 4 we detail the method used in our research, design, research questions and research hypotheses, collection method, and data analysis. In Section 5 we present the results and threats of validity and, finally, the main conclusions and future works are presented in Section 6.

## II. Background

In this section, we present the key concepts for the general understanding of this study. In the next subsections we discuss CT, Basic Education in Brazil, National High School Exam – ENEM and the relationship between ENEM and CT.

### A. CT: definitions, skills and measure

The term CT was established in the literature by Seymour Papert [12] from practical experimental studies and using the LOGO programming language to support the teaching of school children.

Papert defines implicitly the CT as a reflective reasoning process developed from theoretical study and practice of programming language. The author suggests that introducing a programming language in the curriculum can make learning in different disciplines most significant for the student helping him/her to think reflectively and in a self-conscious way [12].

Jeannette Wing popularized the term CT and defined it as being based on fundamental concepts of Computer Science; combines and completes mathematical thinking and engineering; and includes a variety of mental tools. At its core, is the problem solving process, systems design and abstraction, not just programming. Also according to Wing, everyone should learn about CT considering its influence in many fields of study [3].

Currently, there are multiple CT definitions in related literature. According to Chenglie Hu [13], the definition of CT by Wing [3] is very abstract, and even in later works [14] this concept still open to multiple interpretations.

Barr and Stephenson consider that CT practices allow students to become knowledge-builders, once they use concepts such abstraction, recursion and iteration to process information, analyze data and create real and virtual artifacts [5]. These authors have defined 9 concepts that make up the core of CT:

*Data Collection*: the process of collecting information about a problem; *Data Analysis*: making sense out of the data, find patterns, and draw conclusions; *Data Representation*: representing and organizing data in charts, tables, text or figures; *Decomposition Problems*: breaking a complex problem into simple smaller and manageable tasks; *Abstraction*: reducing the complexity of a problem so as to define its main ideas; *Algorithms and Procedures*: sequence of steps to solve a problem or achieve some end; *Automation*: using computers or machines to perform repetitive tasks; *Parallelization*: organizing resources to simultaneously perform tasks to achieve a common goal; *Simulation*: representation or model of a process. The simulation also involves experiments being performed using models.

We can learn CT studying computer science foundations, these may or may not be integrated into curriculum subjects during basic education or higher education. According to Resnick [15] study computer programming is a way to understand the core ideas related to CT, because it is possible to develop and increase skills, such as logical thinking, problem solving and algorithmic thinking.

There are no specific tests or assessments to measure CT skills of a student. This assessment generally uses tests involving logic and problem solving, for example, WASI exam [10] used by [16] or through tests that assess specific knowledge on Computer Science, such as programming [17].

### B. Basic Education in Brazil

In Brazil, basic education consists of three stages: early childhood or preschool, elementary school and high school. Its purpose is to ensure the student the necessary common training for the exercise of citizenship and provide them with the means to progress at work and in later studies [18]. Public schools administered by municipalities, states and the federal government guarantees free basic education to all people, however, it is also possible completing any stage of basic education in private schools.

Early childhood education or preschool is the first stage of basic education. Its goal is the development of children up to five years by complementing the action of the family and community. However, it is not a required step [18].

Elementary school is compulsory for children from 6 years old, its goal is to provide basic training to the students and its duration is nine (9) years beginning in the 1st year and ending in the 9th year. Basically, the following compulsory subjects are addressed during the formation of the student at this stage: Portuguese Language, Foreign Language, Art, Physical Education, Mathematics, Natural Sciences, History and Geography [18], [19].

The high school is the last stage of basic education and requires, as a prerequisite, the completion of elementary education. The objective of this stage is to deepen the knowledge acquired in elementary school aimed at basic preparation for work and citizenship. The regular duration of this stage is 3 years divided into grades, starting in 1st grade and ending in the 3rd grade of high school [18].

The curricular organization of high school is divided into three major areas of knowledge: *Languages, codes and its technologies:* Portuguese Language and Literature, Foreign Language, Art, Physical Education and basic concepts of Information Technology and Communication; *Humanities and its Technologies:* History, Geography, Philosophy and Sociology; *Natural Sciences, Mathematics and its technologies:* Mathematics, Chemistry, Physics and Biology [20].

### C. The National Exam of High School – ENEM

The National High School Exam – ENEM is an optional assessment used to evaluate the performance of students at the end of basic education in order to improve the quality of the brazilian basic education. In addition, the ENEM results are also used as a selection criteria for access to higher education in public and private institutions [11].

TABLE I. Common reference matrix to all areas of knowledge assessed in the ENEM. Extracted from: Inep

| Competence | Description |
|---|---|
| Master Languages (ML) | Mastering the standard Portuguese language and make use of mathematical language, artistic, scientific and languages Spanish or English. |
| Understanding Phenomena (UP) | Construct and apply concepts from several areas of knowledge to understand natural phenomena, historical and geographical processes, technological production and artistic events. |
| Confront problem-situations (CP) | Select, organize, relate, interpret data and information represented in different ways, to make decisions and confront problem situations. |
| Construct Argumentation (CA) | Relate information, represented in different ways, and knowledge available in concrete situations to build consistent arguments. |
| Elaborating Proposals (EP) | Use knowledge developed in school to make proposals for intervention in reality, respecting human values and considering the socio-cultural diversity. |

The ENEM is structured by a matrix of skills (see Table I) for all evaluated knowledge areas. The ENEM assessment consists of one (1) essay (argumentative text) in Portuguese Language about a political theme, social or cultural. Moreover, four (4) objective tests, each containing 45 (forty five) multiple choice questions assessing the following knowledge areas and curriculum components: *Languages, Codes and their Technologies and Essay:* Portuguese Language and Literature, Foreign Language (English or Spanish), Arts, Physical Education and basic concepts of Information and Communication Technologies; *Humanities and its Technologies:* History, Geography, Philosophy and Sociology; *Natural Sciences and their technologies:* Chemistry, Physics and Biology; *Mathematics and its technologies: Mathematics*.

The performance of the ENEM is calculated based on *Item Response Theory – IRT* and varies on a proficiency scale 0 to 1000 [11].

## III. Related works

In order to provide quantitative arguments to the premise that computer skills are essential for students in basic education, Oliveira [7] describes a quantitative study that correlates the student's performance from 6th to 9th grade of a Brazilian elementary school with CT skills. To measure the CT skills he used a logical reasoning test based on a model of Turing machine proposed in his work. As conclusion, the author expresses the existence of a significant positive correlation between the evaluation carried out and the average performance of students on school.

Similar to Oliveira we evaluate the relationship between the performance of students in basic education and CT, however, in our study we analyze this relationship at the end of basic education using the performance of students in ENEM, and the CT is measured in terms of computer programming skills. In addition, we seek to quantify the effect of CT for each area assessed in the ENEM by means of a comparative analysis between control and experimental groups.

Van Dyne and Braun [16] propose and evaluate a CT course called CS0, his goal is to support the development of problem solving skills in college students. The contents covered in the course involved logical reasoning, problem solving and aspects of programming using the robotic kit Lego NXT Mindstorms [21]. To measure effects of the course they used the WASI exam and analysis of the study looked at two groups of students, the experimental group formed solely with CS0 students and the control group formed solely with FESP students. FESP – Foundations of engineering and Science program is a course similar to CS0 facing engineering and science programs. Based on a statistical analysis Van Dyne and Braun concluded that, CS0 students had a significantly improved outcome compared to FESP students. Thus, for the authors, studying CT contents any student can improve their analytical skills [21].

This work differs from Van Dyne and Braun, mainly from methodological viewpoint, since we use a design research ex post facto, Van Dyne and Braun opted for an action research [9]. In addition, our sample is composed of high school students and we adopt a control group without experience in CT.

## IV. Metodology

For this work, we conduct an empirical study featured as *ex post facto* research. This research types is characterized by verifying possible relationships of cause and effect between factors and relevant prior aspects to a population or group [9].

### A. Research design

The research design is exposed in Table II. It's based on [9] and consists of defining two groups of students from the same population (control and experimental) and with similar profiles, where only the factor CT in the experimental group is the unique characteristic that distinguishes both groups. The following variables composed this design:

TABLE II. Research design

| Groups | Independent variable | Dependent variable |
|---|---|---|
| Experimental | *CT* | $P_1$ |
| Control | | $P_2$ |

*Computational Thinking* (CT): We believe that computer programming skills comprises the core of Computational Thinking [15], thus CT independent variable is related with learning computer programming during high school.

*Performance* (P): This variable represents the student's performance in WASI or ENEM exams.

### B. Research questions and hypothesis

The following research questions motivated the development of this work:

RQ1: Can CT increase the problem solving skills of students in high school?

H1.0: There is no statistical evidence that CT can increase the students' problem-solving skills;

H1.1: There is a statistical evidence that CT can increase the students' problem-solving skills;

RQ2: Is CT related to the performance of students in high school?

H2.0: There is no evidence of correlation between the students' performance in high school and skills in CT;

H2.1: There is a correlation between the students' performance in high school and skills CT;

RQ3: Can CT improve the performance of students in high school?

H3.0: There is no statistical evidence that CT can improve students' performance in high school;

H3.1: There is statistical evidence that CT can improve students' performance in high school;

### C. Samples and data collection

We studied samples of students from the population of Brazilian high schools. The samples are comprised of data collected from 149 volunteers and the process of data collection occurred in two phases:

The first phase took place in a public school, involving volunteers from 2th grade of high school. In the experimental group, all students had concluded a course about computer science topics with emphasis on programming through games using Unity-2D engine tool ©.

The control group was similar to the experimental group in terms of profile, in other words, family income, extracurricular activities and type of school (public or private).

TABLE III. Attending high school

| Groups | Male | Female | Total |
|---|---|---|---|
| Experimental | 7 | 9 | 16 |
| Control | 8 | 12 | 20 |
| Total | 15 | 21 | 36 |

In this phase we assessed the problem solving ability of students through the WASI. WASI is an exam that assesses problems solving skills, it is originally formed by a pre and a post-test, with 38 and 37 problems respectively, over different categories: verbal reasoning problems, sequential instructions, training analogies, analysis trends and patterns, mathematical resolution and sentences ratio [10]. In this work, only pre-test has been utilized.

TABLE IV. High school graduates

| Groups | Federal | Public | Private | Total |
|---|---|---|---|---|
| Experimental | 7 | 19 | 36 | 62 |
| Control | 4 | 19 | 28 | 51 |
| Total | 11 | 38 | 64 | 113 |

The second phase of data collection was carried out at the Federal University of Campina Grande - UFCG with freshmen students of the undergraduate course in Computer Science. The data collected was related to the students' previous academic life, and consisted of detailed information about the performance of each student in the ENEM exam (used as selection criterion for admission to college education in Brazil, see Section II); type of school (private or public) attended by the students during high school; whether the student have participated in knowledge Olympics or not; computer programming experience at some stage of their formal education. The level of knowledge in computer programming, was measured by a programming test comprised of 23 problems involving sequential structures, conditionals, loops, data structures and error handling in Python programming.

### D. Data analysis process

For data analysis we used the R programming language. The data analysis process involved hypothesis testing, effect size and correlation. All tests are performed considering statistical assumptions and a confidence level of 95%.

To deal with heterogeneity of the experimental and control groups in the sample (see table IV) we conducted a resampling process based on the *bootstrap* method [22]. Thus, groups were analyzed with the same proportions of students of each school type (private, public state or federal).

### E. Reproducibility

We released our data set, statistical tests and all the necessary scripts to reproducing the results presented in this paper in this link: https:/goo.gl/pUYqvi.

## V. Results and Discussion

### A. RQ1: Can CT increase the problem solving skills of students in high school?

To answer the first research question we analyzed the results of the WASI exam applied to 36 students in high school (first phase of the study). These results are illustrated in Figure 1 and Figure 2.

The performance of a student in the WASI exam is given by the number of correct answers he achieved. According to Figure 1, the frequency (density) of students with performance above average (represented by the dotted line on the graphic) is higher in the experimental group than in the control group.


Fig. 1: Performance distribution in the WASI exam

This can also be noticed in Figure 2 where the median performance of the experimental group is higher than the control group, suggesting a possible correlation between CT factor and the performance of students in the WASI exam.

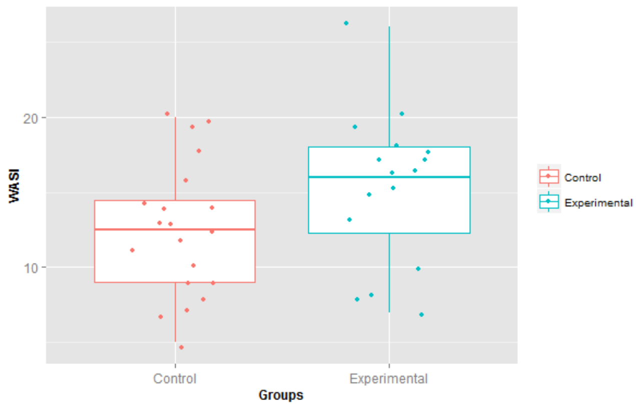


Fig. 2: Boxplot with scores WASI by group

The average difference between the performance of the groups in the WASI exam is shown in Table V. We found that the experimental group had an average performance of 21% better than the control group.

TABLE V. Mean difference between groups in the WASI test

| Groups | WASI Performance | SD |
|---|---|---|
| Experimental (E) | 15.19 | 5.0 |
| Control (C) | 12.55 | 4.4 |
| (E-C)% | 21.5% | 13.6% |

To analyze the effect size of the experimental group over the control group we calculated the Cohen's $d$ [23]. The result was $d$ = 0.56 which means a medium effect [23]. Besides, knowing that the effect size d is exactly equivalent to Z-score of a standard normal distribution [23], we can conclude that 71.2% of students in the experimental group have a higher average performance than the control group.

Considering the results presented above we had evidence to reject the null hypothesis H1.0: There is no statistical evidence that CT can increase the student's problem-solving skills. However, to support decision making, we also executed a hypothesis test Student's t-test.

TABLE VI. $p - value$ of hypothesis tests

| | Shapiro-Wilk | Levene | t-test |
|---|---|---|---|
| $p - value$ | 0.42 | 0.80 | 0.05 |

For ensuring the validity of the t-test, we also evaluated the statistical assumptions of normality and homoscedasticity of the data. Using the Shapiro-Wilk test with significance level of $\alpha$ = 0.05 we evaluated whether the sample came from a normally distributed population or not.

The results for this test are presented in Table VI; we obtained a $p - value > \alpha$, then it is possible to assume, with a confidence level of 95%, that the sample came from a population with normal distribution. The homoscedasticity of the data was performed using the Levene test, considering the significance level $\alpha$ = 0.05. The results of this test are presented in Table VI; we obtained a $p - value > \alpha$, then it is possible to assume, with a confidence level of 95%, homogeneity of the sample variance.

We conducted the Student's t-test to analyze if the difference of means between the experimental and control groups is statistically significant; the test was performed considering a significance level of $\alpha$ = 0.05. As shown in Table VI, we obtained $p - value = \alpha$. Thus, we should accept the null hypothesis H1.0: There is no statistical evidence that CT can increase the students problem solving skills.

Though we do not reject the null hypothesis H1:0 we consider the effect of CT in the experimental group was positive from the aspect of problem solving, since the effect size calculated Cohen's $d$ was average and the percentage difference between the performance groups was 21%.

### B. *RQ2: Is CT related to the performance of students in high school?*

To answer the second research question we analyzed data collected from 113 freshman students who did ENEM (second phase of the study), and additionally conducted a test to evaluate the programming skills of these students. We believe that the performance of students in ENEM is a good indication of the students' performance during basic education, just like skills in computer programming stimulate the core of CT.

Analyzing graph of the overall performance distribution in ENEM (see Figure 3), we observed that the frequency of students above the average performance is higher in the experimental group than in the group control. This represents an evidence that there is some relation between CT and the performance in high school. However, as students came from different types of school, possibly providing different quality education, we also analyzed such a relation considering the type of school as a factor (see Figure 4).

We observed that, regardless the type of school, students in the experimental group outperformed those in the control

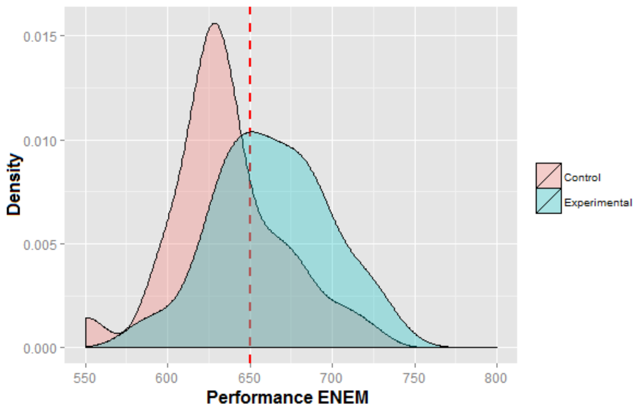


Fig. 3: Distribution of overall performance in the ENEM by groups

group in ENEM. According to the graphic analysis there is evidence of the relation between CT (by means of computer programming) and the performance of students on high school (ENEM). However, to refute the null hypothesis H2.0: There is no evidence of correlation between the students' performance on high school and skills in CT, we carried out hypothesis tests considering a contingency table (see Table VII) containing the variables performance in ENEM and performance in the programming test. By convention, we considered any value above the median of the sample as high score and, any value below the median of sample as low score.

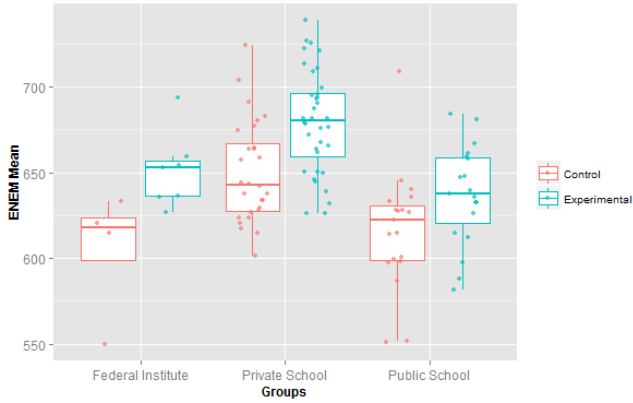


Fig. 4: Overall performance in the ENEM by type of school and groups

TABLE VII. Contingency table 2x2

| | Performance on programming test | |
|---|---|---|
| ENEM Performance | Above the median (+) | Below the median (-) |
| Above the median (+) | 43 (38%) | 14 (12%) |
| Below the median (-) | 19 (17%) | 37 (33%) |

By analyzing Table VII we found evidence of a moderate correlation, as the largest number of scores occurred when both variables are above the median (+, +), 43, and below the median (-, -), 37. To support this evidence we conducted the Chi-Square test using Table VII to test the hypothesis of independence between variables CT and performance in ENEM. As the $p - value$ = 2.192e-05 resulting from Chi-Square test is less than the significance level $\alpha$ = 0.05, then we rejected the null hypothesis of independence and assumed that there is some level of dependence between CT skills and the overall performance of students in ENEM.

This means that, with a confidence level of 95% we rejected the null hypothesis H2.0: There is no evidence of correlation between the students' performance in high school and skills in CT; and we accept the alternative hypothesis H2.1: There is a correlation between the students? performance in high school and skills CT.

Finally, the Pearson correlation test was performed to measure the strength of the correlation between the performance of students in ENEM and in the programming test. In Table VIII is shown the Pearson correlation coefficient $r$ for each area addressed in the ENEM, which are: Mathematics and their Technologies ($Math$), Languages Codes and their Technologies ($Languages$), Humanities and its Technologies ($Human$), Natural Sciences and their Technologies ($Nature$) and Essay Writing ($Text$). The last column ($General$) represents student's overall performance in ENEM.

By analyzing each coefficient $r$ in Table VIII it is evident, the existence of a positive moderate correlation between performance on the programming test and the overall performance in ENEM, as well as for the Math component, since in these cases $0.30 < r < 0.70$. Regarding the other components we observed a weak correlation with the results of programming test as $0.10 < r < 0.30$.

TABLE VIII. Pearson's correlation coefficient

| | Math | Languages | Human | Nature | Text | General |
|---|---|---|---|---|---|---|
| Programming test | 0.44 | 0.12 | 0.06 | 0.25 | 0.07 | 0.34 |

According to this analysis it is clear that there is a positive relation between the CT and the performance of students in high school. Such a relation may be due to the ability of the students in the experimental group on problem solving, as this is a criterion hat comprises the basis of the ENEM assessment matrix I.

### C. Qualitative relations between ENEM and CT

CT and ENEM are strongly related to problem solving. In ENEM, problem solving is evaluated in all disciplines. The relations between the concepts of CT presented in [5] and the skills evaluated in ENEM (see Table I in background section) are described below:

Dominate Languages (DL): CT can be support the development of competence as its core makes use of mathematical language and science;

Understanding Phenomena (UP): This process is related to obtaining information from the observation of a certain environment, so that can clearly involves the concepts of collection, analysis and representation of data, and may also involve the concept of simulation;

Confront Problem Situations (CP): primarily involves the abstraction of concepts for the understanding of the problem; collection, analysis and representation of data on the problem; decomposition, reducing the complexity of the problem and systemic algorithms and procedures for the construction of a solution;

Construct Argumentation (CA): The collection criteria, analysis and representation of data are related to the construction of argument, given the need to support arguments (collection and analysis of information) from a text, images or graphics (data representation);

Elaborating Proposals (EP): Prepare a proposal for intervention or solution to a problem involves Dominate Languages (DL), Understanding Phenomena (UP), Confront Problem Situations (CP) and Construct Argumentation (CA), so it can cover at least 5 of the 9 criteria presented in [5].

Elaborating Proposals (EP): Prepare a proposal for intervention or solution to a problem involves Dominate Languages (DL), Understanding Phenomena (UP), Confront Problem Situations (CP) and Construct Argumentation (CA), so it can cover at least 5 of the 9 criteria presented in [5].

By analyzing these skills we realize that concepts related to CT are present implicitly in Essay Writer, once it deals with developing a proposal for intervention for a particular problem.

### *D. RQ3: Can CT improve the performance of students in high school?*

To answer the third research question we conducted a Student's t-test with the performance of experimental and control groups in ENEM. We consider as an alternative hypothesis test that the mean of the experimental group are higher than the average of the control group. To ensure the statistical accuracy of the analysis of the t-test assumptions (normality and homoscedasticity) were evaluated. All tests were performed with a confidence level of 95%.

As the experimental and control groups have a different number of students in schools public, federal and private IV, we carried out the balancing of the groups using a method of resampling pseudorandom with replacement based on the bootstrap method Efron [22]. So we submit to the t-test groups of students with same proportion of students in public school, federal school and private school.

To verify the null hypothesis of normality, collected data variance homogeneity was performed respectively the Shapiro-Wilk and Levene tests. Table IX shows the $p - value$ of the tests for each set of ENEM averages.

Analyzing the results of the Shapiro-Wilk test and Levene in Table IX, for a significance level $\alpha$ = 0.05, we find that we can not reject the hypothesis of normality and homogeneity of variances, once have obtained a $p - value > \alpha$ in both tests to sample.

TABLE IX. Hypothesis tests

| Test | Math | Languages | Human | Nature | Text | General |
|---|---|---|---|---|---|---|
| Shapiro-Wilk | 0.35 | 0.64 | 0.50 | 0.54 | 0.63 | 0.22 |
| Levene | 0.55 | 0.40 | 0.17 | 0.52 | 0.83 | 0.46 |
| Student's t | <0.01 | 0.10 | 0.15 | 0.01 | 0.08 | <0.01 |

The t-test $p - value$ is observed in Table IX and since $p - value < \alpha$ = 0.05 for the components of Mathematics, Natural Sciences and the general performance on ENEM. Then, reject the null hypothesis H3.0: There is no statistical evidence that CT can improve students' performance in high school. Thus, we assume the alternative hypothesis H3.1: There is statistical evidence that CT can improve students? performance in high school.

We did not observe significant differences in the performance of the experimental and control groups in relation to the curriculum components of Languages, Humanities and Text (Essay Writing), once for those cases $p - value > \alpha$.

The mean difference between groups can be observed in detail in the table X.

TABLE X. Mean difference and percentage between groups

| Groups | Math | Languages | Human | Nature | Text | General |
|---|---|---|---|---|---|---|
| Experimental (E) | 711 | 603 | 648 | 621 | 725 | 662 |
| Control (C) | 645 | 592 | 639 | 599 | 694 | 634 |
| E - C | 66 | 11 | 9 | 22 | 31 | 28 |
| (E - C)% | 10.2% | 1.9% | 1.4% | 3.7% | 4.5% | 4.4% |

Based on data from the table X, it is found that the average percentage difference between the groups ranges from 1.4% to 10.2%. In terms of score means, students with proficiency in CT practices increased by 4 to 66 points from the group of students who did not have any level of knowledge on the subject.

The analysis of the effect size of the experimental group on control group was measured through Cohen's $d$ [23] for all areas evaluated in ENEM. The result of the analysis is presented in Table XI. According to the assessment metric [23] general performance in ENEM had a medium effect, the mathematics component showed a high effect, while the components such as Languages, Humanities, Natural Sciences and Essay Writing presented a small effect.

TABLE XI. Effect size

| Test | Math | Languages | Human | Nature | Text | General |
|---|---|---|---|---|---|---|
| Cohen's d | 0.86 | 0.27 | 0.21 | 0.46 | 0.28 | 0.78 |
| Effect size | Large | Small | Small | Small | Small | Medium |
| Exceeds scores % | 80.5% | 60.6% | 58.3% | 67.7% | 61% | 78.2% |

By the Table XI it is possible to observe that the percentage of subjects in the experimental group have a higher average than the control group ranges from 58.3% to 80.5% depending on the area evaluated in ENEM.

### *E. Threats of validity*

*Constructo validity:* The forms for data collection used and tests have been validated by experts.

*Internal validity:* This study analyzed a random sample of the student population, in addition, it was possible to isolate spurious factors of variation that influence student performance, it sought to decrease the bias and ensure the internal validity.

*Conclusion validity:* The small number of samples increases the probability of Type II error (accepting the null hypothesis when it is false). However, we consider all the assumptions of the statistical tests used in this work, as well as use homogeneous samples in the data analysis.

*External validity:* As the number of samples used in our study is not representative of the student population, we can not generalize the results.

## VI. Conclusions

Starting from the need to acquire problem solving skills even during the basic education, and considering inherent difficulty in building these skills, the aim of our study was to analyze the effect of CT in problem solving skills, as well as in students' performance in high school.

We conducted our analysis with a sample of high school students classified under experimental and control groups, where the CT factor distinguishes both groups. From the analysis, in the context of basic education, we found that students trained in CT topics (experimental group) have better performance in problem solving, Mathematics and Natural Sciences as compared to students who did not develop CT skills (control group). These results support the hypothesis sustained in qualitative studies of literature that claim CT favors students' skills in different knowledge areas.

However, it was not possible to identify significant effects of CT in some areas such as Humanities, Languages and Essay Writing. This may be related to the maturity level of the students in the use of CT skills in subjective areas. Therefore, more studies about the benefits of CT in these areas are needed.

Our results suggest that CT can support the improvement of students' skills on basic education. However, given the small number of samples it is not possible to generalize the results of this work to other students who were not part of this research.

As a future work, we plan to conduct new experiments with representative samples of students in order to further clarify the benefits related to this area of knowledge.

## References


[1] UNESCO, "Teaching and learning: Achieving quality for all. efa global monitoring report." Tech. Rep., 2014, access: March 20, 2016. [Online]. Available: http://unesdoc.unesco.org/images/0022/002256/225660e.pdf

[2] O. PISA, "Results in focus: What 15-year-olds know and what they can do with what they know," 2012.

[3] J. M. Wing, "Computational thinking," *Commun. ACM*, vol. 49, no. 3, pp. 33–35, Mar. 2006. [Online]. Available: http://doi.acm.org/10.1145/1118178.1118215

[4] M. Israel, J. N. Pearson, T. Tapia, Q. M. Wherfel, and G. Reese, "Supporting all learners in school-wide computational thinking: Aăcross-case qualitative analysis," *Computers & Education*, vol. 82, pp. 263 – 279, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360131514002759

[5] V. Barr and C. Stephenson, "Bringing computational thinking to k-12: What is involved and what is the role of the computer science education community?" *ACM Inroads*, vol. 2, no. 1, pp. 48–54, Feb. 2011. [Online]. Available: http://doi.acm.org/10.1145/1929887.1929905

[6] S. Grover and R. Pea, "Computational thinking in k–12 a review of the state of the field," *Educational Researcher*, vol. 42, no. 1, pp. 38–43, 2013.

[7] O. L. Oliveira, M. C. Nicoletti, and L. M. del Val Cura, "Quantitative correlation between ability to compute and student performance in a primary school," in *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, ser. SIGCSE '14. New York, NY, USA: ACM, 2014, pp. 505–510. [Online]. Available: http://doi.acm.org/10.1145/2538862.2538890

[8] K. H. Koh, H. Nickerson, A. Basawapatna, and A. Repenning, "Early validation of computational thinking pattern analysis," in *Proceedings of the 2014 Conference on Innovation & Technology in Computer Science Education*, ser. ITiCSE '14. New York, NY, USA: ACM, 2014, pp. 213–218. [Online]. Available: http://doi.acm.org/10.1145/2591708.2591724

[9] L. Cohen, L. Manion, and K. Morrison, *Research Methods in Education*, ser. Education, Research methods. Routledge, 2011. [Online]. Available: http://books.google.com.br/books?id=p7oifuW1A6gC

[10] A. Whimbey, J. Lochhead, and R. Narode, *Problem solving & comprehension*. Routledge, 2013.

[11] INEP, "Sobre o enem (2015). avaliable: http://portal.inep.gov.br/web/enem/sobre-o-enem. acesso em: 12 fev 2015." 2015.

[12] S. Papert, *Mindstorms: Children, Computers, and Powerful Ideas*. New York, NY, USA: Basic Books, Inc., 1980.

[13] C. Hu, "Computational thinking: What it might mean and what we might do about it," in *Proceedings of the 16th Annual Joint Conference on Innovation and Technology in Computer Science Education*, ser. ITiCSE '11. New York, NY, USA: ACM, 2011, pp. 223–227. [Online]. Available: http://doi.acm.org/10.1145/1999747.1999811

[14] J. M. Wing, "Computational thinking and thinking about computing," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 366, no. 1881, pp. 3717–3725, 2008.

[15] M. Resnick, J. Maloney, A. Monroy-Hernández, N. Rusk, E. Eastmond, K. Brennan, A. Millner, E. Rosenbaum, J. Silver, B. Silverman, and Y. Kafai, "Scratch: Programming for all," *Commun. ACM*, vol. 52, no. 11, pp. 60–67, Nov. 2009. [Online]. Available: http://doi.acm.org/10.1145/1592761.1592779

[16] M. Van Dyne and J. Braun, "Effectiveness of a computational thinking (cs0) course on student analytical skills," in *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, ser. SIGCSE '14. New York, NY, USA: ACM, 2014, pp. 133–138. [Online]. Available: http://doi.acm.org/10.1145/2538862.2538956

[17] S. Y. Lye and J. H. L. Koh, "Review on teaching and learning of computational thinking through programming: What is next for k-12?" *Computers in Human Behavior*, vol. 41, pp. 51–61, 2014.

[18] Brasil, *Lei de diretrizes e bases da educação nacional*. Conselho de Reitores das Universidades Brasileiras, 1996.

[19] M. BRASIL, "Sef. parâmetros curriculares nacionais: terceiro e quarto ciclos do ensino fundamental: Introdução aos pcn," *Brasília: MEC/Secretaria de Educação Fundamental*, 1998.

[20] M. SEMTEC, "National curriculum parameters," *Brasília: Secretaria de Educação Média e Tecnologia*, 1999.

[21] M. Van Dyne and J. Braun, "Effectiveness of a computational thinking (cs0) course on student analytical skills," in *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, ser. SIGCSE '14. New York, NY, USA: ACM, 2014, pp. 133–138. [Online]. Available: http://doi.acm.org/10.1145/2538862.2538956

[22] B. Efron and R. J. Tibshirani, *An introduction to the bootstrap*. CRC press, 1994.

[23] J. Cohen, "A power primer." *Psychological bulletin*, vol. 112, no. 1, p. 155, 1992.