REVISTA ELECTRÓNICA DE INVESTIGACIÓN
EN EDUCACIÓN EN CIENCIAS
ISSN 1850-6666

CTRM: A SKILL-BASED COMPUTATIONAL THINKING ASSESSMENT

Walkiria Cordenonzi[1], Cesar Costa Machado1[1], Vanessa Mattoso Cardoso[1]

walkiriacordenonzi@ifsul.edu.br, cesarmachado@ifsul.edu.br, vanessacardoso@ifsul.edu.br

[1]IFSUL, Instituto Federal de Educação, Ciência e Tecnología Sul-rio-grandense, Paul Harris 410, Santana do Livramento, RS,Brasil .

Abstract

Computational Thinking (CT) is a set of skills, understood as know-how, that the individual can develop in order to propose solutions to challenges related not only to computing problems, but also to everyday life. Possible solutions go beyond learning about programming. The Computational Thinking Reference Model (CTRM), based on skills, presents an evaluation method, in addition to instantiating a construct for the validation of both. The objective of this research is to analyze the correlation of the CT skills proposed in the CTRM, in order to answer the hypothesis: how and how much these skills are related to each other, through the Pearson coefficient and determination coefficient. Five case studies were performed with students from a higher education course, the assessment being applied objectively, based on tests. As a result of this process, it was possible to determine that all the skills analyzed present strong positive relationships, showing that the implication between the skills resulted in a value above 50%.

Keywords: Computational Thinking Ability. Assessment of Computational Thinking. Problem solving. Pearson's coefficient. Determination coefficient.

MRPC: EVALUACIÓN DEL PENSAMIENTO COMPUTACIONAL BASADA EN HABILIDADES

Resumen

El pensamiento computacional (PC) es un conjunto de habilidades, entendidas como saber hacer, que el individuo puede desarrollar para proponer soluciones a desafíos relacionados no solo con problemas informáticos, sino también con la vida cotidiana. Las posibles soluciones van más allá de aprender a programar. El Modelo de Referencia de Pensamiento Computacional basado en Habilidades (MRPC) presenta un método de evaluación, además de instanciar un constructo para la validación de ambos. El objetivo de esta investigación es analizar la correlación de las habilidades propuestas en el MRPC, con el fin de dar respuesta a la hipótesis: cómo y cuánto se relacionan estas habilidades entre sí, a través del coeficiente de Pearson y el coeficiente de Determinación. Se realizaron cinco estudios de caso con estudiantes de un curso de educación superior, y la evaluación se realizó de manera objetiva, en base a pruebas. Como resultado de este proceso, se pudo determinar que todas las habilidades analizadas presentan fuertes relaciones positivas, mostrando que la implicación entre las habilidades está por encima del 50%.

Palabras clave: Habilidade de Pensamiento Computacional. Evaluación del Pensamiento Computacional. Solución de problemas. Coeficiente de Pearson. Coeficiente de determinación.

MRPC: UMA AVALIAÇÃO DE PENSAMENTO COMPUTACIONAL BASEADA EM HABILIDADES

Resumo

O pensamento computacional (CT) é um conjunto de habilidades, entendido como o saber-fazer, que o indivíduo pode desenvolver com o objetivo de propor soluções aos desafios relacionados não só a problemas de computação, mas também à vida cotidiana. As possíveis soluções vão além da aprendizagem relativa à programação. O Modelo de Referência do Pensamento Computacional (MRPC), baseado em habilidades, apresenta um método de avaliação, além de instanciar um constructo para a validação de ambos. O objetivo desta pesquisa é analisar a correlação das habilidades do CT propostas no MRPC, a fim de responder à hipótese: como e quanto estas habilidades estão relacionadas entre si, através do coeficiente de Pearson e do coeficiente de

Recepción: 14/01/2025 *Aceptación: 21/04/2025*

Determinação. Foram realizados cinco estudos de caso com alunos de um curso superior, sendo a avaliação feita de forma objetiva, baseada em testes. Como resultado deste processo, foi possível determinar que todas as habilidades analisadas apresentam relações positivas fortes, evidenciando que a implicação entre as habilidades é acima de 50%.

Palavras chave: Habilidade de Pensamento Computacional. Avaliação do Pensamento Computacional. Resolução de problemas. Coeficiente de Pearson. Coeficiente de Determinação.

CTRM : UNE ÉVALUATION DE LA PENSÉE INFORMATIQUE BASÉE SUR LES COMPÉTENCES

Résumé

La pensée computationnelle (CT) est un ensemble de compétences, entendues comme un savoir-faire, que l'individu peut développer afin de proposer des solutions à des défis liés non seulement aux problèmes informatiques, mais aussi à la vie quotidienne. Les solutions possibles vont au-delà de l'apprentissage de la programmation. Le modèle de référence de pensée informatique (CTRM), basé sur les compétences, présente une méthode d'évaluation, en plus d'instancier un construit pour la validation des deux. L'objectif de cette recherche est d'analyser la corrélation des compétences CT proposées dans le CTRM, afin de répondre à l'hypothèse : comment et dans quelle mesure ces compétences sont liées entre elles, à travers le coefficient de Pearson et le coefficient de détermination. Cinq études de cas ont été réalisées auprès d'étudiants d'une formation de l'enseignement supérieur, l'évaluation étant appliquée de manière objective, à partir de tests. Grâce à ce processus, il a été possible de déterminer que toutes les compétences analysées présentent de fortes relations positives, montrant que l'implication entre les compétences aboutissait à une valeur supérieure à 50 %.

Mots clés: Capacité de réflexion informatique. Évaluation de la pensée informatique. Résolution de problème. Coefficient de Pearson. Coefficient de détermination

1. INTRODUCTION

In the 1980s, Papert (1980a) defined the LOGO language, whose specific commands children used to program a turtle's movements. According to the author, "the intelligent child teaches the dumb computer, instead of the intelligent computer teaching the dumb child" (Papert, 1980b, p. 9).

In 2006, Wing (2006) resumes the proposal for problem solving, algorithms and programming, describing it as computational thinking (CT), but in a broader way, without tying it to age groups and aggregating methods used in computer science. In addition to suggesting that the CT is a set of skills to be developed by any professional, regardless of their area of expertise (Wing, 2008).

From Wing's proposal in 2006, measures were taken by professionals and governments to incorporate computational thinking (CT) into school curricula. It is believed that it would be very important, teaching CT in formal education, from the early years, so that students begin to master the corresponding skills and, consequently, problem solving.

In Brazil, on February 18, 2022, the National Education Council (NEC) approved the "Norms on Computing in Basic Education – Complementary to the BNCC", which means that the fundamentals and computing technologies will become part of the common national basis curriculum of Brazilian schools, as soon as it is sanctioned by the Executive. Until then, research in the area presents itself as courses (workshops), in which they develop activities in this sense, some focusing on programming, others on unplugged tasks (Brackmann, 2017).

The CT can be developed in different ways, including programming. However, developing the CT through programming does not mean just programming, but developing the entire process to find the solution to a problem. In this sense, different skills are proposed (CRTC, 2018; ISTE & CSTA, 2011; Wing, 2008). The Computational Thinking Reference Model (CTRM), defined by Cordenonzi (2020), proposes five skills for CT development. From case studies carried out, this research aims to analyze the correlation of CT skills proposed in this model, answering the following hypothesis: how and how much these skills are related to each other.

This article is organized as follows: in section 2, the theoretical framework is presented, plus related works on the assessment of CT; in section 3, the CTRM is briefly presented. In section 4, the methodology adopted in this research is explained, followed by the results found, in section 5. The final considerations are presented in section 6 and, subsequently, in section 7, the bibliographic references.

2. THEORETICAL FRAMEWORK

Understanding computational thinking (CT) is not an easy task, since 16 years after the definition of this expression by Wing (2006), researchers have not reached a consensus. It is possible to infer some approximations about understanding, such as, for example, being a critical skill, collaborating in problem solving, in logic and abstraction, among others, in addition to being a necessary skill for 21st century citizens.

As already mentioned, Wing (2006, p. 33) defines CT as a skill that "solving problems, designing systems, and understanding human decomposition when attacking a large complex task behavior, by drawing on the concepts fundamental or designing a large complex system. It is separation to computer science". In addition, the author adds that it is a form of recursive thinking, which uses abstraction and decomposition, that is, it makes use of heuristic reasoning to find the solution. Furthermore, the author compares the CT with the ability to read, write and calculate, emphasizing the ability to solve problems, regardless of computing resources,

Recepción: 14/01/2025 Aceptación: 21/04/2025

age or profession of the subjects. In 2008, Wing (2008) adds that the essence of the CT is abstraction to solve problems, adding the ability to analyze solutions.
However, for the Google for Education team (2015), CT is a perspective used for problem solving anchored in the knowledge of Computing, pointing out that the CT is "essential for the development of computer applications, but it can also be used to support problem solving in all disciplines, including math, science and humanities" (Google for Education, 2015, digital text).
The International Society for Technology in Education (2011, p. 5) defines CT as "develop and employ strategies for understanding and solving problems in ways that leverage the power of technological methods develop and test solutions"
The Reference Curriculum in Technology and Computing (CRTC, 2018) defined by the Center for Innovation for Brazilian Education (CIBE), and sent to the national council of education (NCE), defines the CT as the process that comprises "systematizing, representing, analyzing and solving problems" (RCTC, 2018, p. 19) and proposes 4 skills for the CT: abstraction, algorithm, decomposition and pattern recognition.
In contrast, with respect to CT skills, the Computer Science Teachers Association (CSTA) (ISTE & CSTA, 2011) suggests: abstraction, algorithmic thinking, modeling, scale, and pattern recognition.
Brennan and Resnick define the CT through a framework that contains three dimensions: "computational concepts (concepts are employed as they are programmed), computational practices (developed while they are programmed) and computational perspectives (perspectives form about the world around them and about themselves)" (Brennan & Resnick, 2012, p. 3).
As for the CTRM (Cordenonzi, 2020), the author suggests five skills: comprehension, abstraction, problem solving, algorithmic solving and validation (see more details in section 3).
In Table 1, a summary of some researchers about the CT and its abilities is presented. This table was built from the search for related words, in order to highlight the main skills and/or concepts found in the literature.

Table 1

Summary of CT concepts

Conceitos e/ou habilidades	(Wing, 2008)	(ISTE; CSTA, 2011)	(Brennan; Resnick, 2012)	(CRTC, 2018)	(GOOGLE FOR EDUCATION, 2015)	(Cordenonzi,2020)
Problem Solving	X	X	X	X	X	X
Abstraction	X	X	X	X		X
Test	X	X	X			X
Algorithm		X	X	X	X	
Data Processing		X	X		X	
Programming		X			X	X
Patterns		X		X		
Decomposition				X		X

It is important to point out that some authors call skills as pillars. As an example, we can mention the research by Ribeiro, Foss and Cavalheiro (2017), who mention the three pillars: abstraction, automation and analysis.
The analysis of Table 1 allows us to infer that the skill with the highest incidence is abstraction, which is also supported by research conducted by Grover and Pea (2013) and Cordenonzi et al. (2020). Next comes problem solving. It should be noted that the development of CT skills is not limited to programming, which can be understood in this context as the automation of individual abstractions (Wing, 2008), that is, a way to solve a problem.
To reinforce the most important skills, Avila et al. (2017) analyzed 58 articles and concluded that the most used are algorithmic thinking, problem solving and abstraction. These authors also emphasize the use of assessment instruments: qualitative, through observations; and quantitative, through pre- and post-tests. Furthermore, they mention that tools for evaluating code from visual programming environments and collaborative learning environments are widely used.
The following are some surveys that address the CT Assessment process.

2.1 CT Assessment (CTA)

This section presents some works found in the literature on CT assessment, using the keywords: assessment and computational thinking. However, it can be said that there is no standard or generic model for measuring CT. In this sense, it is urgent that evaluative methods be applied, in order to know the results of the CT development.
At the University of Amasya, Turkey, researchers Korkmaz, Çakir and Özden (2017) defined a scale that determines the levels of CT skills and applied it to 1306 subjects. This scale is composed of 29 items, which were distributed in the factors: creativity, algorithmic thinking, cooperation, critical thinking and problem solving. According to the results, the test is statistically valid and reliable; however, the authors did not describe the form of application, thus being difficult to reproduce.
Avila et al. (2017), in a study carried out on 58 articles published between 2011 and 2016, concluded that the CTA with its own evaluation accounted for 65% of the publications, without, however, discriminating which tests were applied or whether they were automated. To clarify, the

tests performed by automation are limited to the software used to develop the solution to the problem.

For Moreno León, Román González and Robles (2018), the most used methods to assess CT, under different aspects, are CT-Test, Bebras and Dr. Scratch. They add that these tests must be used together, as they are compatible and complementary.

The Computational Thinking Test (CT-Test) was proposed in the doctoral thesis of Román González (2016), in order to measure the level of CT development through a questionnaire composed of 28 objective questions, with options of 4 answers (only one correct). It concluded that CT is not only a cognitive problem-solving skill, it is also gender-independent. Dr. Scratch[1], on the other hand, is a web application that automatically analyzes projects developed in Scratch, however, it is limited to this environment and does not support abstraction of functions and procedures in code (Grover & Pea, 2013). As a result, the tool presents a return that indicates the quality of the program, being considered a formative evaluation tool (Román González et al., 2017).

The International Challenge on Informatics and Computational Thinking, called Bebras , is a challenge that consists of a set of questions, whose purpose is to test the level of development of CT skills (Dagiené & Futschek, 2008), among them: problem solving , decomposition, algorithm design, pattern recognition, generalization and abstraction. However, it is noteworthy that this test came to be used in competitions and not as a method of evaluating the CT.

Araújo (2019) proposed to investigate strategies and instruments to quantify the CT, without the use of programming practices. As a result, he developed a model based on empirical studies to quantify CT as a cognitive skill. This model is divided into four competencies and twelve skills. It used Bebras, and the subjects involved were students from some Brazilian universities.

Grover et al. (2017) set out to explore approaches based on hypotheses that can be combined with those based on data, in order to improve the interpretation of results, through log records. These records, stored in the block-based programming environments, were the source for measuring and evaluating the students' CT skills (subjects of the research). For the analysis of the results, they relied on the data of 229 students, whose log records and final programs were classified. They performed a chi-square statistical significance test for each sequence (with the null hypothesis that there is no difference between the groups), in terms of observation of the sequence in question. They concluded that there is a space that has not yet been researched on how programming tasks need to undergo more accurate assessments regarding the understanding of individual concepts and subjects' skill levels.

Recently, Lai (2022) carried out three case studies with 119 students in the UK. To understand the data, the Rasch Model was used and, to evaluate, the in-fit and out-fit mean square statistics of the Mean Square Statistics (MNSQ) were computed for each item. It presented descriptive statistics of the tests performed and a correlation matrix (Pearson's coefficient), to deal with programming and problem solving variables. The relationships between all variables were significant and strong, resulting in Pearson's coefficient in programming of r=0.40 and in problem solving of r=0.45. That is, a strong correlation of the skills assessed.

In summary, it can be seen that the field of CTA is still devoid of research. According to Raabe et al. (2017), most of the proposed methodologies are of the qualitative type and that there are not many evaluation tools that produce immediate feedback, in addition to being restricted to the development environment used. In their research, Tang et al. (2020), based on 77 articles analyzed, prior to August 2019, conclude that there are still few publications in the CT area and suggest the integration of several tools to improve the assessment of CT learning.

Next, the CTRM is briefly presented, for the understanding of its abilities. The description of the model and the evaluation instruments used are detailed in AUTOR (ANO).

3. COMPUTATIONAL THINKING REFERENCE MODEL (CTRM)

The Computational Thinking Reference Model (CTRM) was proposed in AUTOR's doctoral thesis (ANO). Its purpose was to evaluate the development of the CT, with programming being the main focus. By defining a skill set, the model can be used to classify subjects into: code ability[2] (ACod), code literate (LCod)[3] and unplugged computational thinker.

For this model, an assessment method was proposed that collects multidimensional evidence on CT skills. Added to this model, a construct of classes to be developed, to guide the replication of the CTRM. For this construct, APP Inventor was used, but it is up to the researcher to use any programming language.

In Figure 1, the reference model for CT development (CTRM) is represented graphically.

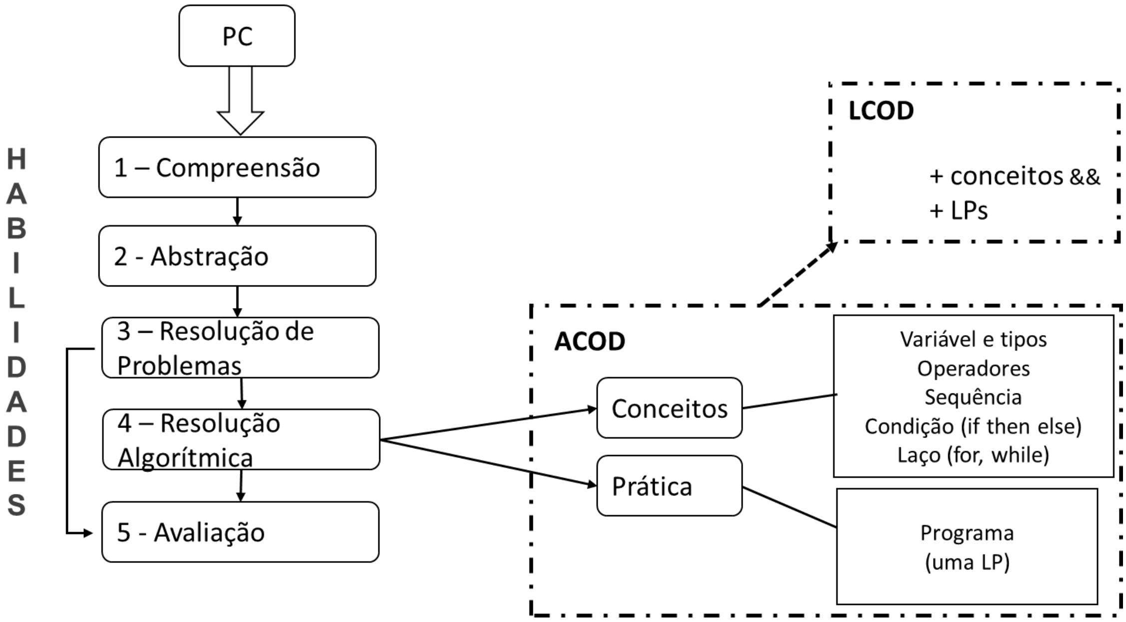
Figure 1

[1] Avaliable in: http://www.drscratch.org/. Acesso em: 16 maio 2022.

[2] Abitity as a synonym of proficiency, adapted from the original. A Code Literate (Acod) individual is one capable of reading, interpreting and write a source code and generate an executable code (program)(Cordenonzi et al., 2020, p. 150).

[3] A Code Literary (LCod) individual is an ACod extrapolating his skills and competencies to code in different programming languages (Cordenonzi et al., 2020, p. 150).

CTRM



Following Figure 1, one can see the existence of a logical dependency, suggested and not mandatory, to develop the CT. That is, from a problem, the first skill to be mobilized is understanding, because, without understanding the problem, the subject will not be able to solve it. Here, it is understood that a problem can be proposed by someone or perceived by the subject himself, as computational thinking is not restricted to solving a problem, but also refers to situations that require a solution.

Therefore, understanding involves significant learning, in which the subject builds his knowledge, and it is up to his cognitive processes to lead to the successful resolution of the problem. In the first step, from reading, the subject needs to create a mental representation of the problem, being as objective as possible. That means, reading and understanding what was read, using their previous knowledge (subsumers), because, for Ausubel (2003, p. 7), it is the "determining factor of the learning process". In short, 'understanding' is a skill proposed in order to perceive the subject's understanding of solving a problem, which is one of the critical factors for the development of CT.

The second skill is 'abstraction', which, in the Reference Curriculum in Technology and Computing (CTRC), was defined as involving the "[...] filtering of data and its classification, ignoring elements that are not necessary, aiming at the that are relevant. It also involves ways of organizing information into structures that can help solve problems" (CRTC, 2018, p. 19). In 2006, Wing (2006) defined 'abstraction' as the process of deciding which details to highlight and which to discard. Eleven years later, the same author adds that abstraction is the highest level process when thinking about CT. At this point, there is a disagreement, as it is believed that, for the subject to reach the abstraction process, he must necessarily have understood the problem, with abstraction being the next step. From the abstraction, one can determine the important points and discard the data that are not relevant, thus advancing to the next step, which is problem solving. Abstraction is important, but difficult to develop, especially when it comes to problem solving, as confirmed by França and Tedesco (2015).

The 'problem solving' skill focuses on the subject's ability to find one or several solutions, regardless of the format of his answer. Therefore, it can encompass problem decomposition, understood as "the process by which problems are divided into smaller and easier to solve parts" (CRTC, 2018, p. 19), that is, breaking them into smaller parts and manageable (ISTE & CSTA, 2011).

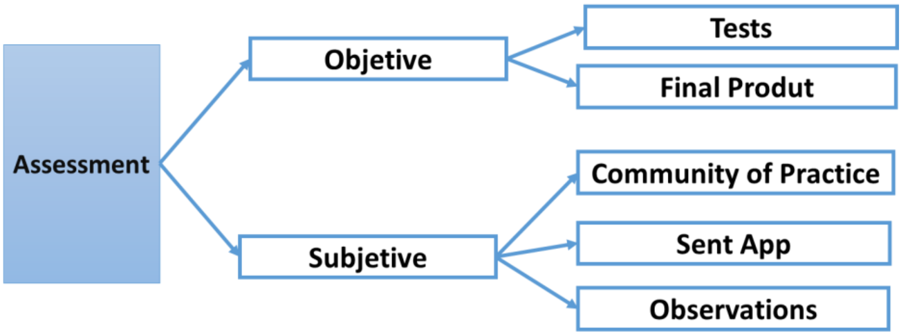
Understanding the skill of 'algorithmic resolution' is in the subject's practice, starting from the previous skill, in which he found the solution to the problem, translating it into an algorithm or software.

The 'assessment', the last skill proposed, is understood as the subject's ability, after presenting the resolution of the problem through an algorithm and/or program, to test and evaluate its correctness. In other words, it is understanding what has been implemented, that is, it is the ability to analyze a source code, executable in terms of its outputs. Evaluation is also understood to be the testing of a proposed solution to a problem presented, regardless of its format.

In order to operationalize the CTRM, an andragogical construct and an assessment method accompany the model (CORDENONZI, 2020). The classes' construct is the development of a course called "Eu Programa 1.0!", with a duration of 20 hours. At each meeting, students develop an App, totaling six, and at the end of the course, subjects develop an application for devices.

As evaluating is not a simple process, several instruments were proposed, which were grouped into two sets: objective and subjective instruments, as can be visualized graphically in Figure 2.

Figure 2
Types of assessment instruments



In this article, the focus is on the tests that make up the objective assessment and which are divided into a pre-test, an intermediate test and a post-test. The questions that make up the first two tests were taken from Bebras; therefore, they are already validated. For each question, the skill to which it is linked is already noted.

Based on the objective of this research to verify how the skills of the CTRM are correlated, that is, focusing only on the Tests, the other possibilities of evaluation of the CTRM will not be explained.
Table 2 specifies the number of questions for each test and, for each question, the skills that were assessed.

Table 1
Question skills

	Pre-test		*Intermediate test*		*Pos-test*				
Question	Q1	Q2	Q1	Q2	Q1	Q2	Q3	Q4	Q5
Skill's	**13**	**123**	**1235**	**145**	**12**	**123**	**12345**	**12345**	**145**

In the next section, the methodology adopted for data collection is detailed, in order to compose the objective evaluation of the tests.

4. METHODOLOGICAL PROCEDURES

Methodological procedures, or methodology, are understood as "the way of thinking and the practice exercised in approaching reality" (Minayo et al., 2002, p. 16). The path will require from the researcher three perspectives and discussions: the first concerns the theories considered; the second refers to the techniques used in development; the third and last concerns the interpretation of the results, that is, "the creative potential of the researcher" (p. 17).
This research was carried out following the inductive method, understood as one that "starts from the particular and places the generalization as a later product of the work of collecting particular data" (Gil, 2008, p. 10). This author proposes that the researcher start his work "from facts or phenomena whose causes he wants to know" (p.10), that is, from concrete facts with the intention of comparing them with each other. Furthermore, Lakatos and Marconi (2003, p. 92) confirm that the "inductive method aims to expand the reach of knowledge".
As for the approach to the problem, this research is quantitative, as they are centered on objectivity, with data analysis centered on numbers and their meanings. Gatti (2004, p. 13) confirms that these meanings "can be very useful in understanding various educational problems".
As for the objectives, it is based on the descriptive method. Descriptive research uses questionnaires and observations to collect data, although it does not exclude other types. Most research in the field of education is carried out in a descriptive way (Gil, 2008).
Still, in the field of technical procedures, the case study was used, which is a rigorous method used by researchers, whose focus is the planning and analysis of data (Yin, 2001). However, he emphasizes that, in order to increase reliability, the researcher must maintain the chain of evidence. The same author argues that, in terms of time, this method "shows that it is possible to carry out case studies in shorter periods and with results that can be confirmed by other studies" (Yin, 2001, p. 55).
From the tests performed - pre-test, intermediate and post-test, and considering that each question is associated with certain skills, for each subject it was verified whether the answer to the question was correct. In the case of a correct answer, one point was counted for the corresponding skills. Subsequently, a sum of the skills of each subject was performed and, subsequently, the sum of the skills by case study. In the end, data analysis was performed by calculating the Pearson coefficient and the coefficient of determination, in order to answer the research hypothesis: if they are related to each other, how are the skills proposed in the CTRM related?
Five case studies (CS) were carried out in the higher course Technologist in System Analysis, at the Federal Institute of Education, Science and Technology in Rio Grande do Sul (IFSUL), Campus Santana do Livramento, at different times, totaling 66 subjects involved, using the CTRM. The classes' construct was followed, with the three tests being presented to the subjects, with the same questions and the same period of 45 minutes for their application. The first CS was held in 2019 and the last one in the first quarter of 2022.
To assess skill 4, the App Inventor software was chosen. The option for this tool was because, from a simple, intuitive interface, the user is able to develop their applications in a short time. In other words, in the first meetings, the subject manages to develop a simple application for the Android operating system (which most users have). Remembering that the App is software developed for mobile devices, the assessment (skill 5) can be performed on the student's own device.

In time, the documents required by Resolution No. 466/2012 of the National Health Council, such as the Free and Informed Consent Term (FICT) and the Letter of Consent of the Institutions involved in the case studies were duly collected.
In the next section, the analysis of the correlations between the abilities proposed in the Reference Model for Computational Thinking (RMCT) is presented, in order to answer the research hypothesis.

5. RESULTS AND DISCUSSIONS

In order to analyze the data, we relied on the support of free software Past[4] and PSPP[5] . The first step to evaluate the data was the study and analysis in order to determine if the data obtained in the Case Studies meet a normal distribution, as this result determines the types of statistical tests that can be applied, complementing that there is a large number of these tests (parametric and non-parametric) and it must be considered that

[...] is the practical implication of a statistically significant difference. A significant difference is a difference that must not have occurred merely by chance, but is not necessarily a practically relevant difference. [...]. The practical analysis remains to verify whether these differences, which can be estimated from the data, are relevant (BARBETTA, 2002, p. 240).

Based on this understanding, the use of the non-parametric Kolmogorov-Smirnov (KS) test was determined for the sample. Statistical detail is described in the next section.

5.1. Correlation and Determination Coefficient

Table 2
Data on the subjects' abilities

Correlation, in statistics, evaluates the association between two qualitative variables (Reis, 2008). In other words, it is analyzing whether and how two or more variables relate to each other – correlation – either in the same or in opposite directions (Barbetta, 2002). The result of the correlation analysis provides a number that "[...] summarizes the degree of relationship between two variables [...]" (Stevenson, 2001, p. 341).. The author goes on to state that "in education and psychology, more emphasis is often placed on the degree or strength of the relationship". There is another estimation technique related to the correlation, which is regression, however, it was not used, since, according to the same author, its result returns "an equation that describes the relationship in mathematical terms" (p. 341), or that is, a prediction equation, widely used in Administration and Economics, among other areas.
In the context of this research, the correlations are related to the skills defined in each question proposed in the tests performed (pre-test, test 2 and post-test). Resuming, in the CTRM, five skills were defined, emphasizing that a question of any test can assess more than one skill.
After correcting each question, the sum of correct answers for each of the skills was performed for all subjects participating in the Case Studies. The result is summarized in Table 3, in which H represents the assessed skill, the next number corresponds to the skill number (according to the CTRM) and the "/number" informs the total number of questions that were assessed according to this skill. The subjects are organized in sequential order, by Case Study performed.

[4] Disponível em: https://folk.uio.no/ohammer/past/. Acesso em: 05 mai. 2022.

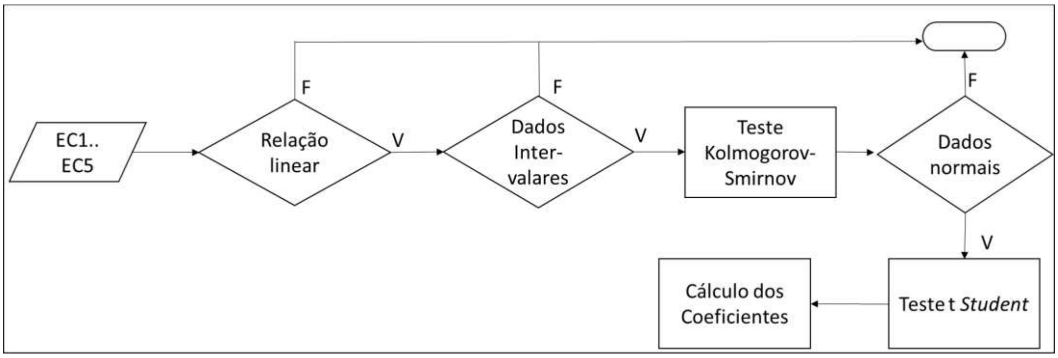
[5] Disponível em: https://sourceforge.net/projects/pspp4windows/. Acesso em: 05 mai. 2022.

Recepción: 14/01/2025 Aceptación: 21/04/2025

		H1/9	H2/6	H3/6	H4/4	H5/4
EC01	A01	5	4	3	1	1
	A02	3	2	1	0	0
	A03	5	4	4	1	1
	A04	6	3	4	1	1
	A05	6	4	4	1	1
	A06	1	2	1	0	0
EC02	A01	5	4	5	1	2
	A02	5	3	5	1	2
	A03	7	5	5	2	0
	A04	5	4	6	2	2
	A05	6	5	5	1	2
	A06	4	3	3	1	1
	A07	0	1	2	0	0
	A08	5	3	3	2	2
	A09	0	1	1	0	0
	A10	5	4	4	1	1
	A11	3	3	2	0	0
	A12	4	4	4	2	2
	A13	4	3	3	0	0
	A14	6	4	4	2	2
	A15	6	4	5	2	3
	A16	6	4	5	2	2
	A17	3	2	3	1	1
	A18	6	4	5	2	3
	A19	2	1	2	0	0
	A20	6	5	5	2	3
	A21	1	1	1	0	0
EC03	A01	6	3	4	1	4
	A02	6	4	4	3	2
	A03	4	2	4	2	2
	A04	7	5	6	3	4
	A05	5	3	4	2	2
	A06	4	2	3	1	2
EC03	A07	5	4	5	2	3
	A08	5	4	4	2	2
	A09	6	4	5	1	2
	A10	7	5	6	2	3
	A11	7	6	6	2	3
	A12	7	5	6	2	3
EC04	A01	3	2	2	0	0
	A02	4	4	4	2	2
	A03	4	4	4	2	3
	A04	6	4	6	3	4
	A05	7	5	6	3	4
	A06	6	4	5	3	3
	A07	5	4	5	3	4
	A08	6	3	5	4	5
	A09	7	5	6	4	4
	A10	8	6	7	4	5
	A11	7	5	5	3	3
	A12	3	3	2	1	1
	A13	8	6	7	3	4
EC05	A01	7	5	6	4	4
	A02	5	3	3	1	1
	A03	4	4	3	0	1
	A04	2	2	2	0	0
	A05	5	5	5	1	2
	A06	6	5	4	3	3
	A07	7	5	6	4	4
	A08	6	5	6	3	3
	A09	5	4	5	2	3
	A10	3	3	2	1	1
	A11	5	3	4	3	3
	A12	4	3	4	1	1
	A13	6	4	5	3	3
	A14	2	1	0	1	1

The most known and used correlation coefficient between variables is the “r of Pearson” (Barbetta, 2002), therefore chosen to analyze the research data. The use of this coefficient requires that the relationship between variables be linear, the data be measured at the interval level, the characteristics follow a normal distribution, and the sampling be random to allow for the application of the significance test. The steps performed can be seen in Figure 3.

Figure 3
Flowchart of correlations



To test the linearity of the data and for each correlation, a scatter diagram was constructed. These diagrams allow visualization of the strength, direction, and nature of the correlation. It is important to point out that, in all the graphs constructed, in the relationship between the variables, there was a strong increasing linear pattern.

As for the second step, explained by Anderson et al. (2019, p. 21): "interval data are numerical and expressed in terms of a fixed unit of measurement"; therefore, this requirement is covered. The next step was to verify if the sample is normally distributed and thus be able to apply the significance test.

For step 3, the Kolmogorov-Smirnov (KS) test was used, considering the data sample size, in this specific case, N=66 elements. Barbetta (2002) clarifies that a sample is considered large enough if it consists of 30 or more observations. The author also states that the mean of a random sample taken from a population of data will have an approximately normal frequency distribution, regardless of the population. The test results are shown in Table 4.

Table 3

Normality Test Result

Skill (H)	Value (p)
H1	0,11
H2	0,002
H3	0,23
H4	0,017
H5	0,105

Therefore, it is evident that all p(values) – see Table 4 – for skills H1, H3 and H5 are greater than α=0.05, that is, $p > 0.05$. Therefore, H0 is not rejected; thus, it can be said that the sample is normally distributed. For the other skills, the central limit theorem is considered

> If the variable of interest does not follow a normal distribution in the population (or it is not known what its distribution is), the sampling distribution of the means of random samples taken from this population will be normal if the size of these samples is sufficiently large. (REIS, 2008, p. 232).

Therefore, these skills will be treated equally.

Still, it is necessary to perform the last step, that is, to test the significance of the correlation measure, in which the following hypotheses are established:

H0: there is correlation between the variables
H1: there is no correlation between the variables.

Using the *t* test with bilateral distribution (the correlation can be positive or negative), and having the t_{critic} = 1.978, the results are shown in Table 5.

Table 4

Abilities t test

t statistical	H2	H3	H4	H5
H1	4,65	2,79	11,85	10,08
H2		-1,81	8,95	6,79
H3			9,62	7,75
H4				-1,54

According to rule P (T*estatistic* < = t*critical*), if it returns true, **H0** is rejected, that is, there is no significant difference. Illustrated in Table 5, the values that appear in red mean that there is a correlation between the variables. In contrast, there are Skills 3 and 5, which do not reject **H0**; therefore, there is a chance that these values are not significant, but nothing can be said with certainty.

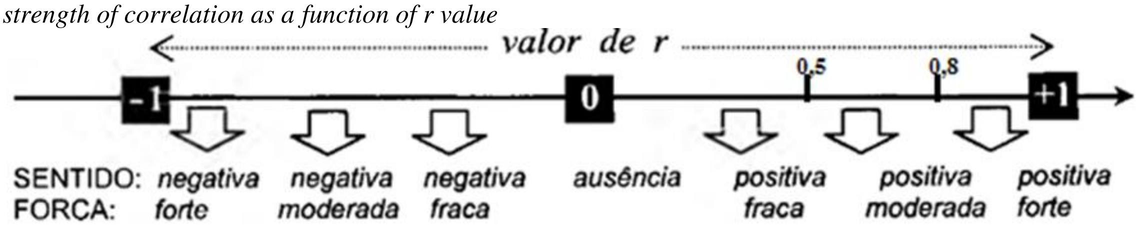
After all the steps performed, in the sequence, Table 6 shows the results of Pearson's coefficient r of all variables (skills) of the CTRM. To explain, a correlation was made between H1 and the other skills, H2 with the others, and so on.

Table 5
Pearson's coefficient

r by Pearson	**H2**	**H3**	**H4**	**H5**
H1	0,877	0,888	0,726	0,729
H2		0,858	0,626	0,627
H3			0,741	0,789
H4				0,869

The interpretation of Pearson's r value is illustrated in Figure 4.

Figure 4
Sense and strength of correlation as a function of r value



To complement the correlation analysis, the Determination coefficient, represented by R2, was applied. That is, a measure of the proportion of variability in one variable that is explained by the variability of the other, obtained by the square of Pearson's r coefficient (Anderson et al., 2019; Barbetta, 2002). The values of these coefficients were calculated and described in Table 7.

Table 6
Determination Coefffient

R^2	H2	H3	H4	H5
H1	77%	79%	53%	53%
H2		74%	39%	39%
H3			55%	62%
H4				76%

Based on the calculations of the coefficients presented so far and using the PSPP, we proceed to analyze them.

Regarding H1 (Understanding), the highest correlation value found was with H3 (Problem Solving), very similar to H2 (Abstraction). That is, r=0.888 and r=0.877 (Pearson's coefficient presented in Table 6) indicate a strong and positive correlation, as shown in Figure 4. Furthermore, 77% (Determination coefficient - Table 7) of the data are correlated, that is, only 23% of H1 cannot be explained by the variability of H3. In addition, in Graph 1, the scatter diagram between H1 and H3 is presented, which points to a positive relationship between the two variables. Interpreting these correlations in the context of CT means considering that the better the understanding (H1) of the problem, the greater the chance of correctly solving a problem (H3), that is, that the latter is related to one (or more) possible solution(s) and with the collection of important requirements for solving the problem, as both, from the point of view of the CTRM, are strongly related. It is important to emphasize that, between H1 and H2 (r=0.877) and H2 and H3 (r= 0.858), the values are close. Thus, it can be said that, from the understanding of a problem, the abstraction and the resolution of it are related. The correlations presented between the variables (or abilities) of the CTRM, as corroborated by Figure 4, are all positive, from moderate to strong strength.

As for R2, the variation of H1 is related to 79% of the variation of H3, which is the highest value. It is understood that understanding influenced the correct resolution of a problem (Graph 1). Skills H2 with H3 and H4 with H5 showed a close coefficient of variation, respectively at 74% and 76%.

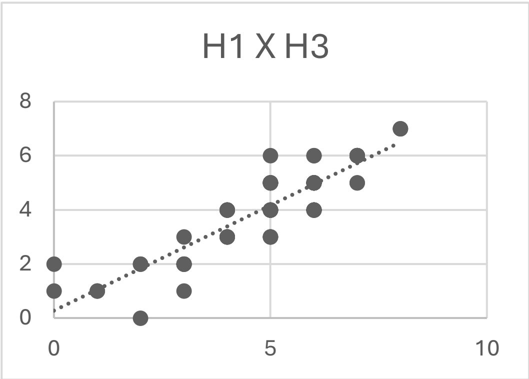
On the other hand, there was a weak correlation between H2 (abstraction) and H4 (Resolution through algorithms) and with H5 (Validation), as shown, respectively, by the values

Recepción: 14/01/2025 Aceptación: 21/04/2025

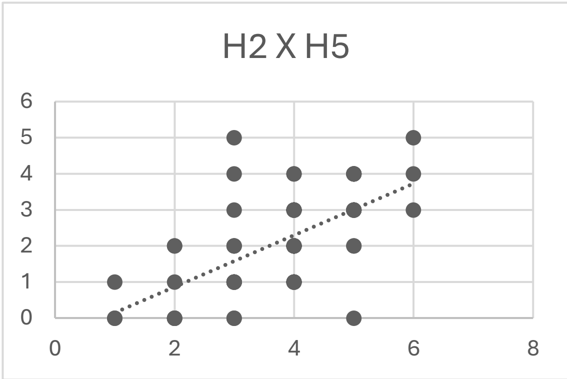
of r=0.626 and r=0.627. The lowest coefficient of determination was R2=39%, between H2 with H5 and H4. Therefore, it can be understood that abstraction is the skill that least influences the validation of data (tests) and the construction of algorithms. In order to validate this diagnosis, the scatter diagram is presented in Graph 2.

For skills 2 and 5, the relationship between the two is considered weak, as there is a greater dispersion of them (Graph 2).

Graph 1
Scatter Diagram between H1 and H3



Graph 2
Scatter Diagram between H2 and H5

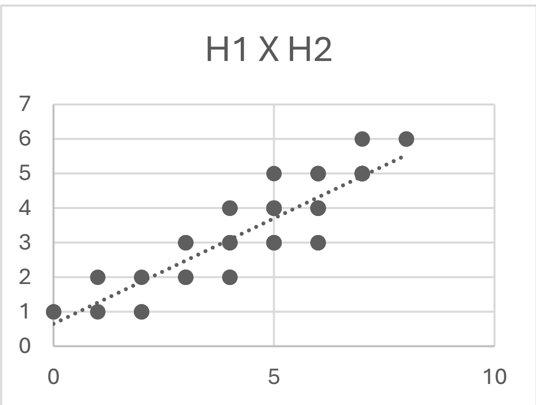


The correlations of H1 with H4 and H5 present similar values; respectively, the coefficients r=0.726 and r=0.729 point to a positive correlation with moderate strength. For the coefficient of Determination, the variation was the same, in the amount of 53%. The same skill (understanding) has a strong correlation with H2 (abstraction), with r=0.877 and R2 = 77% of variability. Therefore, it can be inferred that the more the ability to understand is present, the more the possibility of the correctness of abstraction and problem solving increases (H3).

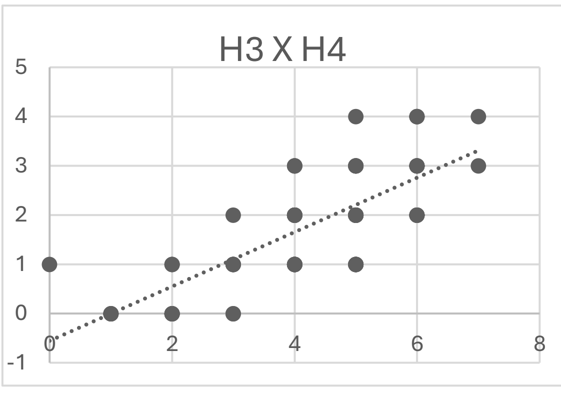
Revisiting Table 6, it can be seen that the highest Pearson r coefficients occurred in the correlations of H1 with H3, since r=0.888, and with a small difference between H1 and H2, with the value of r=0.877. As for the coefficient of determination, values of 79% and 77%, respectively, can be observed. Analyzing from the point of view of the CT and the CTRM, it can be inferred that the correct resolution of the problem (H3) can be explained, in 79%, in the understanding of the problem (H1) and in the ability of algorithmic construction (55%).

Graphs 3 and 4 show the scatterplot between H1 and H2 and between H3 and H4, respectively, thus confirming the linearity of the data.

Graph 3
Scatter Diagram between H1 and H2



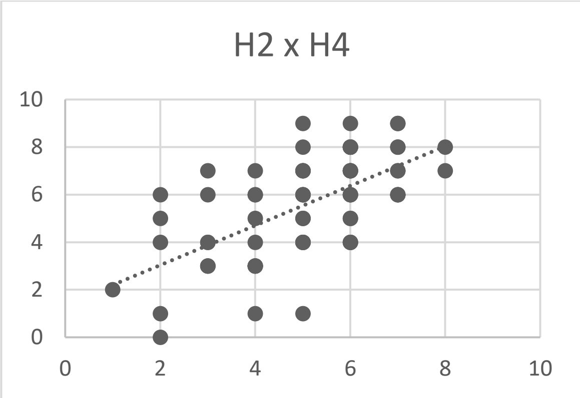
Graph 4
Scatter Diagram between H3 and H4



It is worth clarifying that all scatter diagrams constructed from the values of these variables present positive relationships, at least moderate. The lowest values of r= 0.626 were between H2 and H4 (Graph 5) and r= 0.627 between H2 and H5 (as shown in Graph 2).

In summary, from the statistical tests, it is possible to affirm that all correlations were positive, ranging from very weak (the case of H2 with H4) to a moderate to strong positive relationship, with values above +0.8 (Barbetta (2002). That is, all skills are correlated with each other, some more pronounced, as in the case of problem solving skills (H3).

Graph 5
Scatter Diagram between H2 and H4



Interpreting these values from the point of view of Computational Thinking, it can be inferred that, when the subject understands a problem, he is able to make the necessary abstractions (H1 x H2, for r= 0.877) and is able to propose the most correct possible solution (H2 with H3, where r=0.858) and transform it into an algorithm and/or software (H3 x H4 where r= 0.741 and R2=55%) and, subsequently, is able to validate or test it. lo (H4 x H5 having r = 0.869 and R2 = 76%).
Therefore, it is concluded that the CTRM is correct and can be applied as valid skills, in order to be used to measure the development of CT in the subjects. It should be noted that this model can be used in other case studies, with the possibility of adding more tests to meet the demand and reality of the subjects.

6. FINAL CONSIDERATIONS

The understanding of what computational thinking is and its way of evaluating are topics that still lack discussion and research. Evaluating the CT, in a comprehensive way, through the understanding of the skills involved in the process, can be a way to go.
This research contributed to a step towards a multifaceted assessment, mainly to understand that skills are integrated, related and correlated.
From the statistical analysis, it can be confirmed that the work presented evidence of reliability and correctness of the model, as well as the quality of the complete evaluation. On the other hand, it is suggested that the amount of the sample could be expanded in future works, as well as changes in the test questions. Case studies using other programming languages would certainly provide data to establish comparisons between programming languages and the skills proposed in the CTRM.

7. REFERENCES

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., & Cochran, J. J. (2019). *Estatística aplicada à Administração e Economia* (8th ed.). Cengage.

Araújo, A. L. S. O. de. (2019). *Quantifying Computational Thinking Abilities* [Tese (Doutorado em Ciência da Computação)]. Universidade Federal de Campina Grande.

Avila, C., Cavalheiro, S., Bordini, A., Marques, M., Cardoso, M., & Feijo, G. (2017). Metodologias de Avaliação do Pensamento Computacional: uma revisão sistemática. *Simpósio Brasileiro de Informática Na Educação (SBIE), XXVIII., 1*, 113. https://doi.org/10.5753/cbie.sbie.2017.113

Barbetta, P. A. (2002). *Estatística aplicada às Ciências Sociais* (5th ed.). UFSC.

Brackmann, C. P. (2017). *Desenvolvimento do pensamento computacional através de atividades desplugadas na Educação Básica* [Tese (Doutorado em Informática na Educação), Universidade Federal do Rio Grando do Sul]. https://doi.org/10.1111/j.1469-7610.2010.02280.x

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. *American Educational Research Association - AERA*, 1–25. https://web.media.mit.edu/~kbrennan/files/Brennan_Resnick_AERA2012_CT.pdf

Cordenonzi, W. H. (2020). *O desenvolvimento do pensamento computacional e as evidências da alfabetização em código em adultos* [Tese de doutorado, Universidade do Vale do Taquari – Univates]. Repositório Institucional da Univates. http://hdl.handle.net/10737/2912 CRTC. (2018). *Currículo de Referência em Tecnologia e Computação*. http://curriculo.cieb.net.br/

Dagiené, V., & Futschek, G. (2008). Bebras International Contest on Informatics and Computer Literacy: Criteria for Good Tasks. *Informatics Education - Supporting Computational Thinking - Internacional Conference on Informatics, III .,* 19–30. https://doi.org/10.1007/978-3-540-69924-8

2016França, R. S. De, & Tedesco, P. C. A. R. (2015). Explorando o pensamento computacional no Ensino Médio: do design à avaliação de jogos digitais. *Workshop Sobre Educação Em Computação (WEI), XXIII, January 2015*. https://doi.org/https://doi.org/10.5753/wei.2015

Gatti, B. A. (2004). Estudos quantitativos em Educação. *Educação e Pesquisa*, *30*(1), 11–30. https://doi.org/10.1590/s1517-97022004000100002

Gil, A. C. (2008). *Métodos e Técnicas da Pesquisa Social* (6th ed.). Atlas.

Google for Education. (2015). *Google for Education: Computational Thinking*. https://edu.google.com/resources/programs/exploring-computational-thinking/

Grover, S., Basu, S., Bienkowski, M., Eagle, M., Diana, N., & Stamper, J. (2017). A framework for using hypothesis-driven approaches to support data-driven learning analytics in measuring computational thinking in block-based programming environments. *ACM Transactions on Computing Education*, *17*(3), 14. https://doi.org/10.1145/3105910

Grover, S., & Pea, R. (2013, January 1). Computational Thinking in K–12. *Educational Researcher*, *42*(1), 38–43. https://doi.org/10.3102/0013189X12463051

ISTE, & CSTA. (2011). *Operational Definition of Computational Thinking for K–12 Education*. https://www.iste.org/explore/Solutions/Computational-thinking-for-all

Korkmaz, Ö., Çakir, R., & Özden, M. Y. (2017). A validity and reliability study of the computational thinking scales (CTS). *Computers in Human Behavior*, *72*, 558–569. https://doi.org/10.1016/j.chb.2017.01.005

Lai, R. P. Y. (2022). Beyond Programming: A Computer-Based Assessment of Computational Thinking Competency. *ACM Transactions on Computing Education*, *22*(2), 1–27. https://doi.org/10.1145/3486598

LAKATOS, E. M., & MARCONI, M. de A. (2003). *Fundamentos de Metodologia Científica* (5th ed.). Atlas.

Minayo, M. C. de S., Deslandes, S. F., Neto, O. C., & Gomes, R. (2002). *Pesquisa Social: teoria método e criatividade* (21st ed.). Vozes.

Moreno León, J., Román González, M., & Robles, G. (2018). On computational thinking as a universal skill: A review of the latest research on this ability. *IEEE Global Engineering Education Conference (EDUCON), 2018-April*(April), 1684–1689. https://doi.org/10.1109/EDUCON.2018.8363437

Papert, S. (1980a). *LOGO: Computadores e Educação* (J. Valente (ed.); 1st ed.). Editora Brasiliense.

Papert, S. (1980b). *Mindstorms: children, computers and powerful ideas*. Basic Books.

Raabe, A., Santana, A. L. M., Ellery, N., & Gonçalves, F. A. (2017). Um Instrumento para Diagnóstico do Pensamento Computacional. *Congresso Brasileiro de Informática Na Educação (CBIE), VI .,* 1172–1181. https://doi.org/10.5753/cbie.wcbie.2017.1172

Reis, M. M. (2008). *Estatística aplicada à Administração* (UFSC (ed.)). Departamento de Ciências da Administração- UFSC.

Román González, M. (2016). *Codigoalfabetización y Pensamiento Computacional en Educación Primaria y Secundaria: Validación de un Instrumento y Evaluación de Programas* [Tese (Doctorado En Educación)]. Universidad Nacional de Educación a Distancia.

Román González, M., Pérez González, J. C., & Jiménez Fernández, C. (2017, July). Which cognitive abilities underlie computational thinking? Criterion validity of the Computational Thinking Test. *Computers in Human Behavior*, *72*, 678–691. https://doi.org/10.1016/j.chb.2016.08.047

Tang, X., Yin, Y., Lin, Q., Hadad, R., & Zhai, X. (2020). Assessing computational thinking: A systematic review of empirical studies. *Computers & Education*, *148*, 103798. https://doi.org/10.1016/j.compedu.2019.103798

Wing, J. M. (2006, March 1). Computational thinking. *Communications of the ACM*, *49*(3), 33–35. https://doi.org/10.1145/1118178.1118215

Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *366*(1881), 3717–3725. https://doi.org/10.1098/rsta.2008.0118

Yin, R. K. (2001). *Estudo de Caso: Planejamento e Métodos* (2nd ed.). Bookman.

Recepción: 14/01/2025 *Aceptación: 21/04/2025*

Walkiria Helena Cordenonzi

Possui graduação em Informática pela Pontifícia Universidade Católica do Rio Grande do Sul; Licenciatura em Formação Pedagógica para Graduados não Licenciados pelo IFSul (2023); Especialização Em Ciência da Computação pela Universidade de Passo Fundo, Brasil; mestrado em Ciência da Computação pela Universidade Federal do Rio Grande do Sul e doutorado em Ensino (2020). Docente do Instituto Federal de Educação, Ciência e Tecnologia Sul-Rio-Grandense, Câmpus Santana do Livramento, nos cursos binacionais e docente no Mestrado Profissional em Educação Profissional e Tecnológica em Rede – ProfEPT - Brasil. Orientação de tesis de mestrado no programa de Educação, na Universidad de la Empresa (Uruguay). Tem experiência na área de Ciência da Computação, com ênfase em Qualidade de Software. Áreas de interesse: alfabetização em código, pensamento computacional (PC) e suas avaliações, Informática na Educação e Andragogia.