IDEODIGITAL IMPACT EVALUATION

Computer Science in the Classroom

Kodea, 2024 - 2025

Kodea authors:
Andreas Hein - Head of Research
Claudio Farías - Research Analyst
Francisca Avendaño - Research Analyst
Larratiz Otamendi - Research Analyst

Focus authors:
Laura Ramaciotti
Antonia Arrendondo
Bernardita Canals
Isidora Fuenzalida

Editing and design (Kodea):
Camila del Solar
Karla Cantuarias
Alejandra Moya

Initiative
Developed by
Supported and Endorsed by
In Collaboration with

CONTENTS

1 – Introduction 05

2 – What are Computer Science and Computational Thinking? 07

3 – What is IdeoDigital? 09

4 – Impact Evaluation of IdeoDigital's Transfer Model 13

5 – Results and Impact Estimation 22

6 – Conclusions, Reflections, and Recommendations 40

Photo: Hour of Code 2023, Colegio San Fernando, Municipality of Peñalolén, Santiago.

EXECUTIVE SUMMARY

In different countries around the world, the need to incorporate Computer Science (CS), Computational Thinking, and programming into school systems is increasingly being discussed. Global trends show two main pathways: teaching CS as a specialized subject or integrating it across into the curriculum, as a way to promote higher-order skills—particularly problem-solving. Regardless of the strategy, there is consensus that teacher training is an essential and critical component. In Chile, systematic State efforts to promote digital literacy began in the 1990s, following a logic shared across Latin America: reducing the access gap to computers and the internet and training competent users in basic applications such as office software and web navigation.

Programs such as **Enlaces, Yo Elijo Mi PC y Me Conecto para Aprender** focused on infrastructure, networks, digital resources, and technical assistance for teachers. However, by 2018, the el **Plan Nacional de Lenguajes Digitales** introduced a more ambitious goal: fostering the teaching of Computational Thinking and programming, although with limited reach.

In 2021, Fundación Kodea developed **IdeoDigital**, a nationwide initiative supported and endorsed by **BHP Foundation**, seeking to continue that vision by developing a comprehensive teacher training and mentoring model inspired by the**Code.org** approach. This model ensures effective classroom transfer through training, free resources, and mentoring, aiming to move beyond basic digital literacy toward the development of complex skills that enable students not only to consume technology, but to understand it, question it, and create with it.

This document presents the **final results of the impact evaluation of the IdeoDigital Basic program, conducted by Focus.** It is a quasi-experimental panel study that included three measurements over 18 months. The final sample consisted of 67 treated schools and 36 schools in the control group, selected through Propensity Score Matching (PSM), a methodology that ensures data comparability by matching schools with similar characteristics (treated and untreated). In both groups, students' Computational Thinking skills from 3rd to 6th grade and teachers' knowledge in Computer Science were measured at three points in time: a baseline in May 2024, an intermediate measurement in December 2024, and a final measurement in July 2025.

Student results suggest a positive, statistically significant, and increasing impact on Computational Thinking skills. Specifically, the program produced an average 7% increase in the percentage of correct answers. This impact is significant at a confidence level above 99% across all model specifications used, providing strong and compelling evidence of its effectiveness.

For teachers, findings also indicate positive and significant impacts on strengthening CS knowledge, perceived ease of technology use, and advanced technological skills. This impact is significant at a confidence level above 99%.

These results confirm the program's potential as a strategy to accelerate the integration of Computer Science into the Chilean school system, addressing the gaps left by first-generation digital literacy policies and aligning with the demands of digital transformation driven by artificial intelligence.

La Hora del Código 2024, Festival Puerto de Ideas en Concepción

1. INTRODUCTION

Across the world, there is discussion about the need to incorporate Computer Science (CS), Computational Thinking, and/or programming into school systems. Within global trends, one approach is to teach CS as specialized subjects; another is to integrate it across the curriculum as a way to promote higher-order skills, particularly problem-solving. Whichever path is chosen, teacher training remains an essential and critical component

Since the early 1990s, the Chilean State began implementing its first systematic efforts to promote digital literacy in schools.. As in much of Latin America, the first generation of digital literacy policies focused heavily on reducing the access gap to computers and the internet, and on training competent users of technological applications such as office software and web browsing[1].

Among the first-generation digital literacy policies and programs implemented in Chile, the programs "Enlaces", "Yo elijo mi PC" (YEMP) and "Me Conecto para Aprender" (MCPA) stand out. These programs were strongly focused on expanding access to infrastructure, networks, and digital resources (including educational software, productivity tools, and internet resources), as well as providing teacher training, and technical assistance.[2,3] (Enlaces) They also aimed to promote access to and use of technological resources by 7th-grade students in public schools to support learning processes[4,5,6] (YEMP, MCPA).

In 2018, the National Digital Languages Plan was launched with the objective of fostering the teaching of Computational Thinking and programming in the education system, although its reach was limited. Fundación Kodea—whose fundamental educational vocation is to promote Computer Science education in Chile—participated from its inception through the execution of an initial pilot.

Each of these milestones has shown that, alongside advances in digital technologies, new educational needs emerge. The social impact of technology, the need to prepare new generations to thrive in an increasingly digital world, and understanding the new ethical and scientific dilemmas posed by these advances are all fundamental to forming citizens who are capable of actively participating in the knowledge society.

With the disruption brought about by artificial intelligence, it is no longer enough to teach office software and internet use. Instruction must be expanded to equip children and young people with advanced Digital Skills to enable them to understand how digital technologies work and the foundational principles behind them, with special emphasis on Computer Science (CS) and Computational

1 – OECD.(2020). Making the Most of Technology for Learning and Training in Latin America. https://doi.org/10.1787/ce2b1a62-en. © 2020 OECD.Paris.
2 – Quiénes Somos. (n.d.). Retrieved July 27, 2021, from Enlaces website: (http://www.enlaces.cl/sobre-enlaces/quienes-somos/)
3 – Universidad Diego Portales. (2012). Evaluación de Impacto programas TICS'S Ministerio de Educación. Final report.
4 – Dirección de Presupuestos. (2020). Evaluación Ex Ante Proceso Formulación Presupuestaria 2021. Becas de Acceso TIC. Retrieved from: https://programassociales.ministeriodesarrollosocial.gob.cl/pdf/2020/PRG2020_2_80944.pdf
5 – Pontificia Universidad Católica de Chile. (2017). Final Report: Evaluation of Implementation of the Me Conecto Para Aprender Program. UNESCO. Center for Studies on Policies and Practices in Education (CEPPE UC), Dirección de Estudios Sociales (DESUC).
6 – Katalejo. (2019). User satisfaction evaluation of the TIC Scholarships Program Yo Elijo mi PC and Me Conecto para Aprender: Final report. Commissioned by JUNAEB, Chile: Feller, C., Alvarado, P. & García, I.

> The contribution made by IdeoDigital in its five years of implementation has been to develop a model for the ongoing training, support, and education of teachers...

Thinking[7]. Future generations must be the architects of the digital world and not only consumers of it. This is not only about education today, but also about future employability: basic digital literacy is no longer sufficient to meet the demands of ongoing digital transformation.

Given this scenario, many countries[8] have revised school curricula to incorporate the teaching of CS concepts and to develop students' Computational Thinking. Since 2021, Fundación Kodea, with the support and endorsement of BHP Foundation, has developed and implemented IdeoDigital to create the necessary conditions for integrating Computer Science into the Chilean school system.

The challenges faced during the pilot, implementation, and scaling phases include: limited digital skills among students, teachers, and schools overall; widespread technological infrastructure that remains fragile and insufficient for instructional purposes; an outdated technology-related curriculum; and a lack of adequate teacher training in this field.

Over its five years of implementation, IdeoDigital has contributed by developing a model for ongoing teacher training, mentoring, and capacity-building, based on the Code.org teaching–learning model for Computer Science and programming in school education (www.code.org). Additionally, this model provides teachers with free resources that ensure effective transfer of skills to the classroom and promote the development of Computational Thinking among students.

This document presents the results of the impact evaluation of the IdeoDigital transfer model, focused on strengthening teacher learning for Computer Science instruction and fostering the development of Computational Thinking among students in 3rd to 6th grade.

7 – Computer Science is an academic discipline that develops knowledge related to computers and algorithms, encompassing fundamental principles, hardware and software design, practical applications, and social impact. This body of knowledge includes topics such as problem analysis, programming and algorithms, data storage structures, computer architecture, networks, cybersecurity, robotics, artificial intelligence, and machine learning. Understanding why and how computers work—i.e., Computer Science—provides the foundation for a deeper understanding of computer use and its associated relevant rights, responsibilities, and applications. Computer Science includes related concepts and content such as Computational Thinking and programming (or coding). Computational Thinking refers to a mental process that allows problems to be formulated such that their solutions can be carried out by computers. Developing Computational Thinking involves developing skills of conceptualization, analysis, and solution development for complex problems by selecting and applying strategies and tools characteristic of Computer Science. It implies thinking in terms of abstraction and generalization; modeling and breaking down problems into subproblems; analyzing processes and data; as well as creating digital artifacts—both virtual and real—among other abilities. Meanwhile, programming or coding refers to the ability to define a set of instructions for a computer to perform a specific task (Jara, I., & Hepp, P., 2016). Computer Science is also connected to a range of disciplines and related concepts such as Digital Citizenship, Robotics, Cybersecurity, Computational Thinking, Programming, Machine Learning and Artificial Intelligence, Data Science, Networks and Communications, and Video Game Design.

8 – For more information, see: Hein, A., Otamendi, L., Farías, C. (2022). Current conditions for the implementation of initial teacher training programs in Computer Science. Research Area, Fundación Kodea. This document reviews CS education policies and teacher professional profiles in Israel, the United Kingdom, Argentina, Uruguay, and Estonia.

Hour of Code, AI for Oceans 2023 – Colegio Francisco Arriarán

2.

WHAT ARE COMPUTER SCIENCE AND COMPUTATIONAL THINKING?

Computer Science is an academic discipline that develops knowledge related to computers and algorithms, encompassing their fundamental principles, hardware and software design, practical applications, and impact on society. It covers topics such as problem analysis, programming and algorithms[9], data storage structures, computer architecture, networks, cybersecurity, robotics, artificial intelligence, and machine learning.

Knowing why and how computers work—that is, Computer Science—provides the basis for a deep understanding of computer use, as well as the associated rights, responsibilities, and applications[10].

A misconception about what Computer Science (CS) truly is can be a major obstacle to teaching this discipline in school. To address this, the K–12 Computer Science Framework explains what CS includes and what students should learn from early childhood through upper secondary education.
According to this framework, CS relates to four areas: computer literacy (knowing how to use computers and basic software), educational technology (using tools to support learning across subjects such as drafting and editing an essay online), digital citizenship (using technology safely and responsibly), and information technology (focused on installing and managing software in workplace contexts). Although the latter overlaps with CS, it differs in that it focuses on applying computing rather than creating new solutions. .

Computer Science also contains related concepts and content such as Computational Thinking and programming. Computational Thinking is the mental process that enables problems to be formulated in such a way that their solutions can be carried out through a computer. It involves thinking in terms of abstraction and generalization; modeling and decomposing problems into subproblems, analyzing processes and data, and creating digital artifacts, both virtual and physical,, among others. In other words, developing Computational Thinking means learning to understand problems, breaking a problem into smaller steps, identifying patterns, focusing on what truly matters, and designing a clear set of instructions to reach a solution that can be executed by a computer—this is programming.

9 – These are step-by-step procedures that computers must follow in order to acquire, represent, structure, process, and communicate data, as well as perform calculations.
10 – K–12 Computer Science Framework. (2016). Retrieved from http://www.k12cs.org

2.1 Why promote the development of Computational Thinking?

Computer Science is not an end in itself, but a tool for solving problems. Therefore, it is not limited to learning how to code. Studying and learning CS allows children and adolescents to exercise a broad range of cognitive and non-cognitive skills.

Evidence shows that teaching CS and coding helps exercise cognitive skills such as mathematical problem-solving, critical thinking, social skills (e.g.,communicating with others), self-management learning, planning skills, logical thinking, reflective thinking, and problem-solving. Among non-cognitive and/or socio-emotional skills, persistence, frustration tolerance, and collaboration stand out.

There is also evidence that skills developed through teaching coding in school can strengthen other abilities—for example, planning skills. Likewise, it has been observed that some of the aforementioned skills can be developed through early coding instruction, even from the ages of 4 or 5.

Developing Computational Thinking can also enhance student empowerment by providing tools to express creativity and shape projects that respond to real-world problems. In particular, through project-based learning methodologies, Computer Science promotes motivation, enjoyment, and student engagement with their own learning. Studies conducted by Code.org report that students enjoy Computer Science classes more than other subjects and that 70% of them want to continue learning CS beyond elementary school.

There is also empirical evidence supporting the relationship between Computer Science and creativity, understood as the process by which a problem emerges—whether imagined or visualized—and then an idea, concept, notion, or scheme is created along unconventional lines to solve the problem. In this way, Computer Science can strengthen creativity development based on students' interests, needs, and knowledge, It teaches thinking strategies such as: solving unexpected problems, formulating alternatives, proposing and implementing designs, making observations, abstraction across various topics, drawing exercises, and the using metaphors and analogies to address problems.

11 – Kalelioğlu, F. (2015). A new way of teaching programming skills to K-12 students: Code.org. Computers in Human Behavior, 52, 200–210.

12 – Popat, S., & Starkey, L. (2019). Learning to code or coding to learn? A systematic review. Computers & Education, 128, 365–376.

13 – Arfé, B., Vardanega, T., & Ronconi, L. (2020). The effects of coding on children's planning and inhibition skills. Computers & Education, 148, 103807.

14 – Kalelioğlu, F. (2015). A new way of teaching programming skills to K-12 students: Code.org. Computers in Human Behavior, 52, 200–210.

15 – Scherer, R., Siddiq, F., & Sánchez Viveros, B. (2019). The cognitive benefits of learning computer programming: A meta-analysis of transfer effects. Journal of Educational Psychology, 111(5), 764.

16 – Serdar Çiftci & Ahmet Bildiren. (2020). The effect of coding courses on the cognitive abilities and problem-solving skills of preschool children. Computer Science Education, 30(1), 3–21. DOI: 10.1080/08993408.2019.1696169

17 – Weintrop, D., & Wilensky, U. (2019). Transitioning from introductory block-based and text-based environments to professional programming languages in high school computer science classrooms. Computers & Education, 142, 103646.

18 – Sáez-López, J. M., Román-González, M., & Vázquez-Cano, E. (2016). Visual programming languages integrated across the curriculum in elementary school: A two-year case study using Scratch in five schools. Computers & Education, 97, 129–141.

19 – Code.org is a U.S. NGO that promotes programming education by providing interactive courses through a virtual platform.

20 – Code.org.

21 – Seo, Y. H., & Kim, J. H. (2016). Analyzing the effects of coding education through pair programming for the computational thinking and creativity of elementary school students. Indian Journal of Science and Technology, 9(46), 1–5.

22 – Pérez Palencia, M. (2017). Computational Thinking to strengthen the development of skills related to creative problem solving. 3C TIC: Cuadernos de desarrollo aplicados a las TIC, 6(1), 38–63.

IdeoDigital launch at Escuela San Nicolás, Ñuble 2022.

3.

WHAT IS IDEODIGITAL?

Initiative that began in 2021 and ends its first phase in December 2025. Its central objective is to create the necessary conditions to implement Computer Science (CS) in Chile's public school system. To achieve this, it has a working structure based on three interconnected components that are developed in parallel and mutually reinforce each other.

1 -Awareness-Raising in the Educational Community

This component seeks to motivate schools to train their teachers in CS and to engage the entire school community in this process. Activities target school authorities, leadership teams, teachers, families, and students, disseminating information about the feasibility and benefits of teaching CS, as well as its impact on skills development. The strategy includes media outreach campaigns, the creation of a network of leading schools to serve as references for others, and the implementation of an official certification, developed in coordination with the Ministry of Education, to recognize schools that integrate CS into their classrooms.

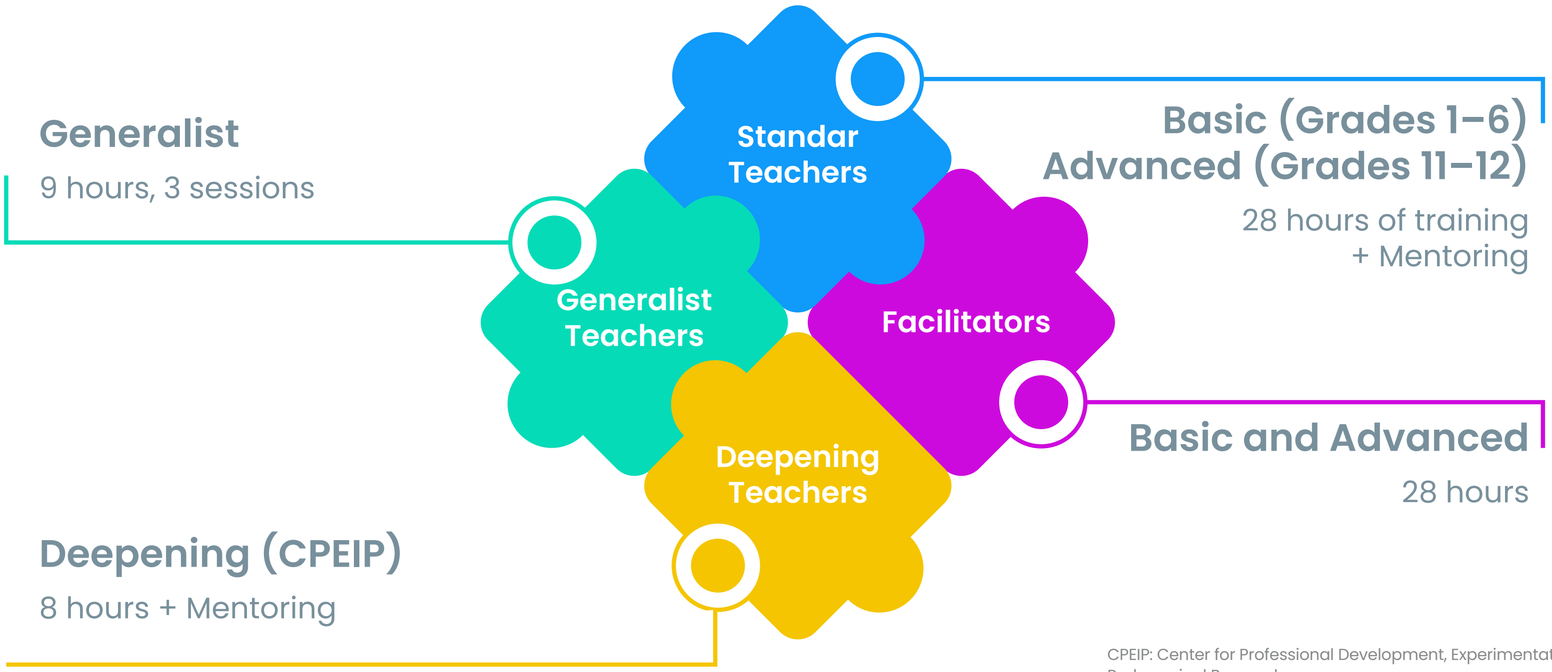
2 - Public Policy Advocacy

This component aims to transform the national Technology subject program by incorporating CS into the official curriculum and developing a model program approved by the Ministry of Education. To support this process, evidence has been generated and disseminated on effective CS teaching experiences and their impact, working closely with education stakeholders and legislators to drive regulatory changes.

3 - Classroom Implementation

Training programs have been designed and adapted to different levels: IdeoDigital Basic (1st to 6th grade), IdeoDigital Advanced (11th and 12th grade), Deepening (for teachers with prior training), and Generalist (open to the general public). The implementation strategy includes preparing facilitators who train and mentor teachers, modeling classes, providing feedback, and supporting planning and evaluation. These programs are supported by a clear theory of change (Figure 2) and a systematized training offer (Figure 1).

Figure 1: Summary of Training Programs



CPEIP: Center for Professional Development, Experimentation and Pedagogical Research

To date, the I**deoDigital initiative has trained more than 1,350 teachers from 250 schools, reaching over 37,920 students.** Impact evaluations results show that the initiative has been successful in transferring knowledge and developing CS teaching skills among teachers who had no prior experience in the field.

Summit Ideando el Aula 2023
Talleres para docentes

Figure 2: Theory of Change

Purpose: To create the necessary conditions to implement Computer Science training in Chilean public schools.

NEEDS	INPUTS/ACTIVITIES	RINTERMEDIATE RESULTS	INTERMEDIATE RESULTS	FINAL RESULTS
• The 4.0 revolution is radically changing most areas of human activity. • To ensure inclusion in technological education, it is necessary to teach Computer Science (CS) in schools. • CS is unknown to principals and teachers, generating fear and resistance. • There are no available materials that facilitate teaching CS and digital skills in the classroom. • There are no trained and experienced teachers in CS classroom teaching. • Children do not receive CS training in school.	1. Inputs • Program management team. • Facilitators trained in the KODEA–CODE.org training system. • Instructional manuals adapted to the Chilean curriculum. • Online training platform. • Monitoring and evaluation system (instruments, data collection and processing platforms). • Mobility and logistical resources. 2. Activities 2.1 Awareness component • Ecosystem awareness-raising (presentations, webinars). • Recruitment of implementing partners. • Recruitment of schools. 2.2 Training component • Training of facilitators. • Training of teachers. 2.3 Mentoring component • Mentoring sessions. • Classroom implementation support. • Recognition awarded to schools for implementation achievements. • Orientation sessions for leadership teams to address implementation challenges and barriers. 2.4 Diagnosis, evaluation, and feedback component • School and teacher diagnosis. • Evaluation of training results. • Satisfaction evaluation. • Monitoring of mentoring processes • Monitoring of teacher competency development. • Monitoring use of the CODE platform. • Evaluation of students' learning outcomes and interest. • Awareness events (webinar, presentation to authorities).	*Awareness-raising events (webinar, presentation by authorities). *Schools sign the collaboration agreement. *Facilitators complete training. *Teachers from schools participate in training workshops *Teachers (90%) complete training. *Teachers (90%) engage in lesson planning, classroom observation, and feedback throughout the academic year. *Teachers register on the CODE platform (85%). *Students participate in connected and disconnected practice activities. *Students carry out activities on the CODE platform. *At least two group orientation sessions on managing implementation barriers are held for management teams. *Schools receive recognition for completing implementation requirements. *Principals receive feedback on school assessments. *Teachers and facilitators receive feedback on initial assessments. *Teachers and facilitators receive feedback on post-training assessments. *Teachers and facilitators receive feedback on post-support assessments. *Teachers and facilitators receive	*Awareness events (webinar, presentation by authorities). Facilitators: Strengthen their knowledge and skills to convey STEM teaching strategies to teachers. Principals: Improve their perception of the acceptability of STEM teaching. Teachers: Develop science knowledge, digital skills, and implementation skills through direct experience, and improve their perception of the acceptability of science teaching. Schools: *Collaborate during project participation and manage implementation difficulties/ barriers.	Students: *Develop digital and CC application skills (depending on age) *Increase their interest in CC, digital skills, and technology applications.

Assumptions

- School principals show interest, understand the nature and scope of the program, and actively support its implementation.
- Teachers demonstrate an interest in learning.

Risks

- Participating schools may fail to allocate or properly organize the necessary time, resources, or infrastructure. Example: using the computer room for other purposes during technology class time. To mitigate this risk, awareness-raising activities are conducted with management teams and infrastructure coordinators, and resource needs are also planned every three months.
- Dropout of participating teachers. To mitigate this risk, two or three teachers per school are trained, and peer-learning activities are organized so teachers can share and showcase their experiences
- Delays and postponement of sessions due to various contingencies (e.g., leave, strikes, health issues). To mitigate this risk, implementation controls are established, and a business continuity plan is activated.
  The project is mainly implemented online, which facilitates the continuity of activities.

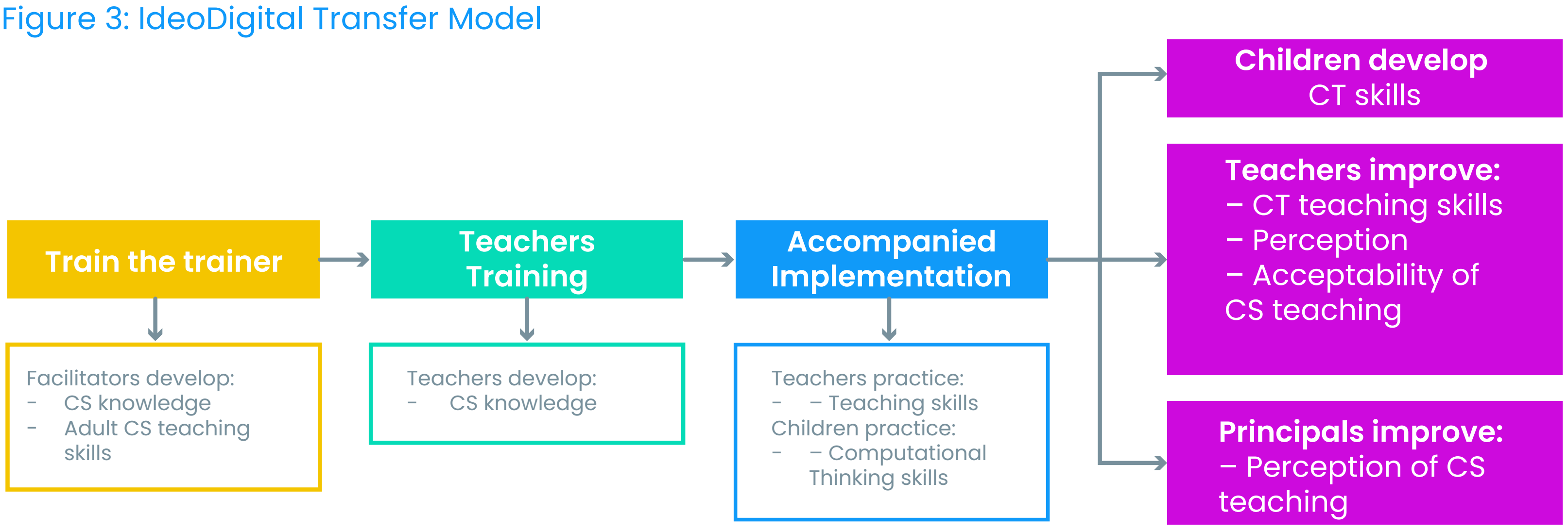
3.1 Why Promote the Development of Computational Thinking?

IdeoDigital's transfer model (Figure 3) seeks to enable teachers to learn and effectively apply Computer Science teaching in the classroom. To do so, facilitators from Technical Educational Agencies (ATEs), previously trained by Fundación Kodea, deliver a **20-hour synchronous online training program.** These sessions are complemented by **10 to 14 practical classroom sessions,** during which teachers teach their students while being observed by the facilitator, who then provides feedback using a structured observation rubric.

During the process, teachers implement a program of guided lessons with their students, supported by the Code.org platform, which offers interactive activities and progressively organized exercises. This approach combines theory, practice, and mentoring, ensuring that the methodology is integrated into everyday classroom instruction.

The model has undergone both internal and external evaluations to measure outcomes, adjust content, and refine the overall strategy. After this process, a decision was made to focus the impact evaluation on the most robust version of the model, one that has **demonstrated consistent and sustained results over time,** thereby consolidating a mature and effective methodology.

Figure 3: IdeoDigital Transfer Model



CS: Computer Science
CT: Computational Thinking

Photo: Technology class, Thomas Jefferson school, Independencia, Metropolitan Region, Chile.

IdeoDigital Los Lagos, Valparaíso, and Metropolitan Region Graduation, 2023

IMPACT EVALUATION OF IDEODIGITAL'S TRANSFER MODELL

4.

The impact evaluation focused on the IdeoDigital Basic model, implemented in schools serving students in grades 3 through 6.

4.1 Preparation and Monitoring

Since 2022, a monitoring and evaluation strategy has been developed, making it possible to systematize learning, generate evidence, and support decision-making. The evaluation model focused on generating and managing knowledge in order to:

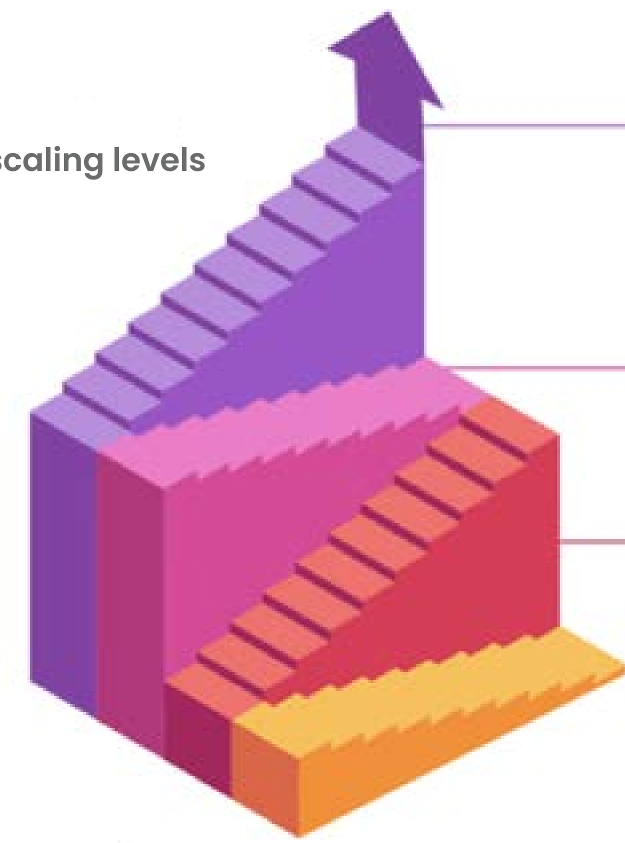
- Identify strategies and practices to overcome implementation barriers.
- Identify critical success factors for implementing different components of the initiative.
- Assess the effectiveness of the initiative, with special attention to strategies and factors that may moderate its results.

The monitoring and evaluation system included an applied research program that combined various activities with internal and external evaluation milestones, which progressively informed the initiative's development and scaling, as well as the systematization and dissemination of key lessons learned. Accordingly, , the main evaluation milestones were organized according to the following stages of program development:

- **Learning and maturation (2021–2022):** This stage included an external evaluation of the pilot, using surveys, interviews, and focus groups. It assessed facilitator and teacher readiness, the mentoring strategy, implementation barriers, and program acceptability. During this phase, measurement instruments and indicators were also developed to assess learning outcomes, teacher skills, satisfaction, and acceptability of the program concept, i.e., characterizing the availability of eligible actors to implement the program and identifying potential implementation barriers[23].
- **Consolidation and scaling (2023–2025):** During this stage, the impact evaluation was designed and implemented, two external process evaluations were conducted (2023 and 2025); and a Computational Thinking test for students was developed and validated (2023–2024).

23 – See, for example, Hein, A., Farías, C. (2022). Toward Installing Computer Science Teaching in the Classroom. Evaluation of the pilot implementation of the Idea Digital transfer model. Research Area, Fundación Kodea.

Figure 4: Relationship Between Evaluation Practices and Intervention Maturity



4.2 Methodological Framework for the Impact Evaluation

The evaluation was implemented between March 2024 and July 2025. Its central questions were:

- Does the IdeoDigital Basic program develop students' Computational Thinking skills?
- Does it improve acceptance of CS teaching among teachers and school leaders?

Beyond measuring effectiveness, the evaluation also sought to identify good practices, critical success factors, and barriers to scaling the teaching of Computer Science in Chile.

4.2.1 Evaluation methodology

a) Evaluation design

To evaluate the initiative's impact, a quasi-experiment was used.

A quasi-experiment (or quasi-experimental design) is a research approach that seeks to establish a cause–effect relationship between an independent variable (the cause) and a dependent variable (the effect). To do so, an experimental group (or treatment group) and a control group (one that does not receive the treatment or intervention) are constructed. Ideally, the only difference between the two groups should be the fact that one receives the intervention and the other does not. Unlike a true experimental design, participants in a quasi. experiment are not randomly assigned to groups. In a true experiment, the researcher has full control: they can randomly assign people to a treatment group (which receives the intervention) or to a control group (which does not). This randomization is key because it balances participant characteristics between groups, eliminating confounding factors. In contrast, in quasi-experiment designs, random assignment is not feasible due to practical or ethical constraints, and group membership is determined by pre-existing characteristics or natural circumstances.

Accordingly, in this evaluation the treated group consisted of public or publicly subsidized schools interested in participating in the program and had an operational computer lab connected to the Internet. The control group consisted of schools with characteristics similar to those of the treated group, selected through Propensity Score Matching (PSM), a statistical technique that ensures group comparability when randomization is not possible. (described in section 4.2.2).

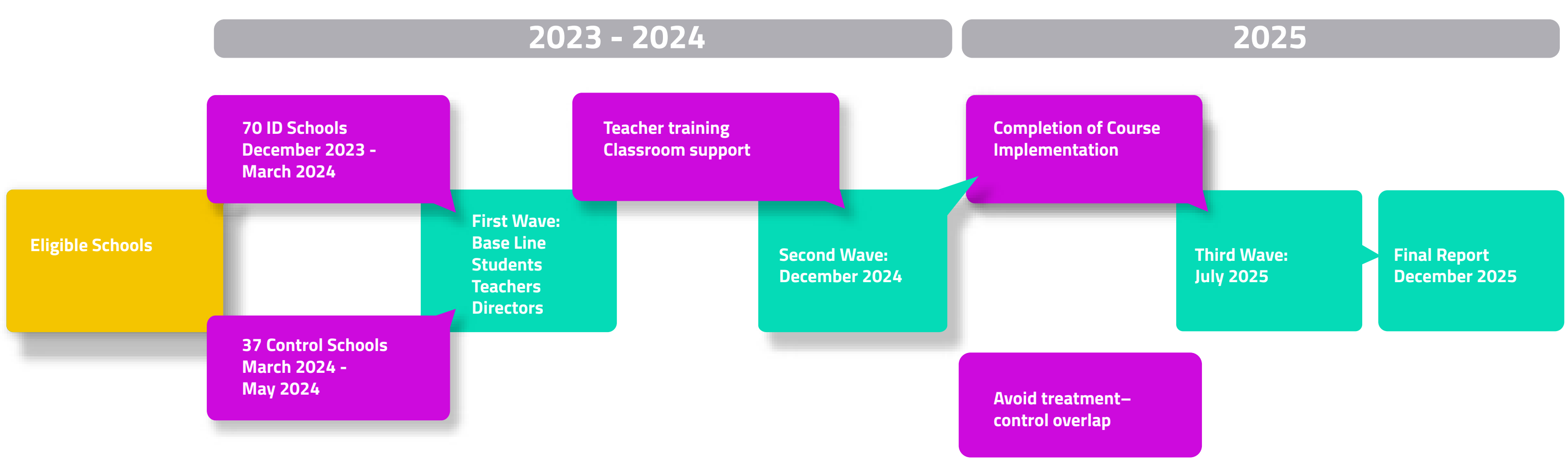
Once both groups were formed, a set of measurement instruments was applied to assess knowledge, skills, and attitudes related to teaching Computer Science and Computational Thinking. These instruments were applied to both groups (control and treatment) in three waves over an 18-month period.

The first measurement (May 2024) corresponded to baseline, the second (December 2024) corresponded to an intermediate evaluation coinciding with the end of the course and the practical classroom implementations for teachers, and the third (July 2025) to the final evaluation, once teachers had completed teaching a full course to the same group of students.

After data collection was completed, control-group schools received compensation consisting of CLP 600,000 to invest in technology, training in digital skills teaching (provided by Fundación Kodea), and **school leadership and management training** (provided by Focus, responsible for managing the control group).

The design summary is illustrated in Figure 5.

Figure 5: Research design



b) Universe and participants

Between 2021 and 2025, 250 schools participated in the IdeoDigital program. All of these schools are public or publicly subsidized (69% and 31%, respectively), most are urban (81%), and are located outside the Metropolitan Region (75%).

Participating schools show a high average School Vulnerability Index (IVE) (83.9%) and are predominantly classified as "Medium–Medium Low" performance schools (76% in 2019). Regarding SIMCE results, these schools perform slightly below the national average in 4th grade (Reading: 269.5; Mathematics: 255.7).

These antecedents suggest that the program has prioritized the participation of schools facing greater socio-educational challenges.

Complementarily, through analysis of the Institutional Educational Projects (PEI) of participating schools, a greater presence of technology-associated words has been confirmed (significantly higher than the national average among public and subsidized schools), which could indicate interests aligned with IdeoDigital's objectives.

4.2.2 Methodology for Selecting the Control Group (Propensity Score Matching).

As mentioned above, Propensity Score Matching (PSM) was used to select control-group schools for the evaluation in order to ensure comparability with treated schools through the identification of cases with similar characteristics.

The first step was defining a possible universe of schools, based on three essential criteria: (i) receiving state subsidies (thereby excluding fully private schools), (ii) not participating in Computer Science (CS) training programs, and (iii) having an operational computer lab. Due to the lack of databases to directly identify conditions (ii) and (iii), a checklist was applied to pre-selected schools to ensure the relevance of their participation.
Next, a propensity score was calculated for each school, representing the probability of participating in the program given its observable characteristics. This score ranges from 0 to 1, where 0 means very low probability and 1 means very high probability of participation. This approach made it possible to compare schools with similar profiles, regardless of whether they ultimately participated in the program.

To calculate the score, a multiple regression model was used considering the following predictor variables (selected on technical and strategic criteria):

- **Potential interest index:** created by Fundación Kodea (range 0–1100), reflecting a school's potential interest in the program. This index was constructed based on the frequency of technology-related terms (e.g., "technology," "computer science") appearing in the school's institutional program.
- **Administrative dependency:** obtained from the Educational Establishments Directory (2023), identifies whether the school is public (SLEP or municipal) or publicly subsidized private.
- **Initiative Factor 2023:** obtained from SNED 2022–2023. Index ranging 0–100 indicating the school's capacity to incorporate educational innovations and secure support from external agents in its pedagogical work.
- **Initiative Factor 2021:** obtained from SNED 2020–2021. Same as above.
- **Total enrollment:** obtained from the Educational Establishments Directory (2023). It is proposed as a proxy for the presence of the infrastructure required for the implementation of the program.
- **Index of Improvement of Conditions 2023:** obtained from SNED 2022–2023. Index 0–100 indicating improvement of working conditions and adequate functioning.

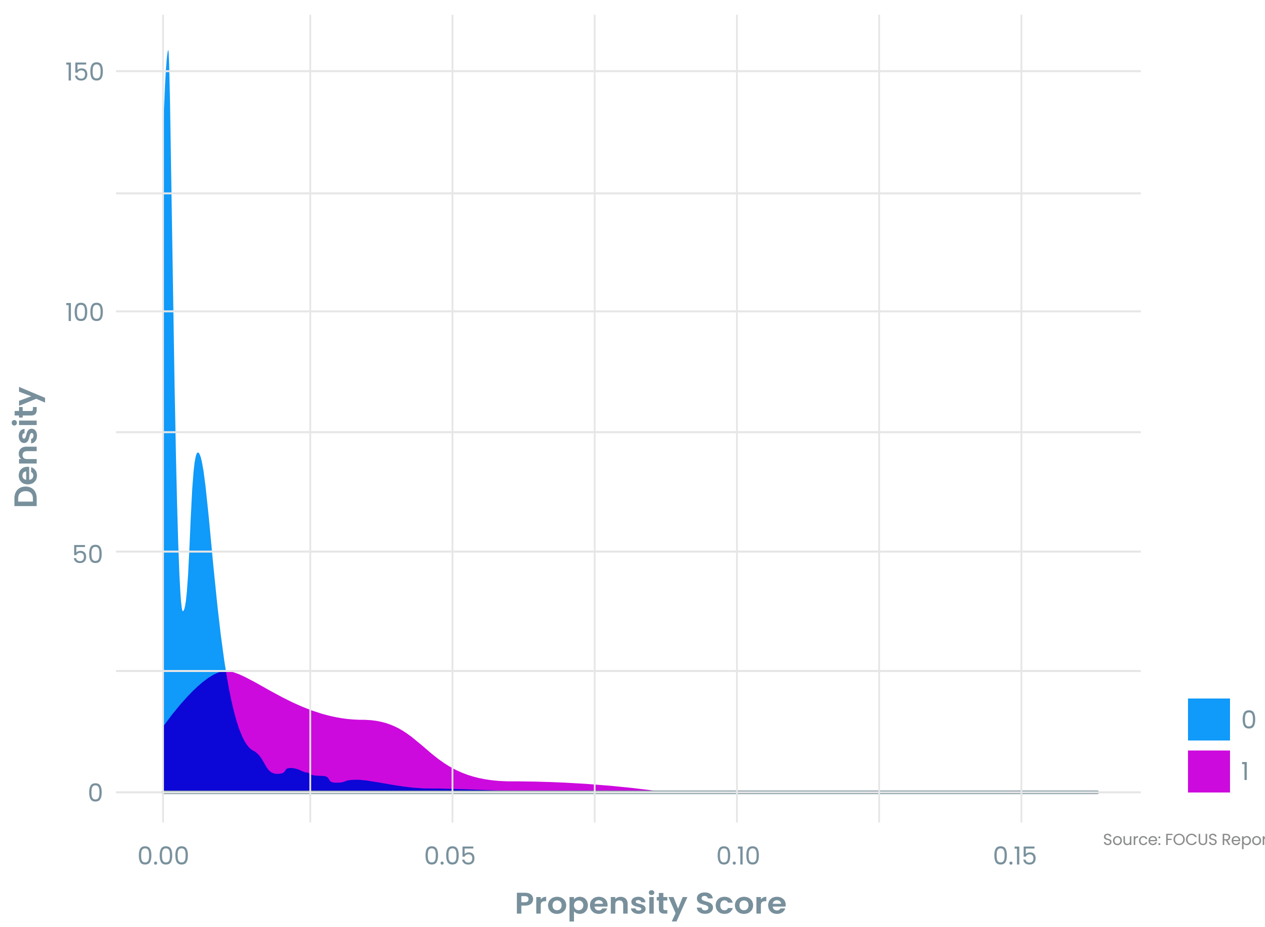
The selected model was the one that minimized the Akaike Information Criterion (AIC) and the Schwartz criterion[24].

24 – Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). Akaike information criterion statistics. Dordrecht, The Netherlands: D. Reidel, 81(10.5555), 26853.

After running the model and estimating scores, a common support region was identified, i.e., the range in which treated and untreated schools overlap in terms of similar scores. . From this region, a list of 171 possible controls[25] was generated, grouped into four priority levels according to proximity of propensity scores. From this list, 38 schools agreed to participate in the study. To correct potential imbalances between the final treatment and control groups, additional adjustment mechanisms and models were proposed, as detailed in section 4.2.6.

The following figure (Kernel density distribution chart) illustrates the degree and quality of overlap between treated and untreated schools. Specifically, the X-axis represents the estimated probability of receiving the treatment (propensity score), while the Y-axis shows the frequency (density) of schools for each propensity score valueThus, it identifies a significant common support region between values 0.001 and 0.164.

Figure 6: Propensity Scores for Treated Schools and Schools from the Untreated Universe (Kernel Density)



Source: FOCUS Report 2

It is important to note that within the common support region, control schools were chosen only within the regions of Antofagasta, Coquimbo, Valparaíso, O'Higgins, Maule, Biobío, Los Lagos, Aysén, Metropolitan Region, or Ñuble, corresponding to the regions where treatment-group schools are located.

Based on propensity score estimates and territoriality criteria, the following procedure was conducted to select control schools (with oversampling to cover potential replacements):

- Using the "second-nearest neighbor method", two potential control schools from the untreated universe are selected for each treatment-group school.
- For each treatment-group school, the primary control was defined as the paired school with the closest propensity score.
- From the list of prioritized controls, 43 untreated schools were selected, of which 36 ultimately participated in the study.

25 – The first-priority group were those with propensity scores closest to the treatment-group schools.

26 – In addition, only schools located in the same regions as treated schools were considered: Antofagasta, Coquimbo, Valparaíso, O'Higgins, Maule, Biobío, Los Lagos, Aysén, the Metropolitan Region, and Ñuble.

4.2.3 Achieved Sample and Attrition

Once data collection was completed, the databases were cleaned and refined. For students, three criteria were established to validate responses. Cases were considered invalid when: (i) the RBD (school identifier) could not be identified; (ii) the total number of responses was two or fewer; or (iii) total response time was under one minute, regardless of the number of questions answered. Variation in the number of cases across waves is presented below:

Figure 7: Student Database Cleaning (Counts)

	Wave 1		**Wave 2**		**Wave 3**	
	Treated	**Control**	**Treated**	**Control**	**Treated**	**Control**
Missing RBD	505	0	525	0	518	0
Invalid responses	24	707	18	873	22	850
Total valid cases	8411		7746		7919	

Subsequently, a new analytical database was constructed identifying grade level and school, where each observation corresponds to a grade level within a school. . In this new database, all observations with incomplete data were removed (i.e.,schools or grade levels with mismatched information in pre- and post-measurements, or vice versa). In addition,the outcome variable was redefined: instead of the percentage of correct individual responses, it became the classroom-level average percentage of correct responses.

For teachers, a validation and cleaning procedure equivalent to that applied to students was implemented, with the additional exclusion of incomplete and duplicate surveys. To ensure longitudinal traceability, an identifier variable was created based on respondents' email addresses, enabling the linkage of each teacher's data across the three waves and discarding cases without a valid match. Variation in the number of cases across waves is presented below:

Figure 8: Cleaning of teacher databases (counts)

	Wave 1		**Wave 2**		**Wave 3**	
	Treated	**Control**	**Treated**	**Control**	**Treated**	**Control**
Incomplete / missing data:	0	4	24	0	2	8
Duplicate responses	0	10	0	0	-	4
No match between instruments (wave 1):	22	24	-	-	-	-
Total valid cases:	181		183		152	

4.2.4 Variables and Instruments

This section details the variables and instruments used in the impact evaluation.

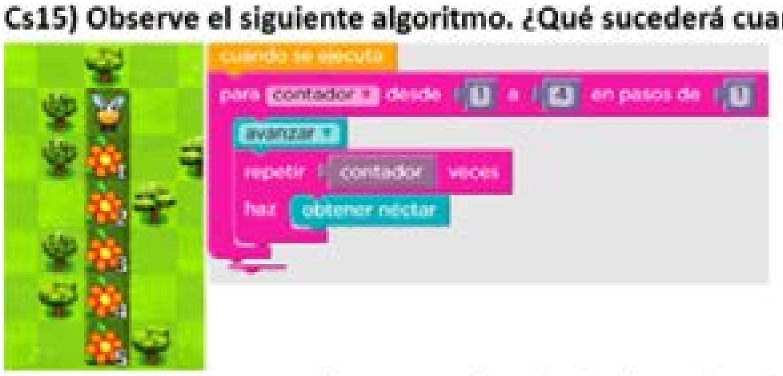
Teacher Knowledge

To measure teacher's knowledge, a 20-question test was implemented, focused on both the theoretical foundations and practical application of Computer Science The questions were designed to measure competency-based learning; therefore, correct answers require require understanding and applying knowledge about programming and use of the Blockly programming language—specifically, the use of sequences, events, loops, variables, functions, and counters.

Figure 9 shows examples of the questions used.

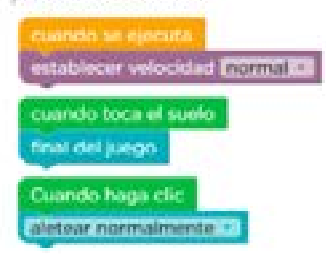
Figure 9: Examples of test questions on teaching knowledge

Cs15) Observe el siguiente algoritmo. ¿Qué sucederá cuando haga clic en "Ejecutar"?



a. La abeja avanzará y recogerá todo el néctar de cada flor
b. La abeja avanzará, pero no recogerá el néctar
c. La abeja avanzará y recogerá el néctar sólo de la primera flor
d. La abeja avanzará y recogerá el néctar de cada flor, excepto de una

Cs7) Observe el siguiente algoritmo. Esta línea de códigos en bloque, ¿a qué concepto de la programación hace referencia?



a. Loop/ciclo
b. Secuenciación
c. Condicional
d. Evento

Acceptability of Computer Science Teaching: Principals and Teachers

To measure the acceptability of CS teaching in the classroom, a 16-item Likert-type questionnaire was applied, with differentiated versions for principals and teachers. The instrument focused on evaluating indicators related to perceived usefulness and ease of teaching Computer Science.

- How likely do you think it is that you will teach Computer Science in the next 12 months? **[Implementation intention]**
- How useful do you think it will be for your students to learn about Computer Science? **[Perceived usefulness]**
- In general, how easy would you say it is for you to use technologies for teaching purposes? **[Perceived ease of use]**
- Would you say that your school's leadership values the teaching of Computer Science in the classroom? **[School relevance]**

CS Teaching Skills (Teachers): Observation Rubric (9 Dimensions)

During the teacher mentoring stage, facilitators use an observation rubric to assess Computer Science teaching practice. This rubric evaluates performance across five levels (Advanced, Intermediate, Initial, Basic, or Not Observed) and covers 9 dimensions of teaching practice. Each level and dimension includes observable descriptors to ensure consistent and reliable assessment. The dimensions are:

1. Communication during the learning process
2. Evaluation of the learning process
3. Management of Computer Science concepts
4. Debugging
5. Lead-learner attitude
6. Use of learning logs or metacognitive strategies during lesson closure
7. Management of student questions
8. Management of resilience
9. Adaptation to the educational context

AIDA TPC test: Computational Thinking Skills

To evaluate the development of Computational Thinking skills among students, two versions of the AIDA TPC test were administered by grade level: one for 3rd and 4th grade, and another for 5th and 6th grade.

In the test, the character AIDA asks the student to help by giving instructions to a robot so it can complete a series of tasks involving drawing, collecting objects, and helping animals in distress. The test assesses problem-solving skills related to sequences, variables, loops, and complex loops. Students must select the set of instructions that correctly solves each problem.

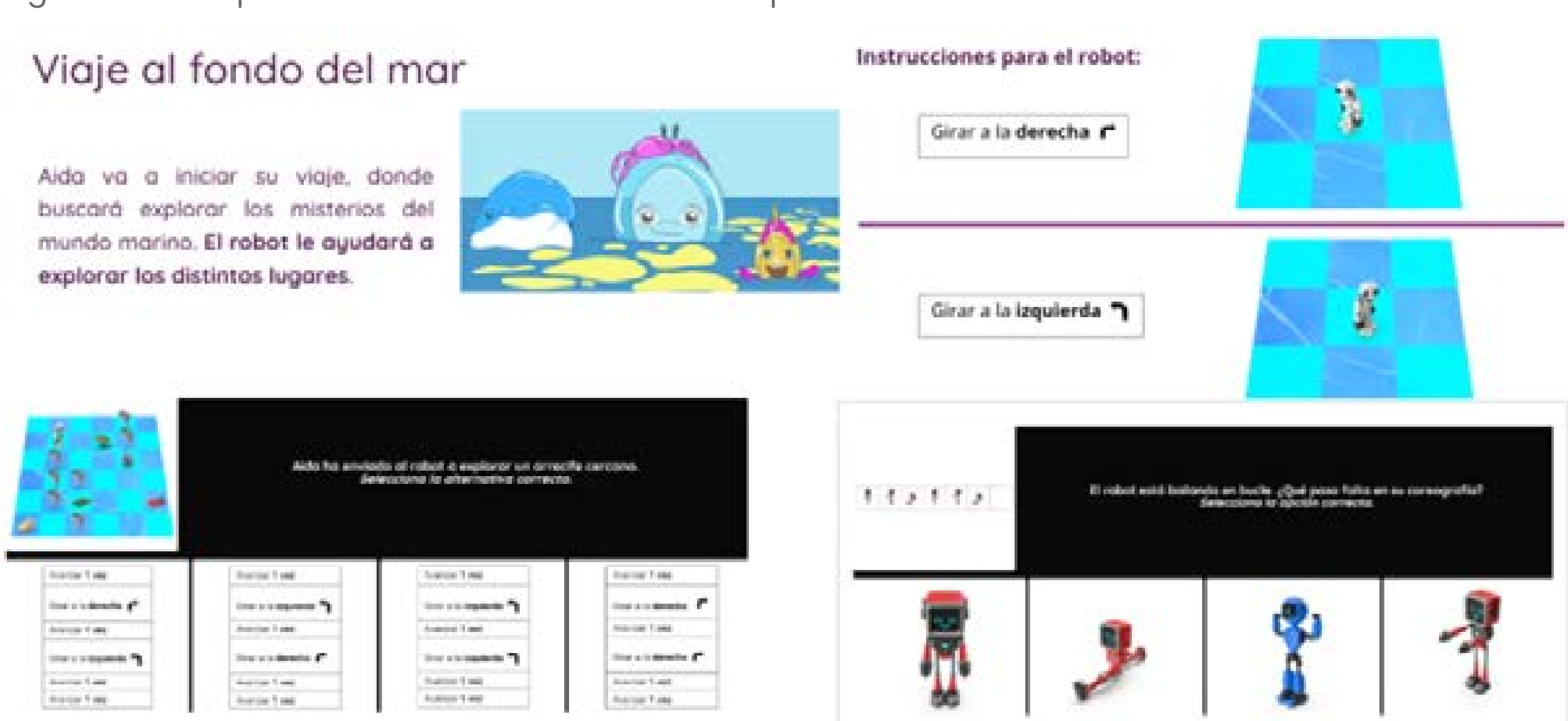
A key advantage of the AIDA TPC test is that it measures Computational Thinking without requiring prior knowledge of a formal programming language.

The test was developed in 2023 and underwent a pilot application (N = 88), followed by face-validity assessment by experts and classroom teachers (N=5), and an experimental application (N=458). At each stage, item properties were examined in terms of difficulty and item-test correlation.

The test was administered online in each school's computer lab through a link on Quizzies platform. Teachers received a 30-minute induction session to conduct the application correctly.

Figure 10: illustrates examples of AIDA TPC instructions and questions.

Figure 10: Examples of AIDA TPC instructions and questions



4.2.5 Methodological Considerations and Scope of the Evidence

The interpretation of results must consider limitations inherent to a quasi-experimental design, which is a methodologically appropriate choice given the context of this program. In this sense, the combined use of Difference-in-Differences and Propensity Score Matching (PSM) enabled a reasonable estimation of intervention effects. The main methodological considerations and the scope of the evidence are outlined below.

- In the case of students, grade-level coverage was unequal. In treated schools, only grades taught by trained teachers were assessed, whereas in the control group the application was census-based across grades. On average, treated schools covered between 1.7 and 1.9 grade levels per school, compared to almost four grade levels in the control group, which could introduce a composition bias between observations.

- For teachers, selection was non-random in both groups: in treatment, only those who participated in the training were assessed; in control, those defined by leadership teams were assessed. This kind of asymmetry was corrected through PSM, balancing observable characteristics between groups.

- Because it was not possible to track individual students over time, measurement was conducted at the classroom or grade level (unit of analysis), using the average proportion of correct answers per group as the outcome variable. While this approach preserves comparability across waves, it limits the ability to capture within-classroom variability and individual-level effects.

- Although matching procedures were applied at the school level, baseline imbalances persisted in certain structural variables, such as total enrollment, contracted teacher hours, proportion of Chilean students, or the SNED equality-of-opportunity indicator. These variables were therefore included as controls in the estimation models to reduce residual bias.

The following section presents the estimation models used to strengthen the validity of results and to test the stability of effects under different modeling assumptions.

4.2.6 Procedures and Mechanisms Implemented to Address Potential Biases

A key step was analyzing baseline balance between treated and control classes in Wave 1. In general, the two groups had similar characteristics, suggesting adequate baseline equivalence. However, statistically significant differences were detected in some variables.
To enhance the robustness of the findings, multiple estimation models were specified, incorporating additional controls and combining DiD with PSM. This strategy allowed for greater control of baseline differences, omitted-variable bias, and other confounding factors that could influence the results. Section 5 presents balance-evaluation results.

For both student and teacher estimates, six models were defined and applied.

- **Model 1 – Simple DiD:** Estimates program effects without additional controls.

- **Model 2 – DiD with imbalance controls:** Replicates Model 1 but adds control variables that show baseline differences between treatment and control groups.

- **Model 3 – DiD with expanded controls:** Incorporates all available control variables, including those with baseline imbalance and those that vary over time.

- **Model 4 – DiD + basic PSM:** Combines DiD with PSM-matched schools only, incorporating as controls the variables with baseline imbalance. This model is estimated on the matched sample considering Waves 1, 2, or 3 depending on data availability.

- **Model 5 – DiD + PSM with unbalanced controls:** Replicates Model 4 but adds additional controls for variables exhibiting baseline imbalance.

- **Model 6 – DiD + PSM with expanded controls:** Replicates Model 5 but includes all observable control variables, providing the most conservative specification

Photo: Girls in ICT Day at Colegio Haras Los Cóndores in San Bernardo, 2025.

5. RESULTS AND IMPACT ESTIMATION

Below are the main evaluation results, focusing specifically on the **impact of the initiative on teacher knowledge and on the development of Computational Thinking skills among students.**

HIGHLIGHTED ASPECTS

Impact on Students
Results confirm a positive, statistically significant, and increasing impact on the development of Computational Thinking skills among students who participated in the Program. Specifically, participation in IdeoDigital Basic led to an average increase of 7 percentage points in the proportion of correct answers. This impact is significant at the 99% confidence level across all model specifications used, with magnitudes above 0.8 standard deviations (SD) in all of them

Impact on teachers
The Program strengthened teachers' capacities, thereby validating the implementation design of IdeoDigital Basic. Specifically, positive and statistically significant effects were observed in Computer Science knowledge, perceived ease of technology use, and advanced technological skills. In addition, digital anxiety exhibited a downward trend among participating teachers. In addition, Digital Anxiety shows downward trends among program participants.

Relevance of the Model
The evaluation confirms the effectiveness of the intervention for both students and teachers in the evaluated group. It is relevant to mention that strengthening teacher capacities translates into improvements in students' skills thus validating the program's causal chain. In this line, IdeoDigital emerges as a viable and effective strategy for integrating Computer Science into the school system.

Results for Principals
School leaders in the treated group presented scores somewhat lower than expected under a no-intervention scenario, both in terms of integrating Computer Science into institutional planning instruments (PEI/PME) and in initial intention or interest. These results should be interpreted with particular caution given sample size limitations and attrition during the follow-up period.

Considerations, Implications, and Limitations

Although evidence from this evaluation indicates positive, statistically significant, and increasing effects, it is important to recognize that the scope of this evidence is limited to the cohort analyzed, due to the assumptions and constraints inherent in a quasi-experimental design. However, having this type of evidence makes it possible to understand that the Program is a feasible and effective alternative for training citizens capable of understanding, creating, and navigating a digital world.

As mentioned above, results of this evaluation apply only to the specific measured cohort and cannot be extrapolated to the program's total universe.

Results for school leadership teams must be interpreted with particular caution given the high level of sample attrition (28% in the treated group) and lack of statistical power in the estimates.

Because it was impossible to identify students individually over time, measurement was conducted at the classroom or grade level. While this approach maintains comparability across waves, it reduces the ability to capture within-classroom variation and control for individual effects; therefore, the evaluation could not estimate which factors moderate effects.

5.1 Results in students' Computational Thinking

5.1.1 Sample Balance

As mentioned above, a fundamental step in impact evaluation is the analysis of baseline sample balance, which compares treatment and control groups before the program intervention. For class characterization, variables collected through student questionnaires were used (averaged by grade level within each school).

In general, most baseline characteristics are comparable between treatment and control groups. However, statistically significant differences were observed in the number of students per class, proportion of Chilean students, whether the grade corresponds to 5th grade, the SNED equality-of-opportunities indicator, number of contracted teacher hours, total school enrollment, and 4th-grade enrollment. These variables were included as controls in the estimation models used.

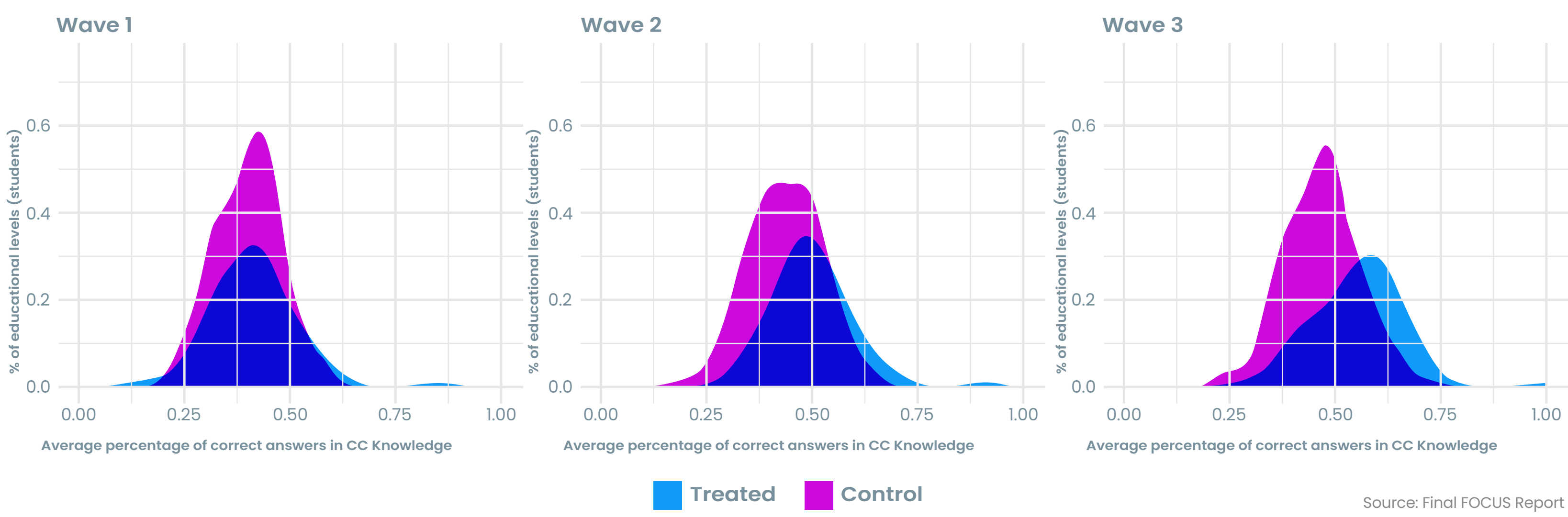
5.1.2 Descriptive longitudinal analysis

When comparing the distribution of the average proportion of correct answers per grade level across Waves 1, 2, and 3 (distinguishing between treated and control schools), it is observed that, compared to Wave 1, the distribution of treated schools progressively shifts to the right in Waves 2 and 3, indicating an improvement in the average proportion of correct answers over time. In contrast, the control group's distribution shows less variation between waves.

The following chart shows the distribution of the average proportion of correct answers per grade level in Waves 1, 2, and 3, differentiating between treatment and control schools.

Photo: The programming party! With Equinix, Pudahuel, Santiago, Chile 2025

Figure 11: Examples of instructions and questions – AIDA TPC test (reference)



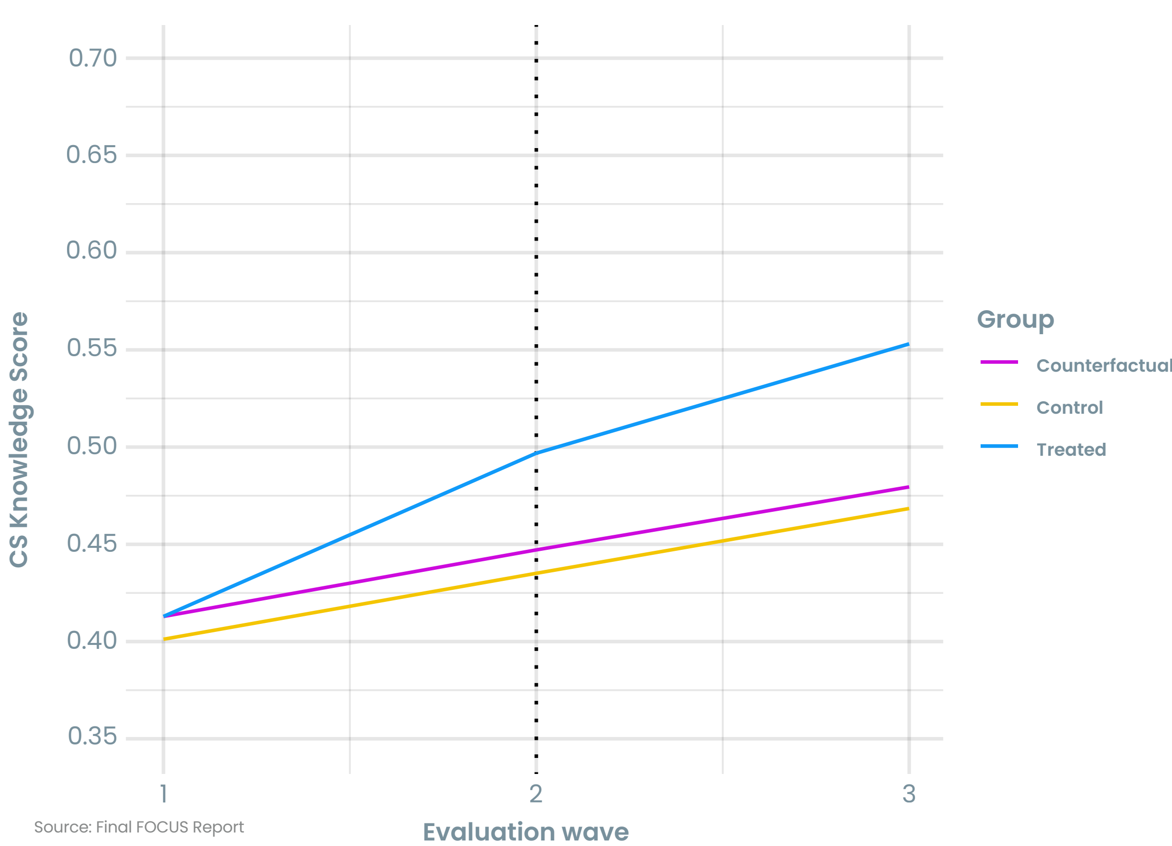
Source: Final FOCUS Report

The progressive shift shown in the previous chart has a direct counterpart in the following figure, which shows the evolution of the average proportion of correct answers for treatment and control groups between Wave 1 and Wave 3.

In both cases, an increase in performance is observed, indicating an overall improvement in results over time. The red line corresponds to the counterfactual scenario, i.e., the hypothetical trajectory the treated group would have followed in the absence of the program. The effect attributable to the intervention is interpreted as the difference between the observed value in the treated group and the counterfactual value in Wave 3.

It is worth mentioning that the estimated effect in Wave 2 increased by Wave 3, from 0.04 to 0.07 standard deviations, which would imply that participation in the program is associated, on average, with an increase of 7 percentage points in the proportion of correct answers.

Figure 12: Average proportion of correct answers (students)



Source: Final FOCUS Report

5.1.3 Impact Estimation

Figure 13 shows the results of three DiD models comparing the three measurement waves(corresponding to Models 1, 2, and 3 specified in Section 4.2.7).

Figure 14 presents the results of Models 3, 4, and 5, corresponding to the combination of DiD and PSM. In addition, this figure reports two sets of comparisons: the first considers only Waves 1 and 3; the second includes all three waves, allowing observation of the effect trajectory over time. Each comparison is estimated in three versions: one without controls, one controlling for variables with baseline imbalance, and a third that adds all available control variables.

Figure 13: Estimated results for standardized outcome (various models). Up to Wave 3

	DiD		
	Model 1	Model 2	Model 3
Standard deviations (wave 2)	0.754** (0,135)	0,613*** (0,129)	0,559** (0,134)
Standard deviations (wave 3)	0.846** (0,167)	0,846*** (0,158)	0,789** (0,159)
Control by variables with initial imbalance	No	Yes	Yes
Control by variables without initial imbalance	No	No	Yes
Observations	681	681	681

Standard errors in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

Source: Final FOCUS Report

Figure 14: Estimated results for standardized outcome (various models). Up to Wave 3

	DiD + PSM Wave 1 & 3			DiD + PSM Wave 1, 2 & 3		
	Model 3	Model 4	Model 5	Model 3	Model 4	Model 5
Standard deviations (wave 3)	1,151** (0,223)	1,137*** (0,225)	1,065*** (0,229)	0,650*** (0,173)	0,622*** (0,179)	0,583*** (0,176)
Control by variables with initial imbalance	No	Yes	Yes	No	Yes	Yes
Control by variables without initial imbalance	No	No	Yes	No	No	Yes
Observations	416	416	416	504	504	504

Source: Final FOCUS Report

Robust standard errors in parentheses

The values in Figure 14 indicate that the program's effect not only persists but increases over time (a constant increase in observed values across waves). When incorporating Propensity-Score Matching (PSM), as shown in Figure 16, this trend is confirmed and strengthened: when comparing only Waves 1 and 3, the estimated impact reaches between 1.065 and 1.137 standard deviations (SD). When Wave 2 is included as an intermediate measurement, the impact remains positive and significant, with values between 0.583 and 0.650 SD.

In sum, it is possible to establish that results demonstrate a sustained and increasing positive effect of the program on the performance of students in treated grades and schools. The impact is already observable a few months after implementation (Wave 2) and intensifies in the subsequent measurement (Wave 3), suggesting that the learning processes and teaching practices promoted by the program generate cumulative and persistent effects over time.

5.2 Results in Digital Anxiety, Technology Acceptability, and Techno-Pedagogical Competence among Teachers

5.2.1 Sample Balance

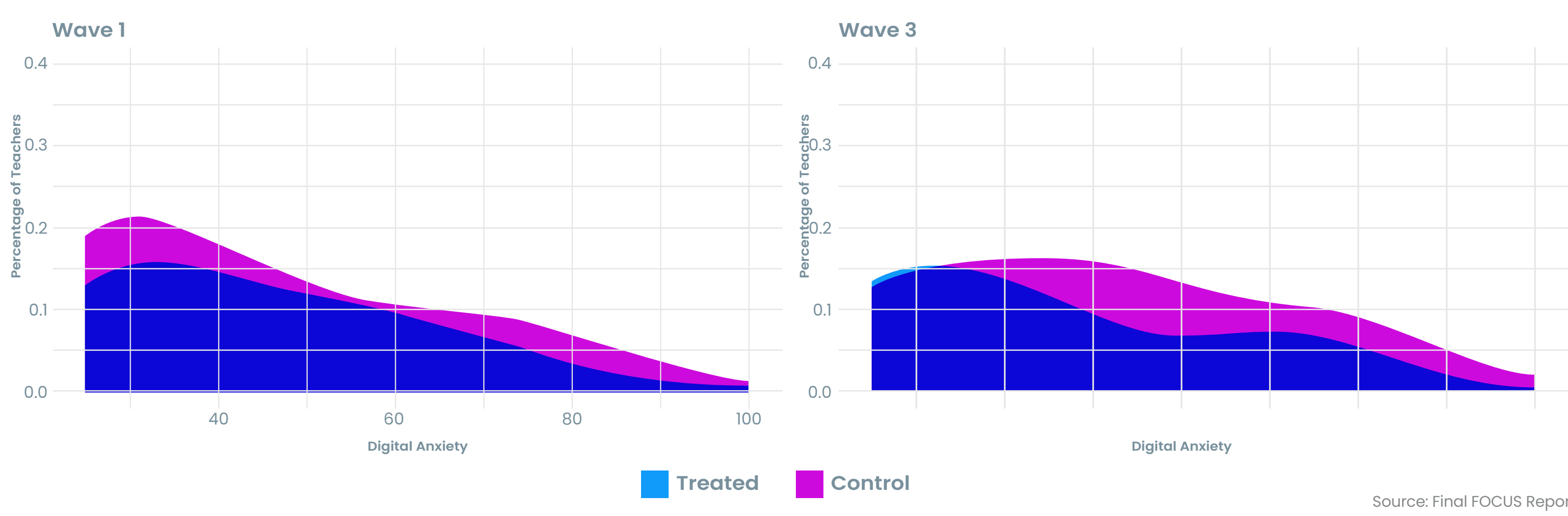
In general, institutional characteristics and teaching staff characteristics are comparable between the two groups. Statistically significant differences are observed in acceptability of technology (overall and in the subdimensions of interest and perceived usefulness) and in teachers' basic competencies, with higher baseline values in the treated group. These variables were incorporated as controls in estimation models to avoid potential biases associated with baseline differences.

5.2.2 Descriptive Longitudinal Analysis

Digital Anxiety

When examining the distribution of results across waves, a similar pattern is observed in both groups, with a slight rightward shift in the Wave 3 distribution for control-group teachers, indicating an increase in digital anxiety among this group.

Figure 15: Distribution of the Digital Anxiety variable in teachers (Pre–Post) for treatment and control groups

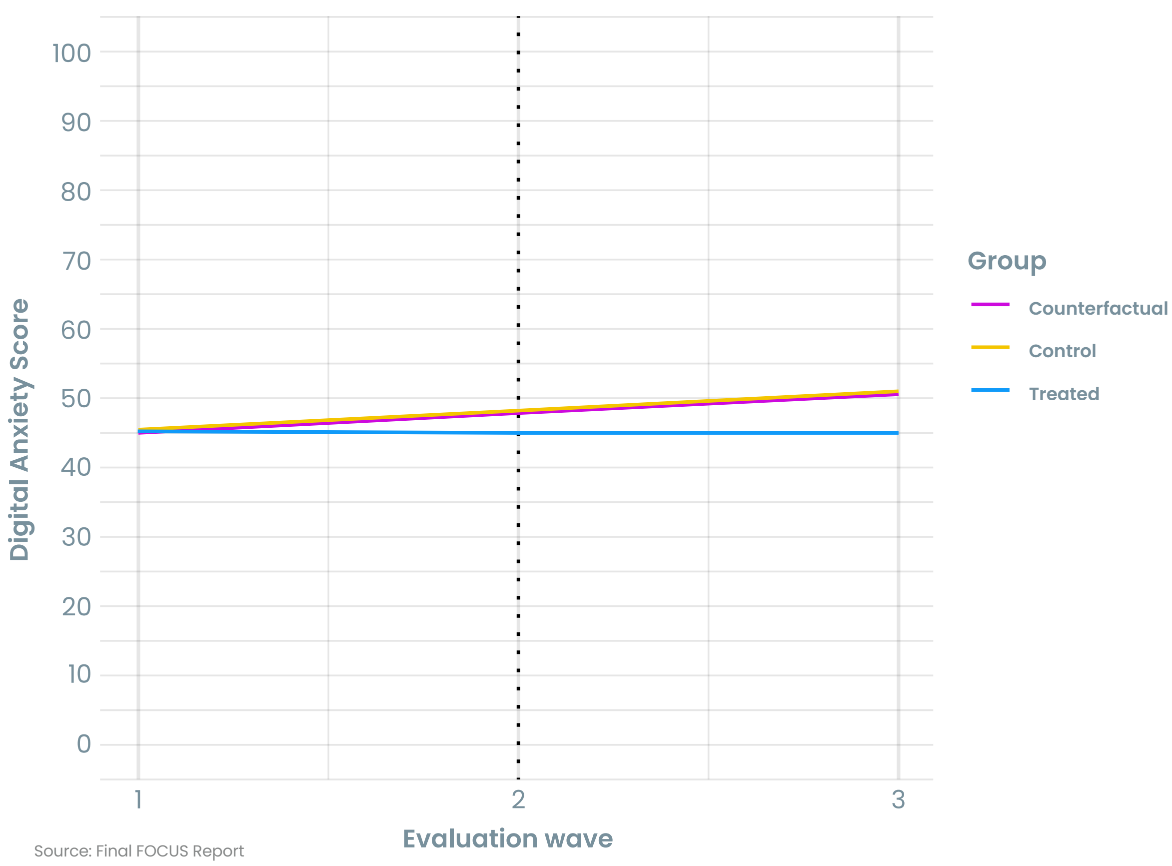


Source: Final FOCUS Report

In line with the above, the evolution of the average digital anxiety score shows that both groups start from similar levels, but in Wave 3, treated teachers remain practically at the same point, while control-group teachers increase their digital anxiety scores. Moreover, compared to the counterfactual trajectory (i.e., what would have been expected without intervention), treated teachers show lower-than-expected levels of digital anxiety.

In quantitative terms, this difference is equivalent to an average decrease of 5.86 points in digital anxiety among teachers who participated in the program (this result is contrasted in the DiD models to assess statistical significance and control for other factors).

Figure 16: Results for the control group, treatment group, and counterfactual (hypothetical) for Digital Anxiety



Source: Final FOCUS Report

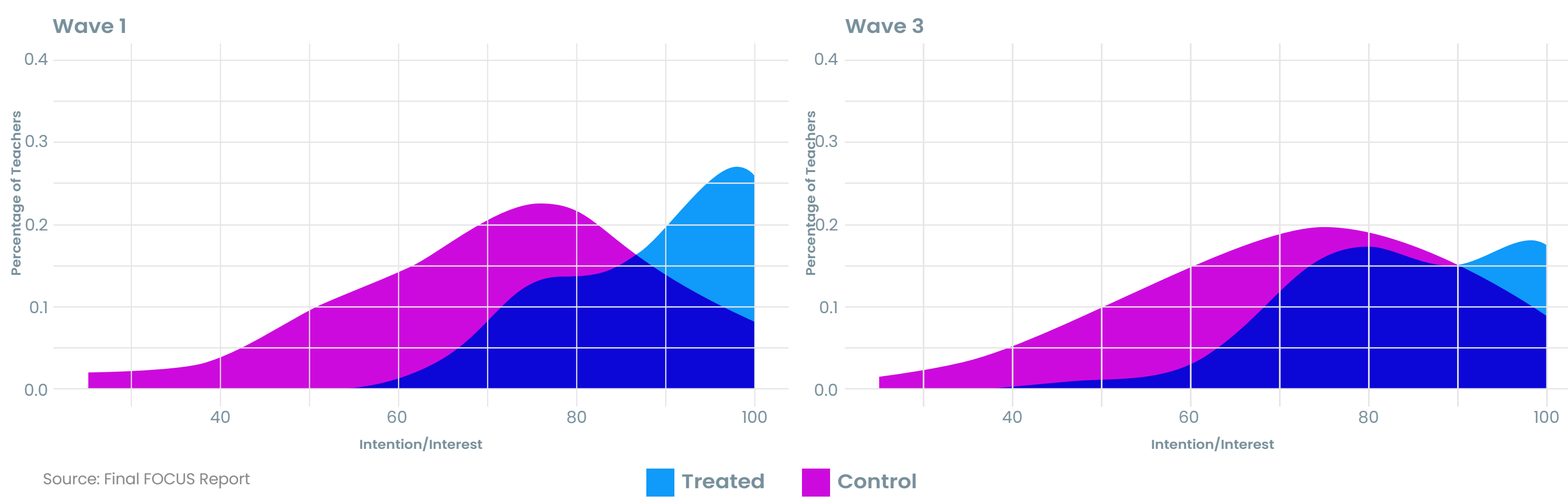
Summit "Ideando el Aula" 2023, workshops for teachers.

Acceptability

The technology acceptability indicator is assessed through three subdimensions: Interest/Intention, Perceived Usefulness, and Perceived Ease.

When analyzing Interest/Intention, in Wave 1 the treated group shows higher levels of intention/interest than the control group, with a distribution shifted toward high values; in Wave 3, both distributions converge, showing a reduction in the initial gap and an overall improvement in both groups.

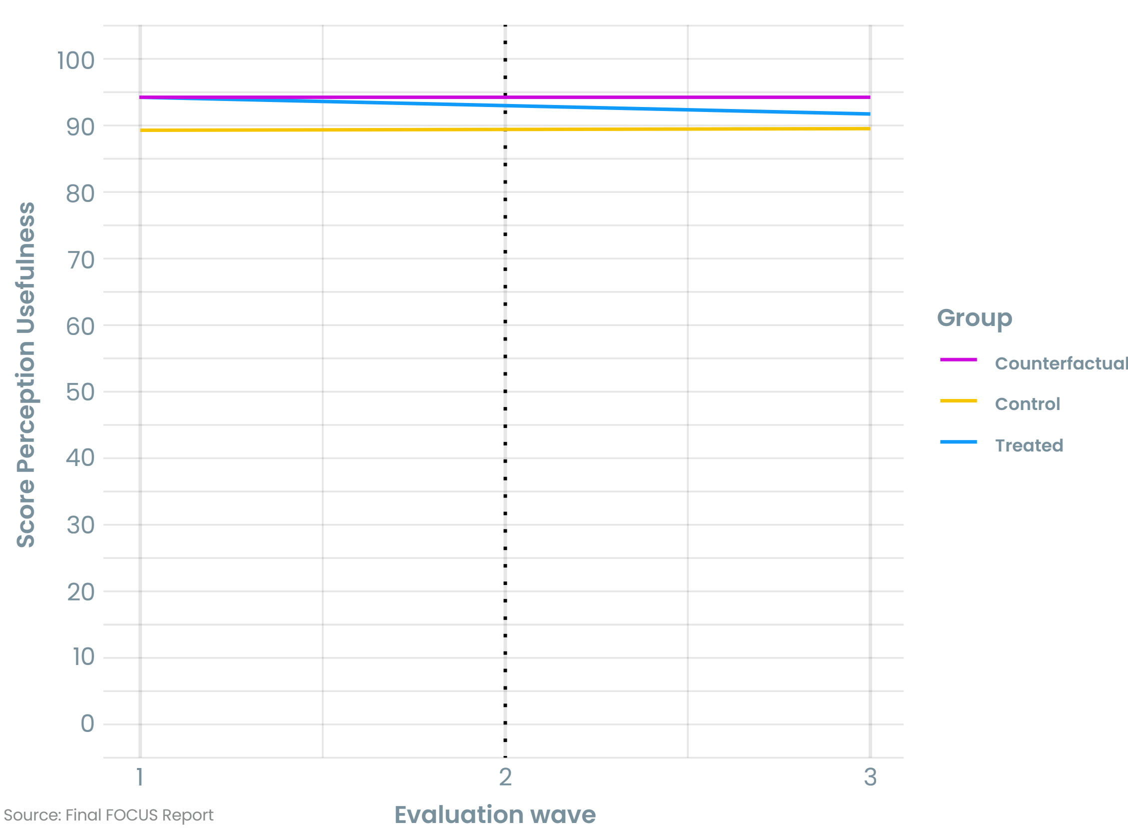
Figure 17: Distribution of the Interest/Intention variable in teachers (Pre–Post) for treatment and control groups



Source: Final FOCUS Report

The counterfactual projection (representing the expected evolution without intervention) suggests that the program does not produce an additional statistically significant increase in this indicator, although both groups maintain high levels of interest/intention over time.

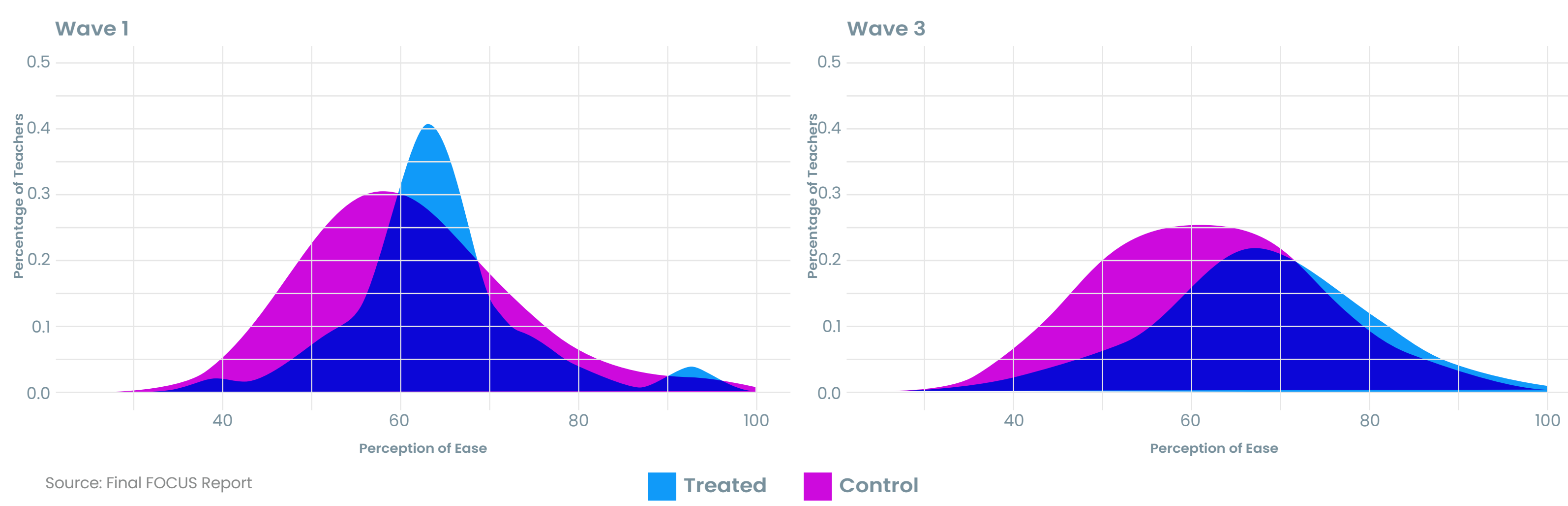
Figure 18: Results for the control group, treatment group, and counterfactual for Perceived Usefulness



Source: Final FOCUS Report

When analyzing Perceived Ease, in Wave 1 treated teachers show a distribution concentrated at medium–to-high values, while the control group shows greater dispersion and a significant proportion of cases with lower scores. In the final measurement, both distributions shift to the right, reflecting an overall improvement in perceived ease. However, the shift is more evident in the treated group.

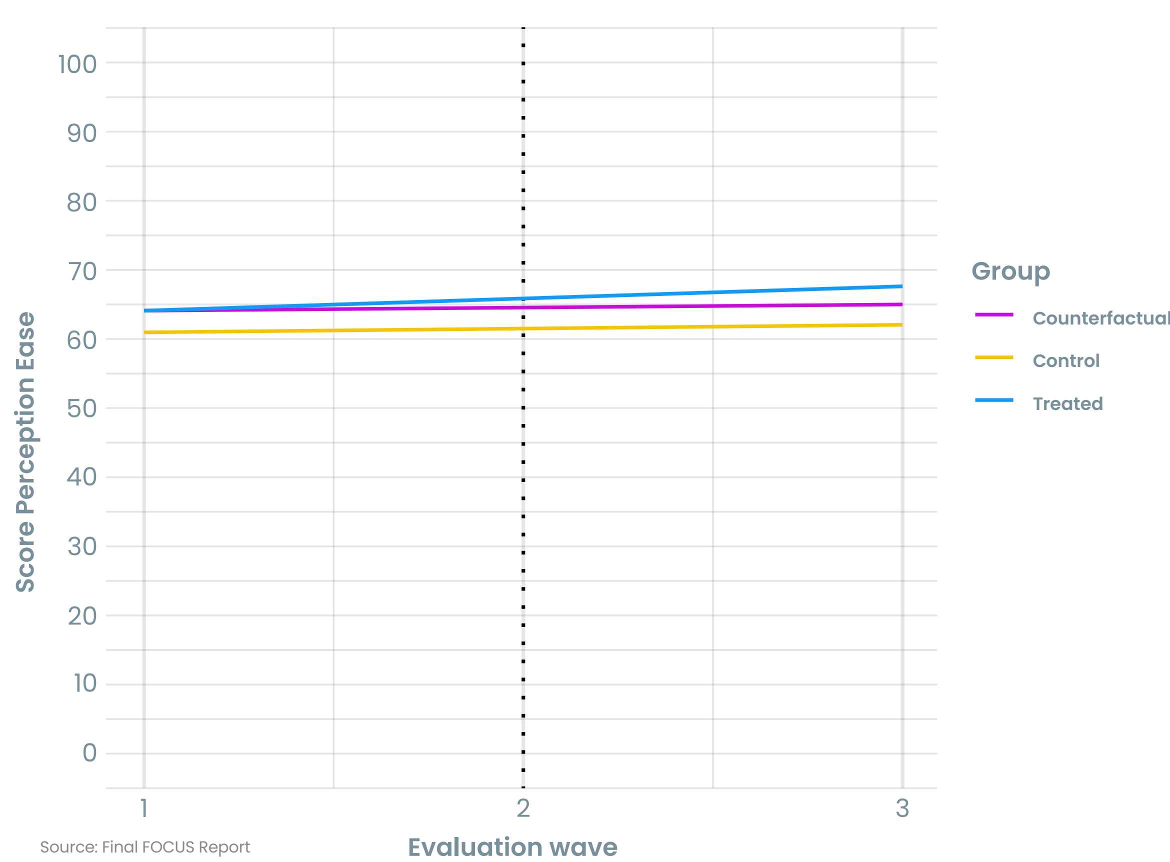
Figure 19: Distribution of the Perceived Ease variable in teachers (Pre–Post) for treatment and control groups



Source: Final FOCUS Report

The counterfactual trajectory (represented by the red line) illustrates the expected evolution of the treated group in the absence of the program; the gap between this trajectory and the observed value indicates a positive effect attributable to the intervention.

Figure 20: Results for the control group, treatment group, and counterfactual for Perceived Ease



Source: Final FOCUS Report

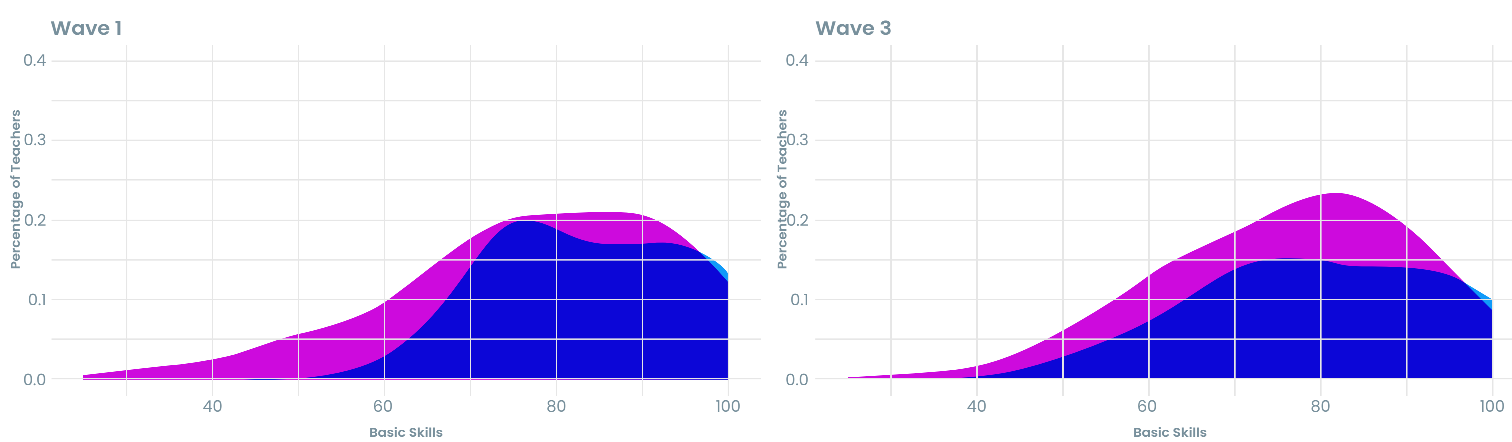
Perceived techno-pedagogical skills

Perceived techno-pedagogical skills are measured through two subfactors: basic skills and advanced skills.

When analyzing basic skills, in Wave 1 both groups show similar distributions, concentrated in the medium–to-high segments of the scale; the control group shows slightly higher density in intermediate values, while the treated group shows a slightly higher proportion of teachers with higher scores.

In the final measurement, distributions maintain shapes very similar to those observed initially. A slight shift to the right is seen in both groups, suggesting generalized improvement, though with an almost parallel pattern across groups.

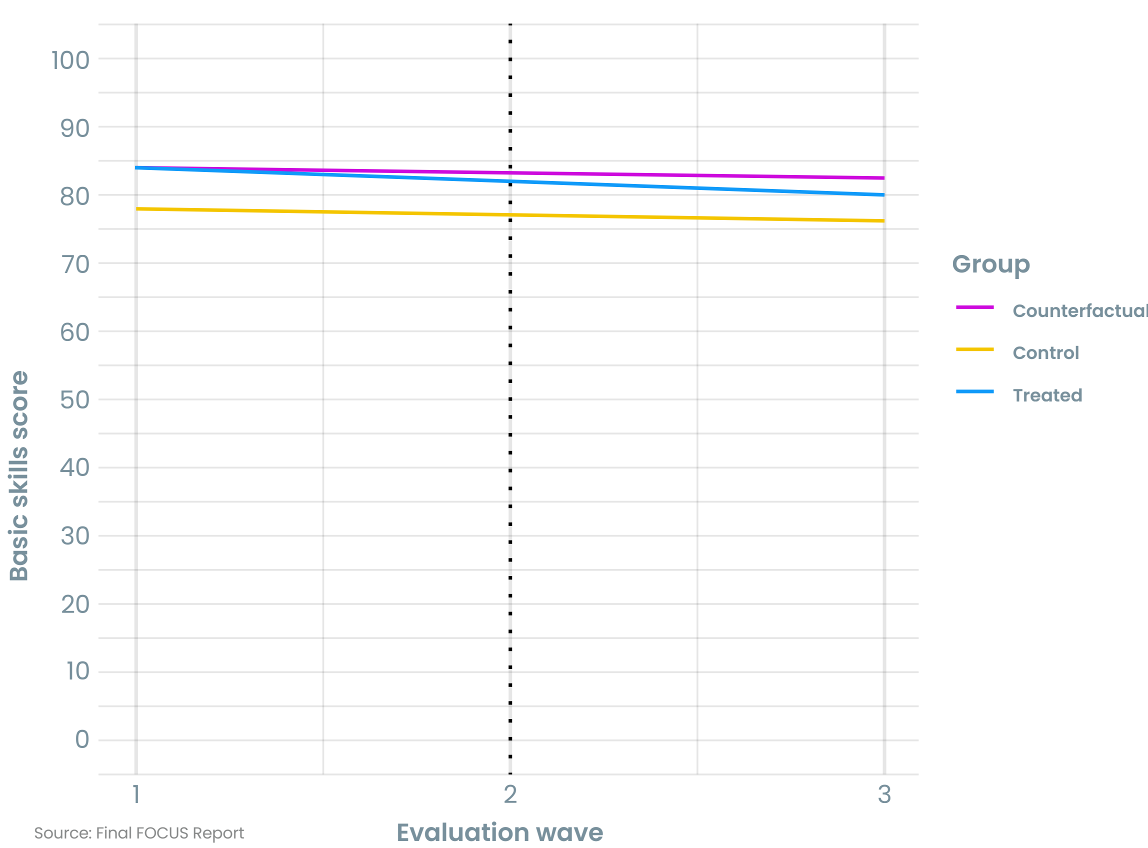
Figure 21: Distribution of Basic Skills in teachers (Pre–Post) for treatment and control groups



Source: Final FOCUS Report

The counterfactual line, slightly above the observed value in the treated group, suggests the intervention did not generate a statistically significant change in this indicator: both groups show a decrease in scores over time, and the gap between groups narrows toward the end of the intervention.

Figure 22: Results for the control group, treatment group, and counterfactual for Basic Skills

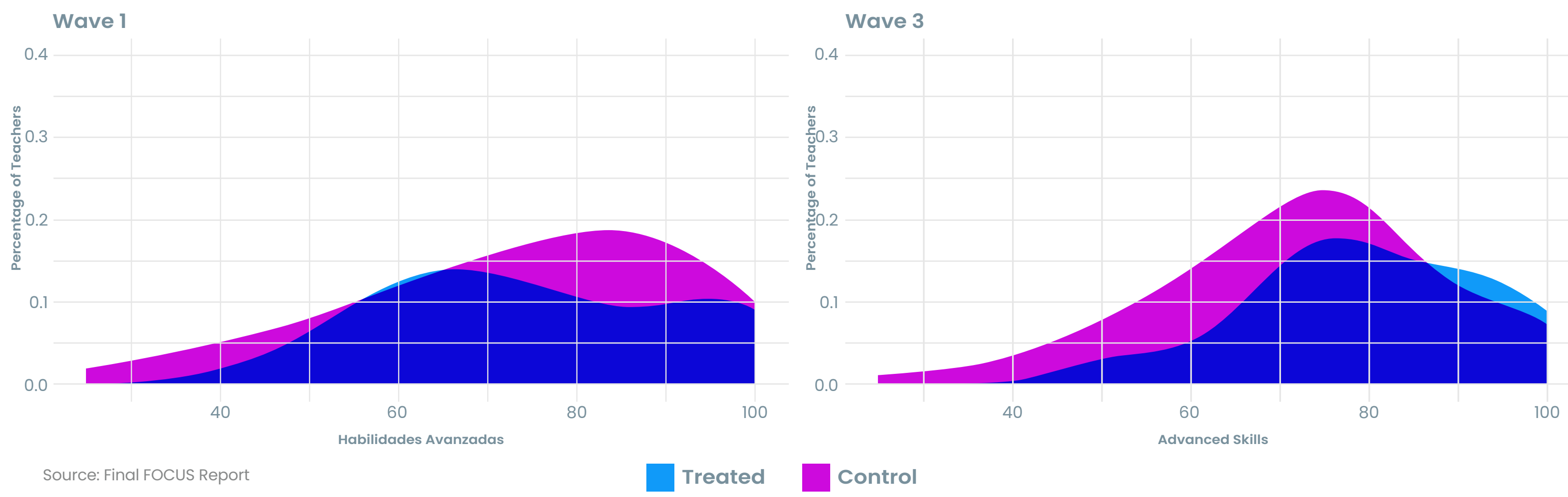


Source: Final FOCUS Report

When analyzing advanced skills, in Wave 1 both groups show similar distributions, with a higher concentration of scores in the medium range and a slight advantage for the control group in the upper segments.

In the final measurement, a shift to the right is observed in both distributions, especially among the treated group, which shows a higher proportion of teachers with high scores. Overlap between curves decreases and the difference in the upper tail widens, indicating a more pronounced improvement in the group that participated in the program.

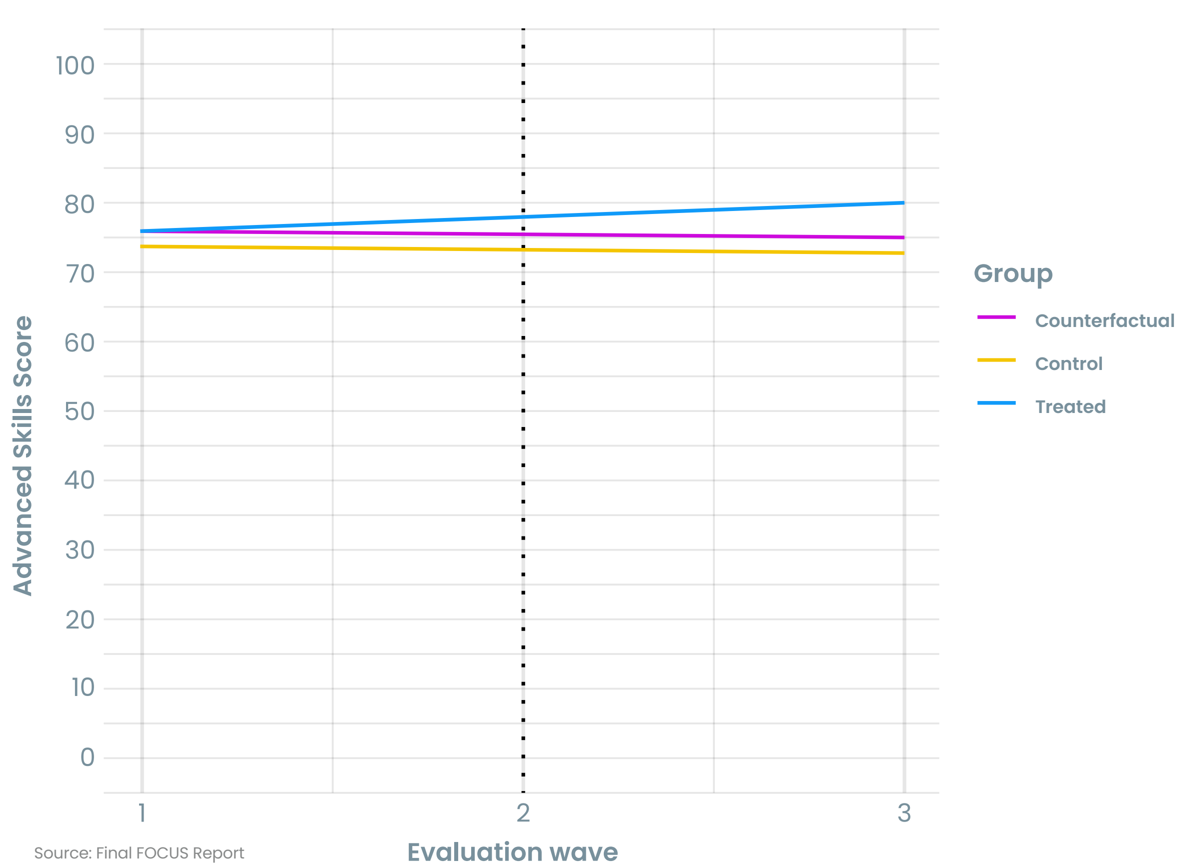
Figure 23: Distribution of Advanced Skills in teachers (Pre–Post) for treatment and control groups



Source: Final FOCUS Report

The counterfactual trajectory (red line) projects the expected evolution of the treated group in the absence of the program; the distance between this projection and the observed value **indicates a positive effect attributable to the intervention: the treated group's average score increases between baseline and final measurement, while the control group shows a slight decline.**

Figure 24: Results for the control group, treatment group, and counterfactual for Advanced Skills



Source: Final FOCUS Report

IMPACT ESTIMATION

Following the logic presented in the student results section, the figures below present two complementary analyses that incorporate the estimation models specified in Section 4.2.7.

Teacher results show differentiated effects depending on the dimension analyzed (Digital Anxiety, Technology Acceptability, or Perceived Techno-pedagogical Skills) and the type of model applied:

- A negative and statistically significant effect is observed in Digital Anxiety (90% confidence level) when using Model 2 (imbalance controls); however, this effect does not hold in any of the other models.
- For technology acceptability, a consistent improvement is observed in perceived ease of use, with positive and significant results in all models. In contrast, the models show negative effects on interest/intention and perceived usefulness, which could reflect more demanding assessments of technology use after participation in the program.
- For perceived techno-pedagogical skills, the results show positive and statistically significant effects on advanced technological skills, which are maintained and reinforced when applying models with more controls (Models 1 to 3). When estimating these effects using models with PSM (Models 4 to 6), positive and statistically significant effects continue, though with slightly smaller magnitudes. On the other hand, no statistically significant changes are observed in basic skills in any of the models used.

Figure 25: Estimation results on standardized outcomes for Digital Anxiety, technology acceptability, and techno-pedagogical competencies (teachers). Waves 1 and 3

	DiD + PSM Wave 1 & 3											
	Model 1	N	Model 2	N	Model 3	N	Model 4	N	Model 5	N	Model 6	N
Digital Anxiety												
Total	-0.309 (0.221)	236	-0.405* (0.227)	234	-0.315 (0.209)	234	-0.268 (0.317)	162	0.293 (0.292)	161	-0.207 (0.273)	159
Technology Acceptability												
Interest/Intention	-0.208 (0.149)		-0.142 (0.142)		-0.143 (0140)		-0.207 (0.245)		-0.176 (0.223)		-0.216 (0.216)	
Perception of Usefulness	-0.271 (0.198)	235	-0.168 (0.196)	234	-0.219 (0.193)	233	-0.124 (0.284)	189	-0.0825 (0.250)	188	-0.0730 (0.248)	188
Perception of Ease	0.298* (0.165)		0.400** (0.155)		0.533*** (0.147)		0.859** (0.312)		0.881*** (0.258)		0.877*** (0.253)	
Technological-pedagogical competencies												
Basic Skills	-0.134 (0-173)	234	-0.102 (0.179)	234	-0.133 (0.187)	232	-0.0585 -0.0585	159	-0.06000 -0.06000	159	-0.0550 -0.0550	157
Advanced Skills	0.341** (0.156)		0.419*** (0.131)		0.500 ** (0.127)		0.377 (0.254)		0.415** (0.170)		0.411** (0.176)	
Control by variables with initial imbalance	No		Yes		Yes		No		Yes		Yes	
Control by variables without initial imba-lance	No		No		Yes		No		No		Yes	

Standard errors in parentheses;
* p < 0.05, ** p < 0.01, *** p < 0.001

Source: Final FOCUS Report

5.3 Results in teachers' Computer Science knowledge

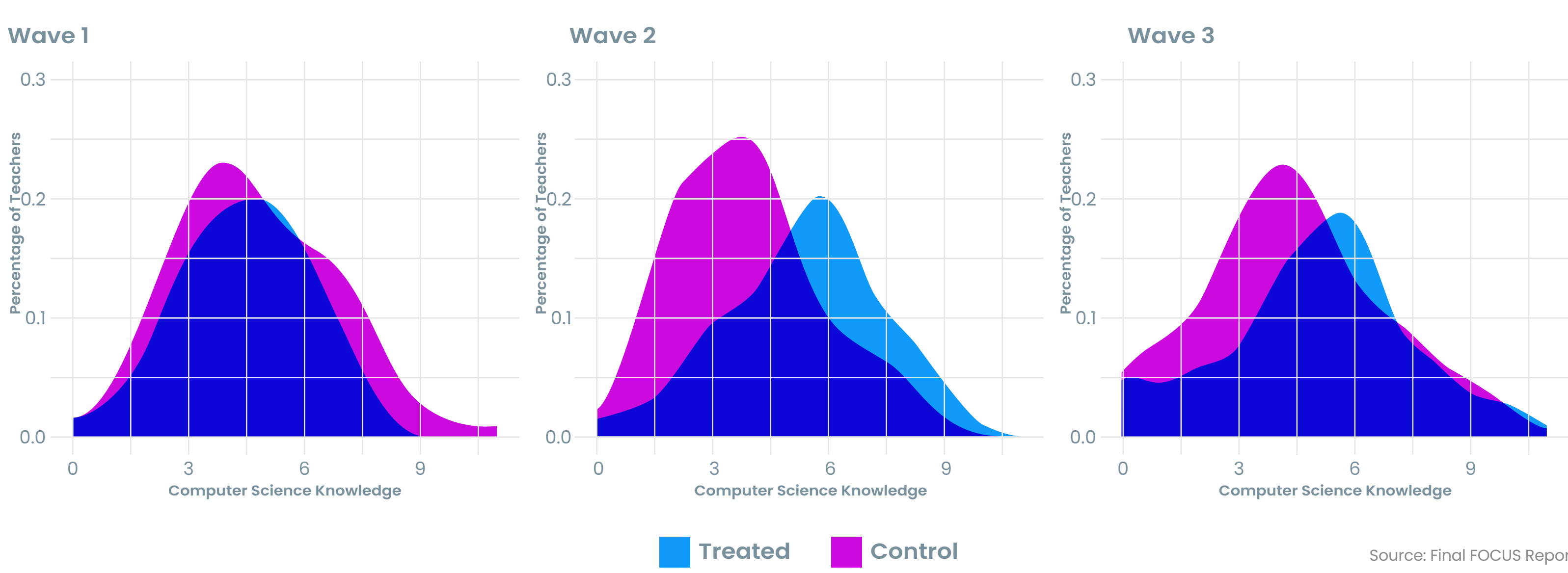
5.3.1 Sample Balance

In general, institutional and teaching staff characteristics are comparable between the two groups. The results indicate that in this estimate, no unbalanced variables were identified in the teacher panel. For this reason, it was not necessary to include Model 2 (DiD with imbalance controls) in the analysis.

5.3.2 Descriptive Longitudinal Analysis

Specifically, teachers in treated schools show improvements in their results, while those in control schools exhibit a decline in scores in the outcome variable (percentage of correct answers). These differences consolidate or remain throughout the evaluation period (Waves 1, 2, and 3). This analysis is evidenced in the rightward shift of the distribution curves for the treatment group in Waves 2 and 3.

Figure 26: Distribution of the Computer Science knowledge outcome variable (Pre–Post) for treatment and control groups

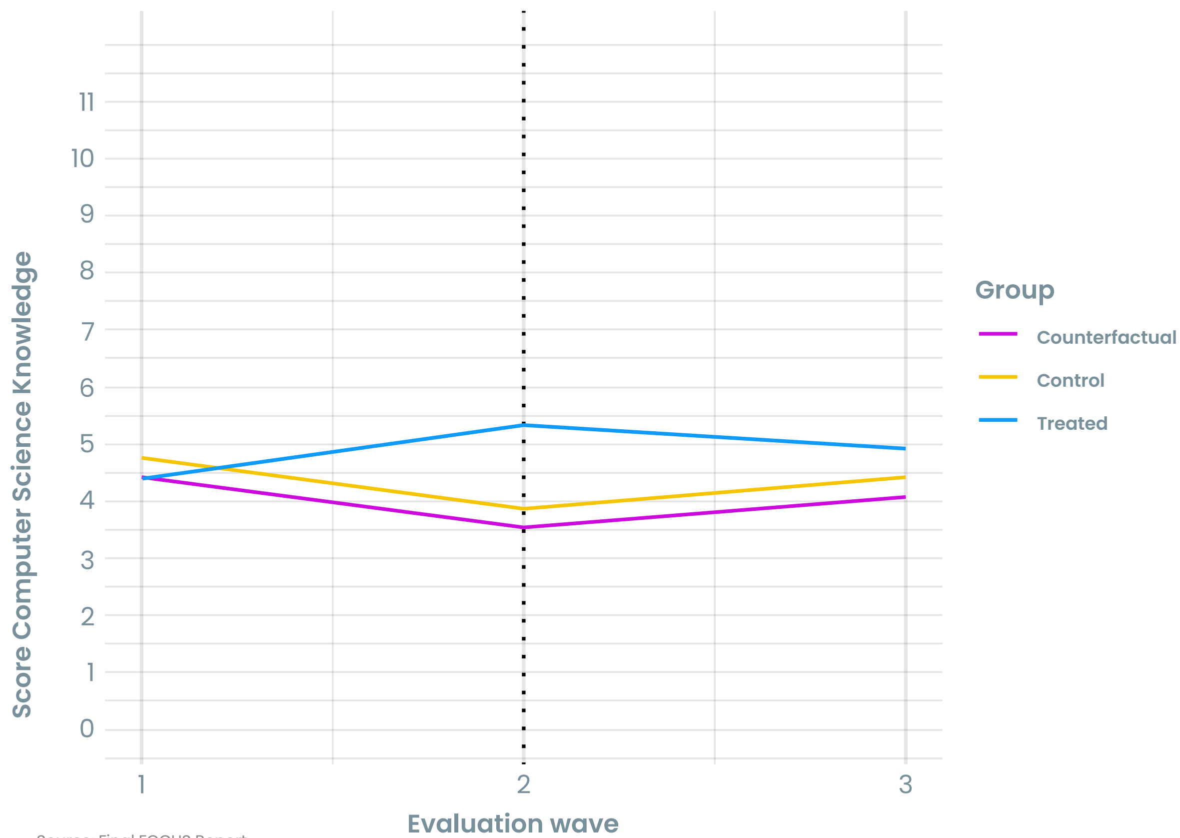


Source: Final FOCUS Report

The previous chart is complemented by the following figure, which shows the evolution of the average proportion of correct answers in the CS Knowledge test between Wave 1 and Wave 3. In Wave 1, the treatment group had a lower average score than the control group; however, from Wave 2 onward, a clear divergence is observed, with the treated group improving significantly, surpassing both the control group and the counterfactual value (i.e., the trajectory that would have occurred in the absence of the program). This difference is also maintained in Wave 3.

The difference between the blue line (treated group) and the red line (counterfactual) represents the effect attributable to the program, corresponding to an Average Treatment Effect (ATE) of 0.84 points on the mean score.

Figure 27: Results for the control group, treatment group, and counterfactual for Computer Science knowledge



Source: Final FOCUS Report

5.3.3 Impact Estimation

The results evidence a significant increase in CS knowledge across measurement waves, with more pronounced effects observed in the treatment groups. The significant increase between Waves 1 and 2 is highlighted and remains positive and statistically significant in Wave 3, suggesting consolidation or maintenance of learning over time.

These results are consistent both in the model without controls and in the model incorporating additional controls without baseline imbalance (Models 1 and 3), reinforcing the stability of the results. As noted above, Model 2 was not estimated because no unbalanced variables were identified.

Models combining DiD with Propensity-Score Matching (PSM) confirm the previous trend, with positive and statistically significant effects (this estimate omits Model 5, since no unbalanced variables were recorded in the group).

Figure 28: Estimation results on standardized outcomes for Computer Science knowledge. Waves 1, 2, and 3

	DiD				DiD + PSM Wave 1, 2 & 3				DiD + PSM Wave 1 & 3			
	Model 1	N	Model 3	N	Model 4	N	Model 6	N	Model 4	N	Model 6	N
Knowledge of Computer Science												
Standard deviation (Treatment x Wave 2)	0.945*** (0.214)	336	0.952*** (0.229)	333	-		-		-		-	
Standard deviation (Treatment x Wave 3)	0.455* (0.233)		0.523** (0.249)		0.698*** (0.251)	252	0.678*** (0.255)	252	0.528 (0.324)	210	0.536* (0.321)	210

Standard errors in parentheses;
* p < 0.05, ** p < 0.01, *** p < 0.001

Source: Final FOCUS Report

5.4 Results of Technology Acceptability in Leadership Teams

5.4.1 Sample Balance

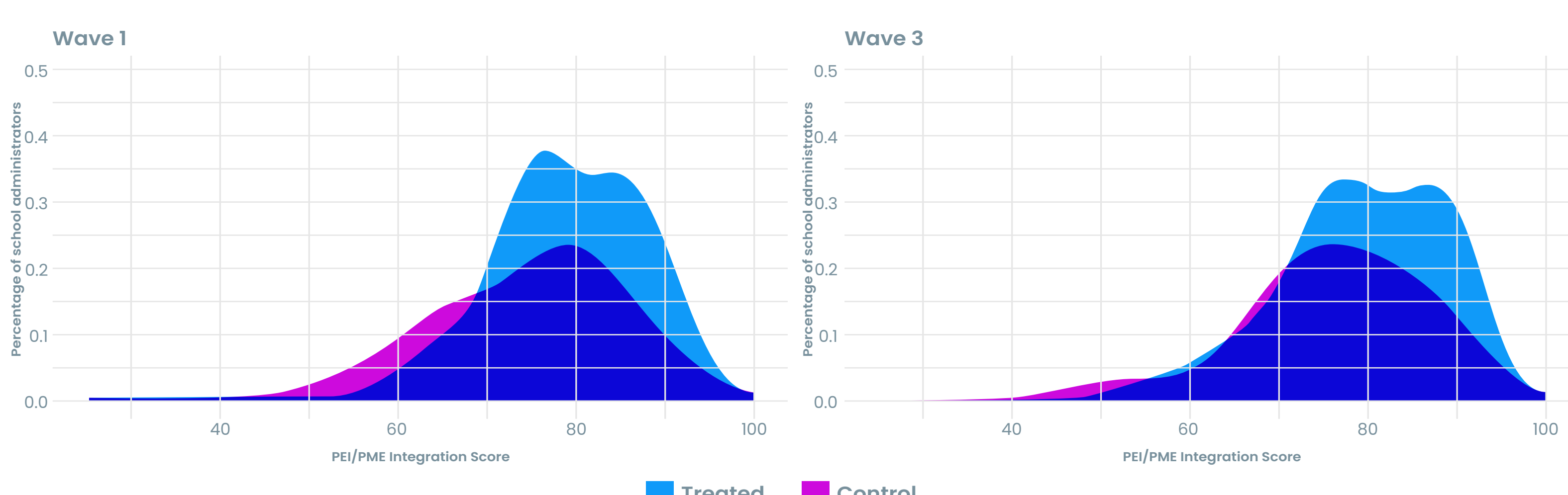
In general, institutional characteristics are comparable between both groups. However, statistically significant differences are observed in the SNED equality-of-opportunities indicator and in 4th-grade enrollment. In the outcome variables, statistically significant differences are also observed, specifically in overall technology acceptability and the subfactors of initial interest and perceived usefulness, with higher values in the treated group. These variables are incorporated as controls in estimation models to avoid bias associated with baseline differences.

5.4.2 Descriptive Longitudinal Analysis

The results for principals in the Technology Acceptability dimension are measured through four subfactors: integration into the PEI or PME, initial interest, perceived usefulness, and perceived ease.

When analyzing integration into the PEI/PME, both groups show similar distributions in Wave 1 , with a slight advantage for the treated group in the upper segments of the scale. In Wave 3, the treated group's distribution shifts toward higher values.

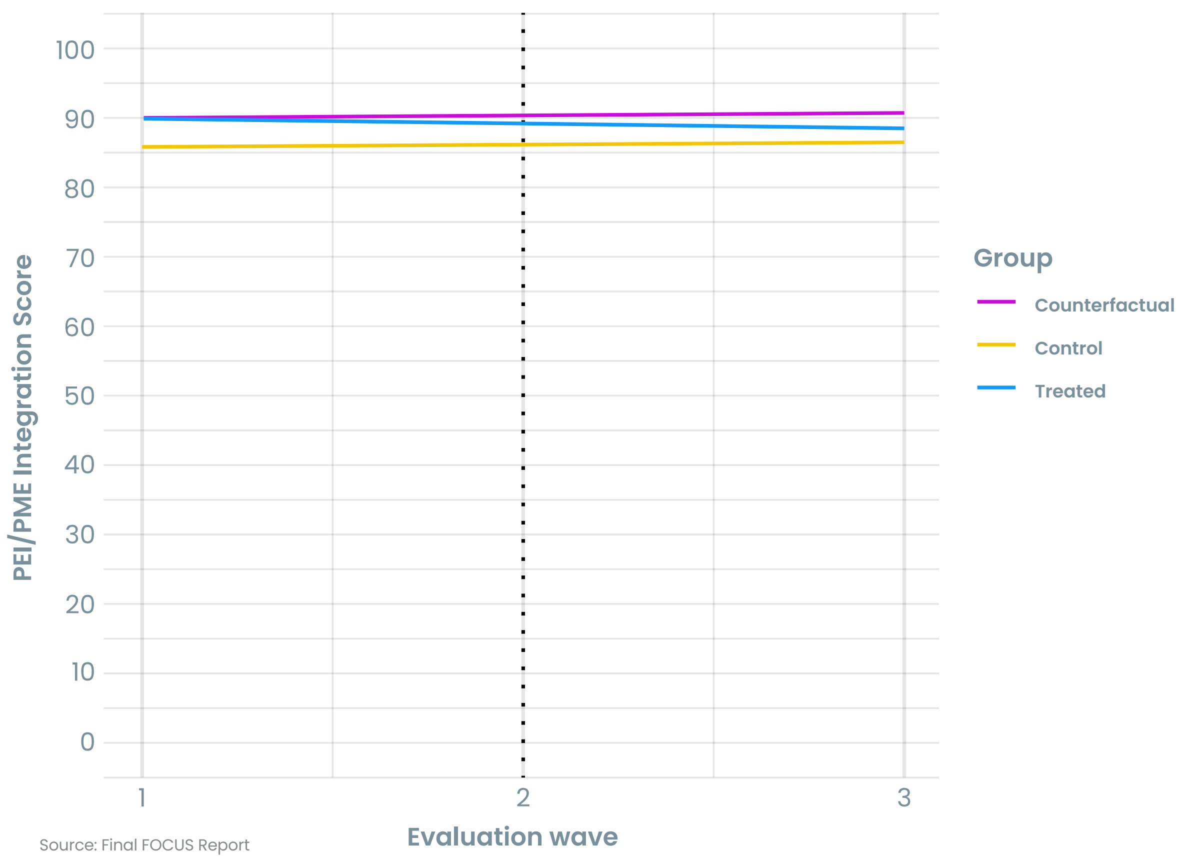
Figure 29: Distribution of Integration into PEI/PME (Pre–Post) for treatment and control groups (leadership teams)



Source: Final FOCUS Report

The counterfactual line (in red) indicates that the observed evolution in treated schools is lower than what would be expected in the absence of the program. However, the descriptive differences between groups are small, making it necessary to contrast them using causal estimation models to determine their statistical significance.

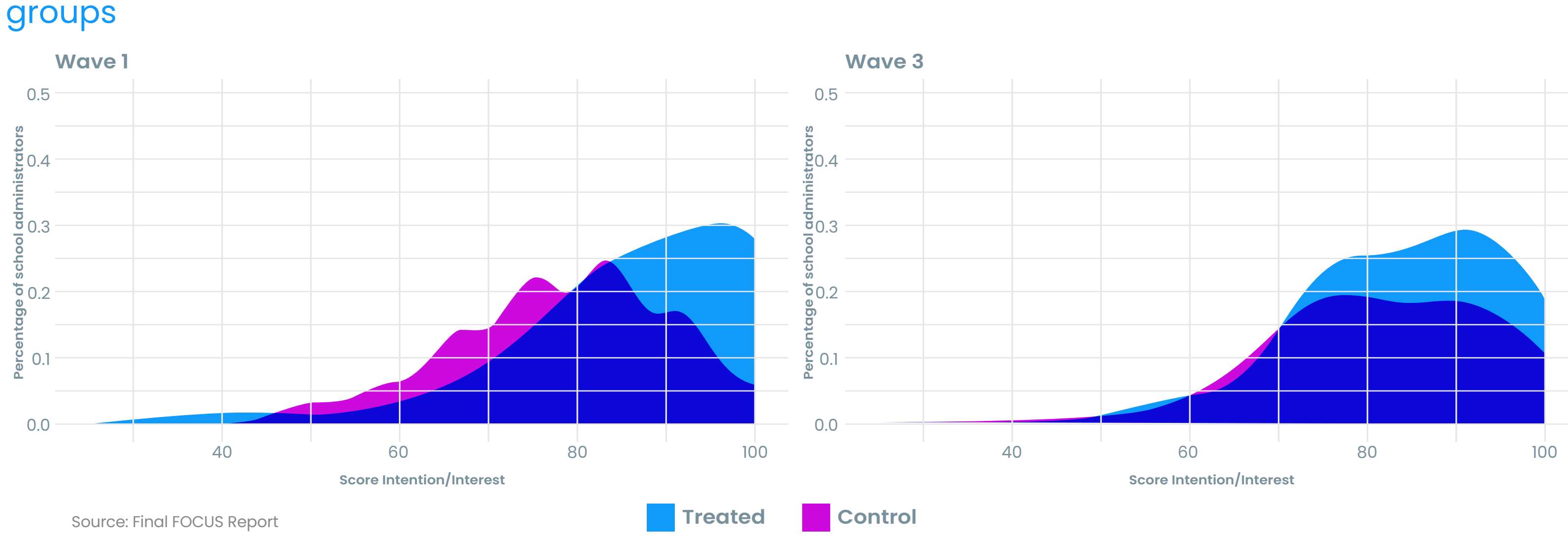
Figure 30: Results for the control group, treatment group, and counterfactual for Integration into PEI/PME (leadership teams)



Source: Final FOCUS Report

When analyzing initial intention/interest, both groups show a similar distribution in Wave1. In Wave 3, the distributions tend to converge, maintaining a high concentration of schools in the upper-range scores.

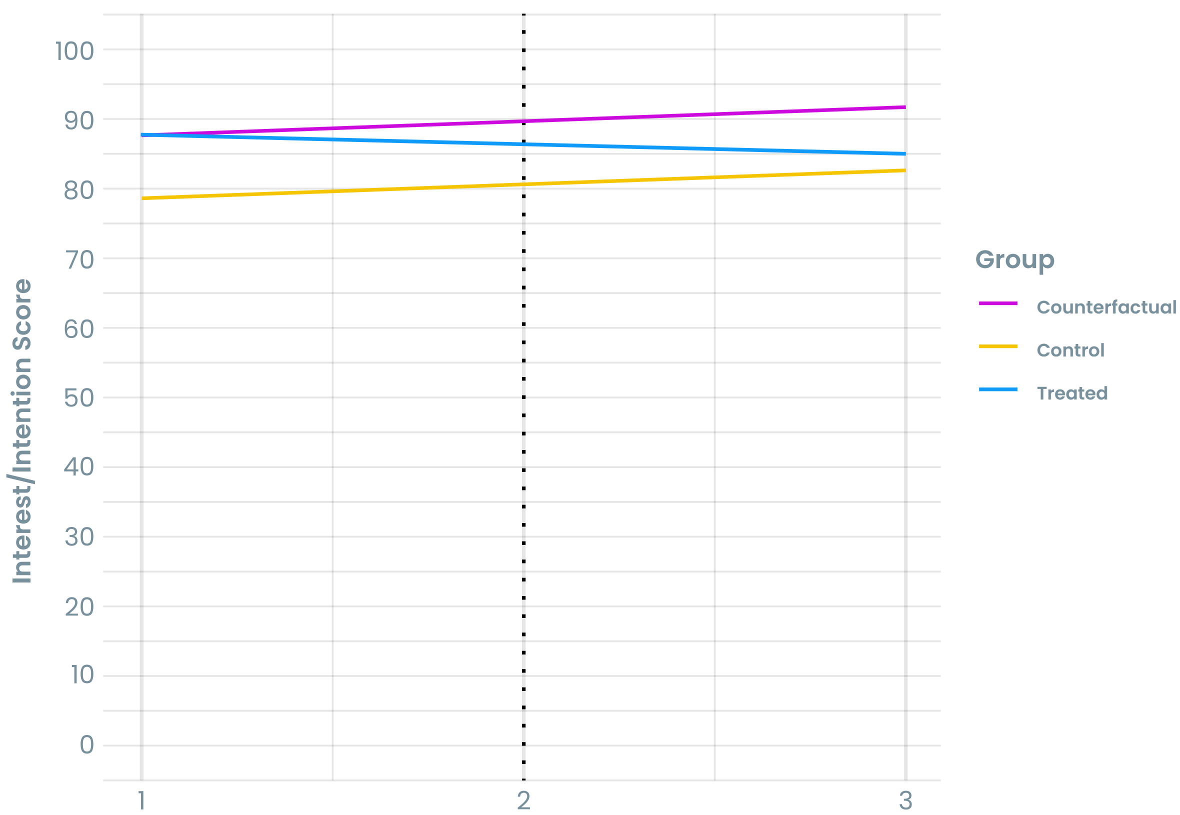
Figure 31: Distribution of Intention/Interest in leadership teams (Pre–Post) for treatment and control groups



Source: Final FOCUS Report

The counterfactual line (in red) shows that observed evolution in treated schools is lower than expected in the absence of the program. However, given that baseline scores for the treatment group are already in the upper part of the distribution. This pattern may reflect a ceiling effect—that is, reduced room for improvement due to high starting levels.

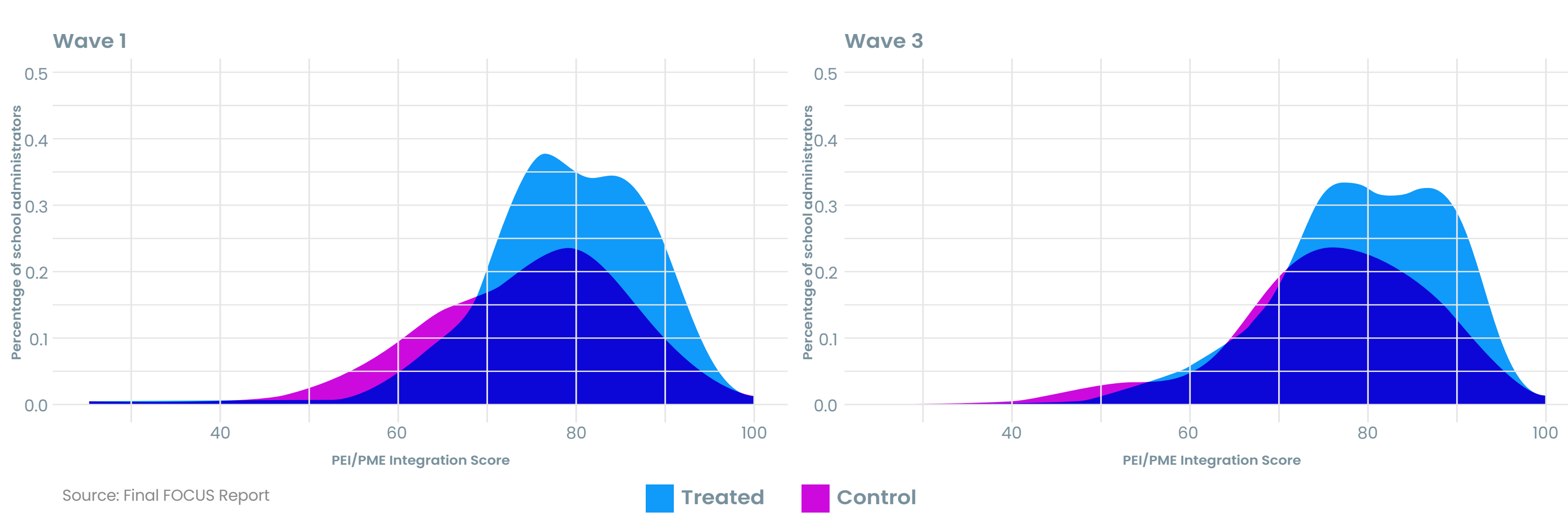
Figure 32: Results for the control group, treatment group, and counterfactual for Interest/Intention (leadership teams)



Source: Final FOCUS Report

When analyzing Perceived Usefulness, the treated group maintains a slightly higher average between waves, but the descriptive gap between groups narrows over time, reflecting improvements in the control group's perceptions and slight stabilization in the treated group.

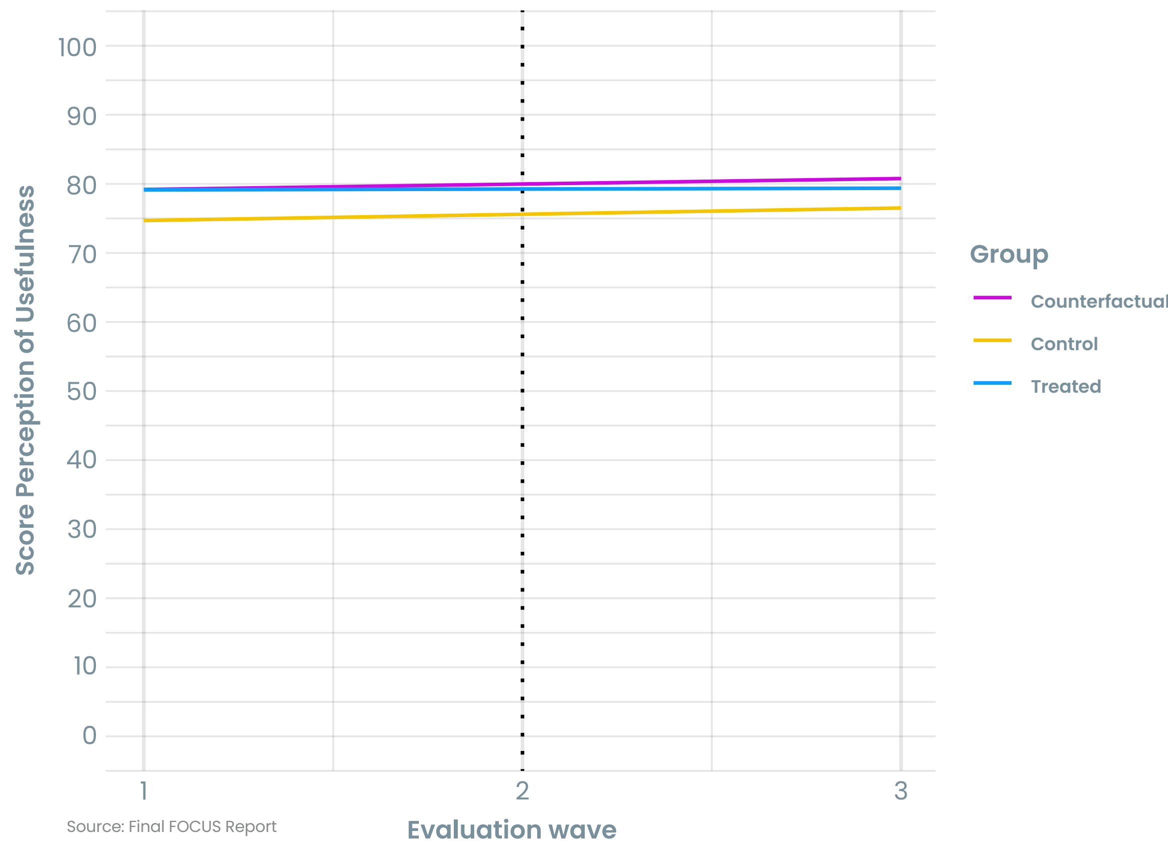
Figure 33: Distribution of Integration into PEI/PME in leadership teams (Pre–Post) for treatment and control groups



Source: Final FOCUS Report

The counterfactual line (in red) shows a slight negative difference: the treatment group's actual scores are below what would have been projected without the intervention. The statistical significance of this difference must be verified in causal estimation models.

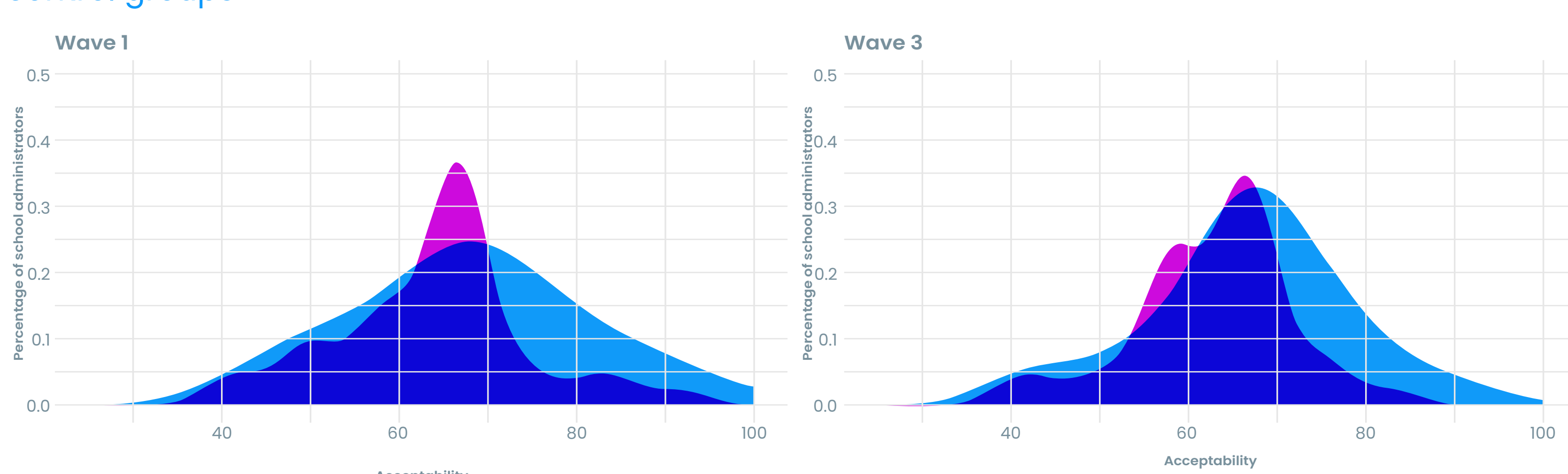
Figure 34: Results for the control group, treatment group, and counterfactual for Perceived Usefulness (leadership teams)



Source: Final FOCUS Report

When analyzing Perceived Ease of use, both groups show similar distributions in Wave 1 , with a slight advantage for the treated group in medium and high segments. In Wave 3, the treated group's distribution remains more extended toward higher values, while the control group's scores become more concentrated around the mean.

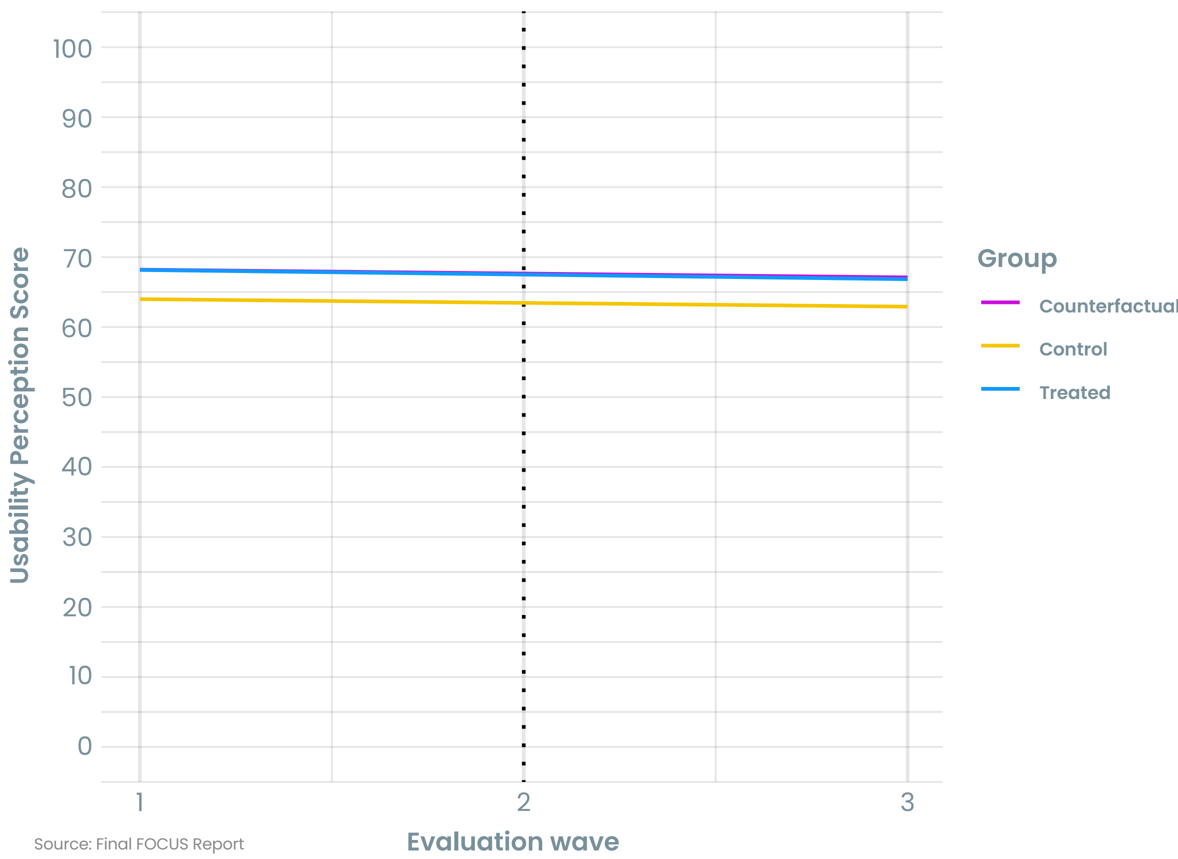
Figure 35: Distribution of Technology Acceptability in leadership teams (Pre–Post) for treatment and control groups



Source: Final FOCUS Report

The counterfactual line (in red) closely overlaps with the observed trajectory of the treated group, indicating the actual evolution does not differ substantially from what would be expected in the absence of the program.

Figure 36: Results for the control group, treatment group, and counterfactual for Perceived Usability (leadership teams)



Source: Final FOCUS Report

5.4.3 Impact Estimation

The results for leadership teams show moderate and non-significant effects across most subfactors associated with technology acceptability. In technology integration into the PEI/PME, coefficients are positive but not statistically significant, suggesting a slight tendency toward improvement that cannot be attributed to the program. The same occurs for Perceived Usefulness and Perceived Usability, where coefficients are also positive and non-significant. By contrast, the Interest/Intention subfactor shows a negative and statistically significant effect in the three models estimated (95% confidence level), even when incorporating imbalance controls and expanded controls.

When combining DiD with PSM, results continue to show no statistically significant program effects for integration into PEI/PME, perceived usefulness, or perceived usability. By contrast, interest/intention shows a different pattern: in Model 4 (without additional controls) a positive and statistically significant effect is observed; however, when incorporating controls and matching (Model 6), the effect becomes negative and marginally significant.

Figure 37: Estimation results on standardized outcomes of Technology Acceptability among leadership teams. Waves 1 and 3

	DiD Wave 1 & 3						DiD + PSM Wave 1 & 3					
	Model 1	N	Model 2	N	Model 3	N	Model 4	N	Model 5	N	Model 6	N
Technology Acceptability												
PEI/PME integration	-0.103 (0.299)		0.0679 (0.313)		0.0991 (0.339)		-0.247 (0.357)	116	-0.181 (0.345)	116	-0.124 (0.354)	112
Interest/Intention	-0.511** (0.227)	158	-0.449** (0.223)	158	-0.500** (0.245)	154	0.432** (0.203)	134	-0.250 (0.268)	136	-0.494* (0.288)	130
Perception of Usefulness	-0.0204 (0.222)		0.134 (0.242)		0.218 (0.229)		0.154 (0.223)	116	0.0335 (0.275)	122	0.179 (0.313)	112
Perception of Ease	-0.154 (0.278)		0.0108 (0.265)		0.101 (0.211)		0.204 (0.236)	124	0.218 (0.306)	130	0.164 (0.334)	120
Control by variables with initial imbalance	No		Yes		Yes		No		Yes		Yes	
Control by variables without initial imbalance	No		No		Yes		No		No		Yes	

Standard errors in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

Source: Final FOCUS Report

Photo: Closing of the pilot "Impulsemos la Educación del Futuro Juntos", Valparaíso, 2025.

6.

CONCLUSIONS, REFLECTIONS, AND RECOMMENDATIONS

Below are the main conclusions, framed by the program's theory of change, the scope of the evidence, and the relevant methodological limitations.

The Model Works in Real Contexts

The evaluation demonstrates that meaningful results and measurable learning gains are achieved in average public schools, working with teachers who had no prior training in Computer Science. This finding is key, as it shows that the model does not rely on exceptional conditions, but is feasible for scaling within the Chilean education system.

The evaluation shows that student effects increase between the second and third measurement, reinforcingKodea's core vision: educational transformation requires continuity, mentoring, and time. The impact is not immediate, but it is sustained.

Kodea's IdeoDigital Model Builds Capacities

The study confirms that the transfer model—from Kodea to ATEs, teachers, and territories—succesfully builds local capacities. This allows schools to continue progressing even without the direct involvement of the central team, which is vital for long-term sustainability and territorial equity.

Findings also confirm that strengthening teacher capacities leads to sustained improvement in students' Computational Thinking skills, thereby validating the program's causal chain.

Evidence Shows Where to Improve

A relevant learning from the evaluation is that the program's impact on school leadership teams is weaker. Consequently, the continuity design of the IdeoDigital program will place greater emphasis on strengthening strategies aimed at leadership teams and on initial teacher education.

Recommendations

In a context marked by the accelerated transformation driven by Artificial Intelligence,, Kodea considers it urgent to advance public policy decisions that prepare teachers to teach 21st-century digital skills. In this sense, the following priority actions are proposed:

National Teacher Training Plan

Establish a mandatory training plan for all in-service teachers—across disciplines and educational levels—covering digital skills, Computational Thinking, and pedagogical integration of technology.

Integration into Teacher Training Degrees

Include at least three courses, as a requirement in initial teacher education,:
a) digital skills leveling;
b) pedagogical use of digital technologies and AI;
c) an advanced elective (e.g., robotics, programming, or data science).

Training Specialists

Allocate scholarships in Chile and abroad to train 800 academics who are experts in teaching digital skills and AI integration, ensuring that every teacher training school has at least two specialists.

Mentoring and Technical Assistance Program

Provide targeted support to schools facing greater management challenges, alongside mechanisms to ensure infrastructure and connectivity, including maintenance and replacement in case of damage or theft.

Educational Piloting Hub

Create a systematic space to test, validate, and scale new teaching practices and emerging technologies, with dedicated funding for short- and long-term research.

Strengthen the Ministry of Education's Center for Educational Innovation (MINEDUC)

Consolidate this unit as a national observatory that systematically measures the impact of educational digitalization on learning and skills development.

IDEODIGITAL IMPACT EVALUATION

Computer Science in the Classroom

Kodea, 2024 - 2025

Kodea authors:
Andreas Hein - Head of Research
Claudio Farías - Research Analyst
Francisca Avendaño - Research Analyst
Larratiz Otamendi - Research Analyst

Focus authors:
Laura Ramaciotti
Antonia Arrendondo
Bernardita Canals
Isidora Fuenzalida

Editing and design (Kodea):
Camila del Solar
Karla Cantuarias
Alejandra Moya