Education Tech Research Dev
https://doi.org/10.1007/s11423-026-10675-2

CULTURAL AND REGIONAL PERSPECTIVES

Impact of an educational robot-based intervention on second-graders computational thinking skills: experimental evaluation of the *Irûmi* program in Paraguay

Emma Näslund-Hadley[1] · Juan Manuel Hernández-Agramonte[2] · Humberto Santos[3] · Pablo Zoido[1]

Received: 30 July 2025 / Accepted: 26 June 2026

Abstract

In this article, we present the impact evaluation of *Irûmi*, an educational robot-based intervention aimed at developing second-grade students' computational thinking (CT) skills in Paraguay. Our results indicate that the program had an effect of 0.089 standard deviations on the students' CT skills, focusing on abilities such as abstraction, algorithmic thinking, and evaluation. These findings suggest that with age-appropriate instructional design, young children could develop CT skills, and that smart devices and electronic toys can contribute to their development. Beyond its effects on CT skills, *Irûmi* generated additional benefits for students. In particular, it increased preferences for science, technology, engineering, and mathematics (STEM)-related toys and promoted greater gender flexibility regarding who can play with them, particularly among girls, while improvements in attitudes toward technology were observed for both boys and girls. In addition to its impact on students, the intervention also affected teachers who implemented the activities. First, *Irûmi* increased the likelihood that teachers would use educational technology in the classroom, including devices not explicitly provided by the program. Second, the program contributed to the development of teachers' CT skills, possibly due to the novelty of the curriculum and methodology in the Paraguayan context. Third, *Irûmi* may have increased teachers' motivation to use technology, thereby strengthening their positive attitudes toward its integration into teaching. Our study contributes to the literature by focusing on a middle-income country, using an experimental design with pre- and post-intervention measures and a large sample of students, and analyzing the impact on both students and teachers.

Keywords Computational thinking · Early childhood education · Educational robots · Randomized controlled trial (RCT) · Paraguay

✉ Emma Näslund-Hadley
emman@iadb.org

1 Education Division, Inter-American Development Bank (IDB), Washington, DC, USA

2 Innovations for Poverty Action (IPA), Madrid, Spain

3 Independent Researcher, Imola, BO, Italy

Published online: 18 July 2026

Introduction

In the past decade Computational Thinking (CT) has drawn interest from researchers, practitioners, and policy makers in education. Although the term is widely used, its definition is still debated. Many studies reference Wing's (2011) definition, which describes CT as: "the thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively carried out by an information-processing agent". This definition highlights two key elements: CT is a way of thinking that doesn't necessarily rely on technology, and it involves problem-solving skills where solutions can be executed by a computer or other information-processing agent.

The first element highlights that CT is a set of transversal skills that go beyond using computers or programming, making it relevant to more than just computer scientists (Voogt et al., 2015; Wing, 2006). Indeed, evidence suggests that these skills can also be developed through unplugged activities (Zhang et al., 2024). CT is seen as an essential analytical skill—like reading, writing, and math—that helps with communication, social interaction (Wing, 2008), and everyday tasks (Tang et al., 2022). This is why many governments have made CT a key 21st-century skill in school curricula (Bati, 2022; Bocconi et al., 2016).

For the second element, different authors outline various concepts and skills linked to CT. Based on a review of widely cited studies, Bocconi et al., (2016, 2022) identify the following core CT skills:

- *Abstraction:* the process of making an artefact more understandable through reducing unnecessary details.
- *Algorithmic thinking:* a way of getting to a solution through a clear definition of the steps.
- *Automation:* a labor-saving process in which a computer is instructed to execute a set of repetitive tasks quickly and efficiently compared to the processing power of a human.
- *Decomposition:* a way of thinking about artefacts in terms of their component parts. The parts can then be understood, solved, developed, and evaluated separately.
- *Debugging:* the systematic application of analysis and evaluation to find and correct bugs and errors.
- *Generalization:* associated with identifying patterns, similarities, and connections. It is a way of quickly solving new problems based on previous solutions and building on prior experience.

The growing focus on CT in education, especially in early childhood education,[1] is driven by four factors. First, research shows that learning CT positively impacts other skills for using Information and Communication Technologies (ICT) at a young age, such as critical thinking, communication, collaboration, creativity (Bers et al., 2019; Critten et al., 2022), and self-regulation (Yang et al., 2022). Second, CT activities help children to understand the technologies they encounter daily (Sullivan & Bers, 2016). Third, some authors argue that early childhood is a critical period for developing CT, as children are transitioning from concrete to abstract thinking (Flannery & Bers, 2013). Fourth, recent investigations conclude that interventions in early childhood have a higher rate of return

[1] Although the age range for what is considered Early Childhood Education may vary, it is typically defined as the period from birth to eight years of age (United Nations Educational, Scientific and Cultural Organization, 2025).

than interventions at later stages and a long-term impact on economic, health, and social outcomes (e.g. García et al., 2020).

With general agreement on the importance of CT skills, a key question is whether structured interventions can foster these skills in early childhood education. Several studies have reviewed the available empirical evidence (Bakala et al., 2021; Bati, 2022; Su & Yang, 2023), and they generally conclude that: (i) with age-appropriate instructional design, young children can develop early CT concepts and skills; (ii) smart devices and electronic toys (like robots) can aid in developing programming and CT skills in early childhood education; (iii) little is known about the most effective practices for developing these skills in young children, given that most of the available evidence focuses on primary and secondary school students; and iv) existing studies often lack coverage of developing countries, use small sample sizes, rarely employ experimental designs, and overlook factors like gender, socioeconomic status, and adult roles (e.g., teachers) in CT learning.

To contribute to this literature, in this article, we evaluate the impact of *Irûmi* ("companion" in the Guaraní language), an educational program designed to develop CT skills in second-grade students in Paraguay's public primary schools. *Irûmi* was created by the Multidisciplinary Organization to Support Teachers and Students[2] (OMAPA), Paraguay's Ministry of Education and Sciences (MEC), and the Inter-American Development Bank (IDB). The program included a curriculum of 18 classroom activities, supported by a technological kit featuring a robot. These robots were developed and donated by SK Telecom,[3] with materials adapted by the Teacher's Association of Primary Computer Education in Korea. The program also offered teacher support through training and tutor visits. It was implemented in 104 schools across seven Paraguayan regions (*departamentos*) over approximately 11 weeks (August 15 to October 29, 2022). To evaluate *Irûmi*, we conducted a Randomized Controlled Trial (RCT) with 207 schools (104 in the treatment group and 103 in the control group), 2040 students, and 160 teachers.

In addition to evaluating the effect of the program on CT skills, which is the program's main objective, the study also analyzes its impact on other dimensions considered in the literature on technology-based interventions in education. First, evidence shows that there are gender biases in preferences for and attitudes toward STEM-related toys in general, and robots in particular. For example, the literature suggests that toys are cultural products that are often gender-stereotyped (Cherney & London, 2006; Spinner et al., 2018). Even at age 5, boys and girls exhibit different preferences for toys (Cherney & London, 2006; Weisgram et al., 2014), with toys like robots traditionally associated with males (Seo, 2022). On the other hand, research indicates that exposure to programming and robotics activities can increase girls' interest and attitudes in these fields (Gomoll et al., 2016; Master et al., 2017). Additionally, counter-stereotypical situations (e.g., girls playing with robots) can change boys' and girls' attitudes toward technology (Spinner et al., 2018), which is relevant given that such attitudes can be considered a proxy for students' likelihood of pursuing STEM-related careers (Svenningsson et al., 2022).

Second, a growing body of literature suggests that programs introducing technology in education can also have effects on teachers, who are responsible for delivering the curriculum in the classroom. For example, interaction with robots may influence teachers'

[2] OMAPA is a non-profit mathematical organization based in Asunción, Paraguay. More information at https://www.omapa.org/

[3] SK Telecom is a South Korean company, specialized in the manufacturing of mobile phones and telecommunications services. More information at https://www.sktelecom.com/index_en.html

attitudes toward technology, which in turn have a significant impact on the use of new technologies in classrooms (Atkins & Vasu, 2000). At the same time, the integration of CT-related content into the curriculum and training may contribute to the development of these skills among teachers (Kong et al., 2020), which is particularly relevant given that teachers' deeper knowledge of the content they teach can enhance student learning (Sadler et al., 2013).

The *Irûmi* program

Program design

Curriculum

The *Irûmi* curriculum (see Table 1), developed by the Teacher's Association of Primary Computer Education in Korea in collaboration with coding education specialists, is delivered through a teacher guidebook and student workbooks, and is grounded in literature on the instructional design of educational interventions that promote computational thinking in young children (e.g., Bakala et al., 2021; Bati, 2022; Zhang et al., 2024). The curriculum includes 18 activities where students learn basic CT concepts by solving problems. Students create a sequence of instructions for the robot to follow, then scan coding cards using the robot's sensors to execute the sequence on a navigation board. The curriculum includes three types of activities: *Disconnected Activities*, which introduce programming concepts without using the robot or coding cards; *Familiarization Activities*, which train students to use the technological kit's components; and *Reinforcement and New Topics*, which strengthen understanding of sequences, loops, and debugging while introducing new robot functions and CT concepts like conditions and algorithmic thinking. The following materials are available in the online supplementary materials on the journal's website: (i) the 18 learning activities included in the *Irûmi* curriculum (Online Resource 1); (ii) the Teacher Guidebook, which includes lesson plans and assessments for all curriculum activities (Online Resource 2); and (iii) the Student Workbook, which contains the activities to be completed, their solutions, and the materials required to carry them out (Online Resource 3).

Educational kit

The program activities use *Irûmi* smart robots, coding cards, robot navigation tracks, a teacher's guide, and student workbooks. This technology was developed by SK Telecom, which, through the Korea Poverty Reduction Fund (KPR[4]), donated 900 educational kits to the MEC.

Robots The robot used in the program corresponds to the commercially available Smart Robot Albert and weighs 230 g and is 13 cm tall (see Fig. 1). It features autonomous navigation, LED lights, optical, light, temperature, touch, proximity sensors, as well as speakers and Bluetooth. These features allow it to move on smooth surfaces, detect obstacles from

[4] The KPR Fund focuses on poverty reduction and social development in Latin America and the Caribbean.

Table 1 *Irûmi*'s curriculum

Type of activity	Activity number	Name of the activity	Description of the activity	Programming concept or robot function involved	CT skill supported	Most aligned CT test task(s)
Disconnected Activities	1	Move! Let's be a robot	The student acts like a robot and moves following classmates' instructions	Sequence	Decomposition/Abstraction/Algorithmic thinking	Bird house/Puzzle 1
	2	Move! Capture of dolls	The student acts like a robot and grabs a doll following classmates' instructions	Sequence	Abstraction/Algorithmic thinking	Bird house/Puzzle 1
	3	Move! Land conquest game	The student acts like a robot and conquers lands following classmates' instructions	Sequence/Repetition	Algorithmic thinking	Puzzle 2
	4	Move! Box construction game	The student acts like a robot and stacks boxes following classmates' instructions	Sequence/Repetition	Algorithmic thinking	Puzzle 2

Table 1 (continued)

Type of activity	Activity number	Name of the activity	Description of the activity	Programming concept or robot function involved	CT skill supported	Most aligned CT test task(s)
Familiarization Activities	5	Say hello to your robot friend!	Presentation of the robot and instructions for using it with the coding cards	Basic function	Abstraction	
	6	Say hello to your robot friend!	The student must find hidden friends using the coding cards	Sequence	Algorithmic thinking/ Automation	Moving ball/Trapped robot
	7	Playing to eat candies	A candy-eating game using robots, coding cards, and navigation boards	Sequence	Algorithmic thinking/ Evaluation	Honey pot
	8	Who is good at this?	A football game using robots, coding cards, and navigation boards	Sequence/Repetition	Algorithmic thinking	Moving ball/Trapped robot
	9	Draw!	A drawing game using robots, coding cards, and navigation boards	Sequence	Decomposition/Algorithmic thinking	Organize the shelf/ Wooden animals
Reinforcement and new topics	10	Ta-Te-Ti	A game where the group that occupies three consecutive spaces on a 3 × 3 board wins	Sequence/Repetition	Decomposition/Evaluation/Automation/ Debugging	Organize the shelf/ Wooden animals
	11	Yes or No! Ladder game	A game where the robot moves on the board depending on the student's answers to mathematics and general knowledge questions	Sequence/Repetition/ Condition	Automation/Debugging	Way home 1/Way home 2

Table 1 (continued)

Type of activity	Activity number	Name of the activity	Description of the activity	Programming concept or robot function involved	CT skill supported	Most aligned CT test task(s)
	12	Let's dance!	A game where students make the robot perform the repetitive movements, sounds, and effects of dance	Sequence/Repetition/Sound	Automation/Debugging	Way home 1/Way home 2
	13	Password Game	A game where students create passwords assigning a correspondence rule between numbers or letters with coding cards	Sequence/Repetition/Sound/LED	Algorithmic thinking/Generalization	Way home 2
	14	Run, Irûmi!	A game where a team aims to remove the label with the name of the opposing team's robot	Sequence/Repetition/Sound/LED	Abstraction/Generalization	Way home 2
	15	Walk through the city	The student determines the best route for the robot, considering the departure point, the waypoints, and the destination	Problem-solving algorithm	Algorithmic thinking/Evaluation	Fix the bracelet
	16	Traveling through South America	On a South America map board, the robot must find and follow the best path to reach Paraguay	Problem-solving algorithm	Algorithmic thinking/Generalization/Evaluation	Fix the bracelet
	17	Travel through space	A game where the robot moves on a solar system board based on the student's answers to questions	Problem-solving algorithm	Algorithmic thinking/Generalization	Fix the bracelet

Table 1 (continued)

Type of activity	Activity number	Name of the activity	Description of the activity	Programming concept or robot function involved	CT skill supported	Most aligned CT test task(s)
	18	Trip to the ocean	A game where the robot moves on an archipelago board based on the student's answers to questions	Problem-solving algorithm	Algorithmic thinking/Generalization/Evaluation	Fix the bracelet

Activity 5 focuses on foundational abstraction skills, including understanding symbolic representations and robot functions. These skills are prerequisites for interpreting and executing the symbolic commands required in subsequent problem-solving tasks but are not assessed as standalone outcomes in the CT test

Fig. 1 *Irûmi* robot

1 to 15 cm away, display various colors, and produce sounds. The MEC provided robots to schools, distributing them based on the number of students per class, with a maximum of three students per robot in most classrooms.

Coding cards Each robot comes with 56 coding cards in six different colors (see Fig. 2). The cards contain 22 instructions that enable the robot to perform simple actions, such as moving back and forth, turning right and left, turning its colored eyes on or off, making sounds, or recording. The cards also allow for more complex actions like iterations (repeating a process a set number of times or indefinitely) and conditional structures (e.g., waiting to detect a hand or repeating an action until a specific event occurs).

Navigation boards The *Irûmi* robot features an autonomous navigation system that enables it to move on boards, where it carries out instructions programmed using the coding cards. The navigation boards are made from a special canvas that can be written on with a

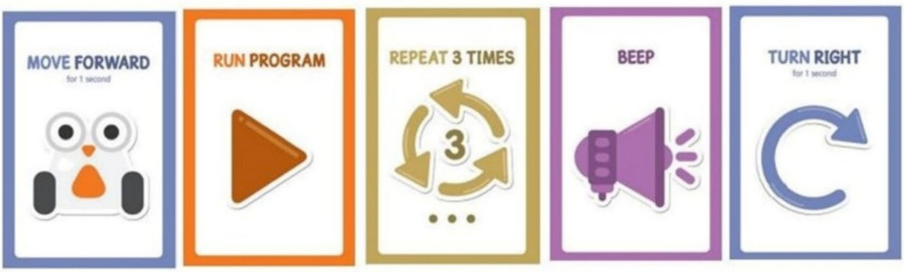

Fig. 2 Coding cards

whiteboard marker and easily erased, allowing them to be customized for each activity (see Fig. 3).

Teacher support and training

Irûmi was designed for second-grade students, who participated in programming and coding classes during the *contraturno*. In Paraguay, most public schools have a short school day of four hours, with students attending either in the morning or afternoon (Mateo-Berganza & Iribarren, 2016). The *contraturno* (countershift) is the opposite shift: afternoon for morning students and morning for afternoon students. A second-grade teacher led the activities, assisted by eight students from the third cycle of primary education (grades 7–9, corresponding to adolescents aged 12 to 14), who provided support. Neither teachers nor assistants received financial compensation for their participation in the training sessions or in the implementation of the program. The classes were structured to last up to three 40-min sessions per week, totaling 36 class hours. The activities were planned to take place over 11 weeks, from August 15 to October 29, 2022.

Each teacher received a monthly follow-up visit from a tutor trained by OMAPA. The tutor observed the teacher's performance in implementing the activities and provided feedback on their teaching practice. During these classroom observations, tutors used monitoring sheets to record information such as: (i) the activities carried out from the workbook; (ii) the classroom environment; and (iii) the main challenges identified and the teacher's needs.

Fig. 3 Navigation boards

The teacher training involved implementing the 18 activities from the program's curriculum through 4 workshops held in schools between July and September, totaling 29 h. The workshops were led by the tutors. The training adopted a Training-of-Trainers (ToT) model, in which lead trainers, previously trained, subsequently replicated the sessions with small groups of five to seven teachers. Each workshop followed an active, experiential approach focused on the practical implementation of the *Irûmi* curriculum, combining pedagogical guidance with hands-on activities using the robot and coding cards, including both unplugged activities and tangible programming. Learning was primarily collaborative, emphasizing group work, role rotation (between the roles of robot and programmer), and trial-and-error problem solving, and was complemented by structured reflection activities. Teachers' learning was assessed formatively through performance observation and qualitative rubrics.

Program implementation

This subsection presents data on the actual classroom implementation of *Irûmi*, collected through a module included in the endline survey administered to school principals and all teachers in the 104 treated schools, which covered questions on program implementation, including teacher training, provision of materials, and classroom delivery of the activities.[5]

In terms of program implementation, 99 percent of participating schools received the complete *Irûmi* kit (robot, coded cards, USB charging cables, navigation boards or canvas track, *Irûmi* student manual, *Irûmi* teacher guide, multicharger, and stationery); 69 percent took part in the four teacher training workshops, 23 percent in three, 6 percent in two, and 2 percent in one workshop. Eighty-seven percent of participants perceived the workshops as useful or very useful and 83 percent were satisfied or very satisfied with their content. Variations in the amount of training received by teachers may have affected treatment homogeneity and should be considered when reading the results discussed in the subsections below.

Regarding the program's implementation in the classroom, 96 percent of teachers reported having carried out the activities contemplated in the school plan. The number of weekly hours dedicated to the program varied, however: 12 percent of teachers reported having conducted one hour of class per week with the *Irûmi* robot, 33 percent two hours, 16 percent three hours, 23 percent four hours, and 11 percent more than five hours. Although 25 percent of schools received the materials with a delay, and 35 percent began implementing the program later than expected, the average number of activities completed was 14.7 (of a total of 18), that is, 82 percent of the expected content. Beyond average implementation levels, the percentage of curriculum activities completed varied across classrooms. Fifty-eight percent of teachers reported having completed all the activities, 21 percent more than 75 percent, 13 percent between 50 and 75 percent, and 7 percent less than 50 percent. Among the teachers who conducted classes using *Irûmi* materials, 77 percent reported that they faced no difficulties during classroom sessions. The remaining 21 percent reported difficulties in developing the classes. The most common issues included low student attendance, difficult coordinating schedules, and lack of time to develop activities. The ratio of students per robot was close to that expected (three students per robot), with

[5] The survey items administered to principals and teachers, together with details on the intervention monitoring process, are available from the authors upon request.

18 percent of teachers reporting a ratio exceeding 3.5 students per robot. Approximately 14 percent of schools faced any difficulties in getting the robots to work. In most cases, these difficulties involved basic operational issues, such as turning on the device, correctly reading the coding cards, executing commands in the appropriate sequence, or battery-related problems due to insufficient charging prior to classroom use. These challenges were typically resolved through on-site technical support and follow-up.

In terms of support, 41 percent of schools received visits from OMAPA tutors during the program's implementation in the classroom (95 percent received one visit and 5 percent two visits). The support component that presented the most difficulties was that of the student assistants (aged 12–14): despite recruitment of the expected number of assistants, 67 percent participated in the two training workshops and 59 percent participated in all classes supporting the teacher. One of the main difficulties was that not all assistants were recruited from the treated schools, as several of these schools do not offer the third cycle of primary education (grades 7–9), from which assistants were selected. This reduced the probability that assistants would be transported to the school where they were to provide support. Another difficulty was that the food service was not present in some schools, requiring student assistants to go home to eat, but some did not return.

The support constraints could explain why 43 percent of the teachers reported that the program was easy or very easy to follow, although this share increases to 64 percent when they were asked specifically about the use of the robots and to 71 percent, when asked about students' interaction with the robots. Despite these constraints, the program was perceived positively by both teachers and students: 100 percent of the teachers reported that the program was useful for teaching children programming concepts, and 99 percent indicated that it boosted students' class participation. Meanwhile, 98 percent of students reported that they liked the classes with the robots, and 99 percent indicated that they would like to continue having classes with the robots. Finally, the data indicate that implementation was more challenging for teachers in rural areas and for men.

Research questions

Given the characteristics of the program described in the previous section, the impact evaluation presented in this article is structured around the following research questions (RQ):

- RQ1: What is the effect of *Irûmi* on students' CT skills?
- RQ2: Does *Irûmi* affect students' preferences and gender attitudes toward STEM toys, and their attitudes toward technology?
- RQ3: Does *Irûmi* have an effect on teachers' classroom practices, attitudes toward technology, or CT skills?

Methods

Sample selection

The MEC selected all 300 schools participating in the Extended School Day Project (*Jornada Escolar Extendida*, or JEE), a pilot program financed by the Inter-American Development Bank ("Lo que dijo el ministro de Educación", 2024). In 131 of these schools, the extended school day model was implemented for both the first and second cycle of primary education, corresponding to grades 1–3 and 4–6, respectively. The remaining 169 schools were those where the JEE model would be implemented, as they met the following criteria: (i) having enough classrooms to implement the JEE modality simultaneously from first to sixth grade; (ii) offering a school lunch service; (iii) having digital connectivity provided by the MEC; and (iv) having at least 60 students enrolled across the first and second cycles. Based on the locations of these 300 schools, the sample was then restricted to seven *departamentos* in Paraguay: *Asunción, Central, Cordillera, Guairá, Paraguarí, Misiones,* and *Itapúa*. After this restriction, 207 schools participated in the study.

The study used a Cluster Randomized Controlled Trial, randomly assigning schools to treatment and control groups. This design helps prevent spillover effects, which could occur if randomization were done at the classroom level, as teachers might share materials or teach multiple classes within the same school. From the sample of 207 schools, 104 were assigned to the treatment group (i.e. they received all components of the *Irûmi* program), and 103 to the control group (i.e. they continued with their normal educational plan), using a stratified randomization strategy. Stratification helps improve the performance of randomizations in small samples (Bruhn & McKenzie, 2009). The strata used for randomization were department and area (rural/urban).

Instruments

To assess the potential impact of the program on second-grade students and teachers, various instruments from the literature were combined and adapted.[6]

Student measures

CT Skills To measure students' CT skills, we combined two instruments. First, we used a set of tasks from the *Bebras International Challenge on Informatics and Computational Thinking*, which is organized by Bebras, an international initiative founded in 2004 to promote informatics and CT among school students of all ages. The Challenge involves solving problems using skills such as logic, algorithmic reasoning, and problem-solving, and has been used in prior research to assess CT skills (e.g., Chiazzese et al., 2019; Lockwood & Mooney, 2018). Second, we used the *Computational Thinking Test (CTT)* described by Román-González et al. (2017), which measures CT levels based on key computing concepts used to solve and formulate problems. From these two instruments, we selected a battery of 11 tasks. The specific tasks and materials included in the CT students' test are provided in Online Resource 4. We then calculated a global score for each student (*CT Score*), representing the percentage of correct answers across all tasks. To disaggregate the results by CT skill, we grouped the 11 tasks according to the specific skills they evaluate. Some Bebras tasks assess multiple CT skills. Table 2 shows the connection between the tasks in the students' CT test and their associated skills. Finally, we measured the CT score for each skill as the percentage of correct answers for the tasks that evaluate it.

Preferences and gender attitudes towards STEM toys To measure students' preferences and gender attitudes towards STEM toys, we used two questions from the evaluation instrument of the *Pequeñas Aventureras* (Little Adventurers) program in Colombia, a gender-

[6] The full questionnaires are available upon request from the authors.

Table 2 Computational Thinking (CT) skills and related tasks in the students' CT test

Task	CT skills						
	Decomposition	Abstraction	Algorithmic Thinking	Automation	Debugging	Generalization	Evaluation
Bird house	x	x	x				
Puzzle 1 (4 pieces)		x					x
Puzzle 2 (5 pieces)		x					x
Honey pot			x				x
Moving ball			x				
Trapped robot			x				
Way home 1				x			
Organize the shelf	x		x				x
Way home 2					x		
Wooden animals		x	x			x	
Fix the bracelet							x

inclusive multimedia program focused on early math and science education. To measure preferences, we asked: “Do you like [name of the toy]? (yes or no)”. The toys included were: (i) computer, (ii) robot, (iii) telescope, (iv) Lego, (v) stethoscope, and (vi) magnifying glass. From this, we created a *STEM Toy Preference Index*, which is the average of the six dummy variables (the percentage of toys preferred). To assess gender attitudes, we asked: “Who plays with [name of the toy], girls, boys, or both?”. We then created six dummy variables (one for each toy), scoring 1 if the student answered “both” and 0 if they answered “girls” or “boys”. Finally, we created a *STEM Toy Gender Flexibility Index*, which is the average of the six dummy variables.

Attitudes towards technology To measure positive attitudes toward technology, we calculated the *Attitudes Towards Technology Index*, which is the average of the student’s level of agreement (1: agree; 0: disagree or neutral) with several statements adapted from the *Computer Attitude Questionnaire (CAQ)* and the *Pupils Attitude Towards Technology Survey (PATT-USA)*. CAQ is an instrument containing 65-item Likert scale, and it is designed to measure attitudes (feelings towards a person, or a thing) and predominant attitudes (dispositions) to computers.[7] PATT-USA is an instrument containing four sections. The first asks the student to briefly describe what they think technology is, the second includes eleven demographic questions, the third has 57 items to measure attitudes toward technology, and the fourth contains 31 items related to technological concepts (Ardies et al., 2013). Higher index values indicate more positive attitudes towards technologies.

Teacher measures

Use of new technologies in the classroom We asked teachers about their use of the following technologies in their teaching: (i) computers, (ii) notebooks, (iii) tablets, (iv) radio, (v) computer tools (Word, Excel, PowerPoint, Google Docs), (vi) file management platforms (Dropbox, Google Drive, OneDrive, etc.), (vii) digital resources such as educational software, and (viii) robots.

Attitudes towards new technologies in education To measure this dimension, we adapted the following instruments: *Teacher Attitudes Toward Technology* and *Teachers’ Attitudes Toward Computers* (adapted to technology). The first is part of a study that explored how teachers’ and students’ technology self-efficacy and attitudes affect technology use in the classroom (Mikusa, 2015), while the second is a questionnaire developed during 1995–97 for a study on the effects of technology integration education on teachers’ attitudes[8] (Christensen & Knezek, 1996). From these, we created the *Index of Attitudes Towards New Technologies in Education* which is calculated as the average of the teacher’s level of agreement (on a scale from 1, “Strongly disagree”, to 5, “Strongly agree.”) with several statements related to the role of new technologies, such as computers, notebooks, and tablets in their teaching work. Higher index values indicate more positive attitudes towards new technologies.

[7] More information at https://stelar.edc.org/instruments/computer-attitude-questionnaire-caq-0

[8] More information at https://iittl.unt.edu/content/teachers-attitude-toward-computers-tac

CT skills To measure teachers' CT skills, we developed a teacher test based on four tasks: *Way Home 1, Trapped Robot, Way Home 2, and Magic Box*. Detailed descriptions of the tasks and materials included in the CT teachers' test are provided in Online Resource 4.

Data collection

The data collection instruments included questionnaires for second-grade students (administered in person using printed teaching materials and a tablet), and teachers and principals (digital and self-administered) from all participating schools (207). The data collection took place in two stages: a baseline, where the instruments were applied before implementing the program (June 27–August 12), and an endline (October 31–December 7), during which data was collected after *Irûmi* had been implemented. All instruments were adapted to the Paraguayan context, piloted with samples from a population similar to that of the study, and validated by the MEC. As part of the data collection preparation, all data collectors and field supervisors received training, and a visit coordination protocol was developed in collaboration with school principals to ensure student attendance on the day of the visit. To guarantee data quality, an audit was conducted on a sample of cases to detect programming errors, potential data falsification, and mistakes made by data collectors during survey administration. In addition, field supervisors accompanied data collectors during school visits to identify any issues in the implementation of protocols and to provide real-time feedback for immediate correction.[9]

The final analytical sample, used to estimate the program's impact, included 1,871 students (925 in the treatment group and 946 in the control group) and 160 teachers (77 in the treatment group and 83 in the control group).

Balance validation

RCTs identify the causal impact of an intervention by creating comparable groups (Duflo et al., 2007). Using baseline data, we tested whether the treatment and control groups were statistically identical. Table 3 shows a set of school, second-grade student, and teacher variables, along with the p-values from a two-tailed test on the hypothesis that the averages of the treatment and control groups are equal. For most variables, no statistically significant differences were found. However, two variables showed statistically significant differences: the control group had a slightly higher proportion of female teachers (86.7% vs. 75.6%) and a higher probability of having received training in science (67.4% vs. 53.4%) compared to the treatment group.

Statistical models

To evaluate the effects of *Irûmi* we estimate the following equation using Ordinary Least Squares (OLS):

$$Y_{ij} = \mu_s + \beta_1 Y_{ij,t-1} + \beta_2 T_j + \beta_3 X_{ij} + \epsilon_{ij} \tag{1}$$

9 Full details of the data collection process are available upon request from the authors.

Table 3 Sample characteristics and balance validation

Variable	Control	Treatment	All	p-value	N
School characteristics					
Infrastructure status: poor	0.087	0.106	0.097	0.619	207
Infrastructure status: regular	0.223	0.24	0.232	0.722	207
Infrastructure status: good	0.689	0.654	0.671	0.518	207
Enough classrooms	0.951	0.952	0.952	0.993	207
Number of classrooms	6.816	6.683	6.749	0.617	207
School global index	7.748	7.448	7.597	0.294	207
Second-grade students					
Age (years)	7.342	7.358	7.35	0.450	2,040
Gender: Girl	0.516	0.508	0.512	0.625	2,040
CT score	0.000	-0.011	-0.006	0.822	2,040
STEM Toy Preference Index	0.000	-0.033	-0.017	0.418	2,040
STEM Toy Gender Flexibility Index	0.000	0.035	0.018	0.497	1,964
Attitudes Towards Technology Index	0.000	-0.009	-0.004	0.806	2,040
Teacher					
Gender: Woman	0.867	0.756	0.811	0.039**	180
Age (years)	43.412	42.716	43.064	0.422	180
Education: uncomplete primary	0.000	0.011	0.006	0.314	180
Education: complete primary	0.122	0.144	0.133	0.713	180
Education: complete secondary	0.022	0.011	0.017	0.662	180
Education: diploma, technician or teacher	0.444	0.467	0.456	0.777	180
Education: uncomplete tertiary	0.067	0.078	0.072	0.729	180
Education: complete tertiary	0.289	0.267	0.278	0.723	180
Education: master's or doctorate	0.056	0.022	0.039	0.238	180
Training: have received training for science teaching	0.674	0.534	0.605	0.060*	177
Training: have received training for programming teaching	0.411	0.411	0.411	0.965	180
Training: have received training for mathematics teaching	0.722	0.700	0.711	0.713	180
Employment: full-time	0.956	0.989	0.972	0.194	180
Employment: part-time	0.044	0.011	0.028	0.194	180
Language of instruction: only Spanish	0.056	0.044	0.050	0.733	180
Language of instruction: mainly Guaraní	0.011	0.011	0.011	0.933	180
Language of instruction: mainly Spanish	0.211	0.278	0.244	0.287	180
Language of instruction: both equally	0.656	0.600	0.628	0.426	180
Language of instruction: Jopará	0.067	0.067	0.067	0.980	180
Index of Attitudes Towards New Technologies in Education	0.000	-0.183	-0.092	0.240	180
CT score	0.000	0.181	0.091	0.232	180

Columns 1–3 present the averages for all observations in the sample (schools, students and teachers). Column (4) presents the p-value of the two-tailed test on the hypothesis that the averages of the treatment and control groups are equal. Column (5) shows the number of observations. Standard errors have been clustered at the school level

***, **, and * indicate the level of statistical significance at 1%, 5%, and 10% respectively

Table 4 Estimated impact of *Irûmi* on students' CT skills (in standard deviations SD)

Variable	Full sample CT score Coefficient/(SE)	Restricted sample CT score Coefficient/(SE)
Treatment (T)	0.089*	0.176*
	(0.052)	(0.080)
CT score (baseline)	0.328***	0.335***
	(0.025)	(0.032)
Observations (N)	1,871	1,193

***, ** and * indicate the level of statistical significance at 1%, 5%, and 10% respectively. SE: clustered standard error. The restricted sample corresponds to schools that implemented over 75% of the curriculum activities and where the teacher attended all four training workshops

where Y_{ij} is the outcome of interest for student or teacher i at school j, $Y_{ij,-1}$ is the same outcome measured at the baseline, T_j is a dummy variable that takes the value 1 if the school is part of the treatment group and 0 when it is part of the control group, X_{ij} is a vector of control variables measured at the baseline (students' age and gender,[10] and schools' number of classrooms and infrastructure status) and ϵ_{ij} is the error term. Including $Y_{ij,-1}$ and X_{ij} in the model helps improve statistical power (Imbens & Wooldridge, 2009). β_2 captures the causal effect of the treatment on the outcome of interest. For all specifications, we include fixed effects by strata (μ_s) to account for the randomization strategy, which enhances efficiency (Bruhn & McKenzie, 2009). Additionally, we clustered the standard errors at the school level, allowing for correlation between the errors of students within the same school. Clustering standard errors prevents the over-rejection of the null hypothesis of no significant impact (Bertrand et al., 2004).

Following the impact evaluation literature β_2 is the Intention to Treat Effect (ITT). The ITT measures the effect of the program on those assigned to the treatment, regardless of their take-up. In many cases, researchers and policy makers are interested in the impact of offering the program, as this will resemble what will be likely to happen if the program is fully implemented. For this reason, the ITT is also referred to as the "policy impact" of the program (Gertler et al., 2016).

To study the heterogeneous effects of *Irûmi*, conditional on the characteristics of the sample, we estimate the following equation:

$$Y_{ij} = \mu_s + \beta_1 Y_{ij,t-1} + \beta_2 T_j + \beta_3 X_{ij} + \beta_4 H + \beta_5 T_{ij} H + \epsilon_{ij} \tag{2}$$

where H are dichotomous variables that define characteristics of students, schools, or teachers, and β_5 captures the additional effect for the included category compared to the effect on the omitted category (β_2). So, in this model with interactive variables, $\beta_2 + \beta_5$ is the total ITT effect on the included group ($H = 1$) and β_2 in the omitted one ($H = 0$).

All the outcome variables presented in the Results section (see the Instruments subsection for definitions) were standardized using this formula:

[10] In this article, the student gender variable is treated as equivalent to biological sex, as it was recorded by data collectors based on observable characteristics, with response options limited to male or female.

$$Z_{ij} = \left(Y_{ij} - \mu_c\right)/\sigma_c \tag{3}$$

where, Y_{ij} represents the value of the outcome variable for student i in school j, and μ_c and σ_c represent the mean and standard deviation of the control group. With this transformation, the standardized regression coefficient β_2 measures the effect size of the program expressed in how many standard deviations (SD) the treated group differs from the control group (Nieminen, 2022). The only exception are the results in Table 8, where the effect of the program is expressed as percentage points (differences in probability). This is because the use of technology in the classroom was measured using dichotomous variables coded as 1 or 0 for each tool.

Results

RQ1: What is the effect of Irûmi on students' CT skills?

Table 4 shows the impact of *Irûmi* on students' CT skills, estimated using Eq. (1) on the student sample. The program improved students' *CT score* by 0.089 standard deviations (SD). To account for heterogeneity in the implementation of the intervention discussed in the previous subsection, Table 4 also reports estimates of *Irûmi*'s effect based on a restricted sample of schools—specifically, those that implemented more than 75% of the curriculum activities and where the teacher attended all four training workshops. The estimated effect of the intervention on students' *CT score* remains positive and increases in magnitude to 0.176 SD. While this estimate measures the program's impact under conditions of high implementation fidelity, the intention-to-treat (ITT) estimate based on the full sample captures the program's effect as implemented in practice, including variation in implementation and training attendance. For this reason, the ITT is widely adopted in the impact evaluation literature as a measure of program effects, as it reflects the impact of an intervention under real-world implementation conditions (e.g., Gertler et al., 2016). As such, the full-sample ITT can be interpreted as a lower-bound estimate of *Irûmi*'s overall impact.

Table 5 breaks down the effect by each CT skill, showing gains in abstraction (0.136 SD), algorithmic thinking (0.086 SD), and evaluation skills (0.143 SD). There was a negative effect on debugging skills (-0.081 SD).

Table 5 Estimated impact of *Irûmi* on different students' CT skills (in standard deviations SD)

Variable	CT skills						
	Decomposition	Abstraction	Algorithmic Thinking	Automation	Debugging	Generalization	Evaluation
Treatment (T)	0.074	0.136***	0.086*	-0.060	-0.081*	-0.009	0.143***
	(0.047)	(0.050)	(0.049)	(0.048)	(0.048)	(0.046)	(0.049)
Observations (N)	1,871	1,871	1,871	1,871	1,840	1,871	1,871

***, ** and * indicate the level of statistical significance at 1%, 5%, and 10% respectively. Clustered standard error in parentheses

Table 6 Estimated impact of *Irûmi* on students' preferences and gender attitudes towards STEM toys by gender (in standard deviations SD)

Variable	STEM Toy Preference Index Coefficient/(SE)	STEM Toy Gender Flexibility Index Coefficient/(SE)
Treatment (T)	0.022	0.009
	(0.063)	(0.071)
Gender (girl = 1)	-0.080	0.014
	(0.056)	(0.065)
Treatment (T) x Gender (girl = 1)	0.130*	0.187*
	(0.077)	(0.096)
Observations (N)	1,871	1,871

This table reports the heterogeneous effect of the program (Eq. (2))

***, **, and * indicate the level of statistical significance at 1%, 5%, and 10% respectively

SE Clustered standard error

Table 7 Estimated impact of *Irûmi* on students' attitudes towards technology (in standard deviations SD)

Variable	Attitudes Towards Technology Index Coefficient/(SE)
Treatment (T)	0.228***
	(0.047)
Attitudes Towards Technology Index (baseline)	0.222***
	(0.025)
Observations (N)	1,871

***, **, and * indicate the level of statistical significance at 1%, 5%, and 10% respectively

SE Clustered standard error

RQ2: Does Irûmi affect students' preferences and gender attitudes toward STEM toys, and their attitudes toward technology?

Table 6 presents the *Irûmi*'s heterogeneous effects on students' preferences (*STEM Toy Preference Index*) and gender attitudes toward STEM toys (*STEM Toy Gender Flexibility Index*), by gender, estimated using Eq. (2). For STEM toy preferences, the program produced a positive but not statistically significant effect on boys (0.022 SD). For girls, the impact was both positive and statistically significant, with *Irûmi* increasing the *STEM Toy Preference Index* by 0.152 SD (0.022 SD + 0.130 SD). Similarly, in terms of gender flexibility, the program effect for boys was not statistically significant (0.009 SD). Among girls, *Irûmi* increased the *STEM Toy Gender Flexibility Index* by 0.196 SD (0.009 SD + 0.187 SD).

Table 7 provides estimates of the *Irûmi*'s impact on students' attitudes toward technology. Students in the treatment group had a statistically significant increase in positive attitudes, with an increase of 0.228 SD in the *Attitudes Towards Technology Index* relative to the control group.

Table 8 Estimated impact of *Irûmi* on the probability that teachers use new technologies in the classroom (in percentage points)

Variable	Technological tools							
	Computers	Notebooks	Tablet	Radio	Computer tools	Digital file management platforms	Educational software	Robot
Treatment (T)	0.076	0.179**	0.048	-0.028	-0.015	0.007	0.082	0.816***
	(0.081)	(0.074)	(0.044)	(0.083)	(0.074)	(0.083)	(0.082)	(0.046)
Observations (N)	160	160	160	160	160	160	160	160

Computer tools: Word, Excel, PowerPoint, Google Docs; Digital file management platforms: Dropbox, Google Drive, OneDrive

***, **, and * indicate the level of statistical significance at 1%, 5%, and 10% respectively. Clustered standard error in parentheses

Table 9 Estimated impact of *Irûmi* on teachers' attitudes towards new technologies (in standard deviations SD)

Variable	Index of Attitudes Towards New Technologies in Education Coefficient/(SE)
Treatment (T)	0.293*
	(0.156)
Index of Attitudes Towards New Technologies in Education (baseline)	0.416***
	(0.074)
Observations (N)	160

***, **, and * indicate the level of statistical significance at 1%, 5%, and 10% respectively

SE Clustered standard error

Table 10 Estimated impact of *Irûmi* on different tasks of the teachers' CT test (in standard deviations SD)

Task	Coefficient/(SE)	Observations (N)
Way home 1	0.155*	160
	(0.081)	
Trapped robot	0.206**	160
	(0.077)	
Way home 2	0.028	160
	(0.074)	
Magic box	-0.104	160
	(0.078)	

***, **, and * indicate the level of statistical significance at 1%, 5%, and 10% respectively

SE Clustered standard error

RQ3: Does Irûmi have an effect on teachers' classroom practices, attitudes toward technology, or CT skills?

Table 8 shows the impact of *Irûmi* on teachers' probability of using different technological tools in the classroom, based on Eq. (1). Teachers in the treatment group were 82 percentage points (pp) more likely to use robots compared to teachers in the control group. Additionally, there was an 18 pp increase in the probability of teachers using notebooks (laptops), despite notebooks not being part of the program's strategy. This increase coincides to recent initiatives by the MEC aimed at providing notebooks to schools ("MEC entregó más de 650 notebooks", 2021).

Table 9 presents the estimated effect of the program on teachers' attitudes towards new technologies, using Eq. (1). *Irûmi* had an estimated effect of 0.293 SD on the *Index of Attitudes Towards New Technologies in Education*. This result reflects a more positive attitude toward educational technology.

Table 10 shows the estimated impact of *Irûmi* on each of the 4 tasks included in the teachers' CT test, based on Eq. (1). The effect of the program is observed in the *Way Home*

1 (0.155 SD) and *Trapped robot* (0.206 SD) tasks, which measure algorithmic thinking and automation skills respectively.

Discussion

In this article we present the impact evaluation of *Irûmi*, an educational program aimed to develop CT skills in second-grade students in Paraguay. Overall, despite some implementation challenges typical of school-based technology interventions, the intervention was implemented with a high degree of fidelity. Most schools received the full set of materials and completed the majority of the planned activities, and both teachers and students reported positive perceptions of the program.

Our results suggest that the program positively impacted students' CT skills (RQ1), with an effect size of 0.089 SD, particularly developing skills like abstraction, algorithmic thinking, and evaluation (according to Bebras' classification). These findings align with existing empirical literature, which suggests that well-designed, age-appropriate instruction can support young children in developing foundational CT skills and that smart devices and electronic toys, such as robots, can contribute to the development of these skills in early childhood education (Bakala et al., 2021; Bati, 2022; Su & Yang, 2023). This study extends the literature by evaluating a developing country context, using a rigorous experimental design with control and treatment groups, pre- and post-intervention assessments, a large student sample, and by examining the program's effects on both students and teachers. Our findings are also relevant to the discussion on how to incorporate CT into school curricula in Latin America. Although there are various initiatives in the region aimed at introducing CT concepts and skills into the classroom—many of which are based on the use of educational robots—the literature highlights a lack of rigorous quantitative research on the impact of interventions designed to develop CT skills in students and teachers (Curasma & Curasma, 2020; Quiroz-Vallejo et al., 2021). Such evidence is crucial for informing the integration of CT into the school curriculum. This study contributes to narrowing this knowledge gap.

Despite the positive effects of *Irûmi* on the students' CT skills, the program had no impact on some specific abilities, such as decomposition, abstraction, and automation, and a small negative effect on debugging skills. Two hypotheses could explain these findings. First, there are challenges in designing and implementing instruments to measure some CT skills in early childhood, partly due to a lack of consensus on CT frameworks and definitions (Bocconi et al., 2016; Su & Yang, 2023)**.** Second, it is possible that young children have difficulties to develop more advanced CT concepts such as iterations, conditionals, or debugging (Bakala et al., 2021; Su & Yang, 2023). Some authors suggest that this is due to differences in mental development between children of different ages. In particular, very young children had less development of working memory, which is defined as the ability to hold and manipulate information over short periods. On the other hand, there is evidence that this ability is not easily developed with interventions at this age (Bati, 2022).

The *Irûmi*'s impact extended beyond students' CT skills, influencing additional dimensions not explicitly targeted by its original design (RQ2). It enhanced students' preferences for STEM toys, including the robot itself, and promoted gender flexibility—challenging traditional notions about who can engage with these toys—particularly among girls. This suggests that educational programs designed to develop CT skills through technology, like robotics, may also help reduce gender biases in early education. This finding is consistent with a recent study showing that a developmentally appropriate robotics curriculum can increase girls' interest in engineering during early elementary education (Sullivan & Bers, 2019). Finally, consistent with the existing literature, we found that the program also contributed to increasing children's positive attitudes toward technology, which may influence their future choices and motivation to pursue STEM-related careers.

In addition to its impact on students, the intervention also had effects on the teachers who implemented the activities (RQ3). First, it increased the probability that teachers would use educational technology in the classroom, even for devices not initially included in the program, such as notebooks, suggesting a spillover effect. This can be partly explained by the program's positive impact on teachers' attitudes toward technology. Second, the program contributed to the development of teachers' CT skills, likely due to the innovative nature of the curriculum and its methodology. This is important from an educational policy perspective, as there is a broad consensus that teachers should be provided with training opportunities focused on CT pedagogy and hands-on learning (Bocconi et al., 2016). According to the teachers' opinions, the training and school planning components were key factors for the positive impact of *Irûmi* on teachers and students.

Despite the relevance and novelty of the results described above, this study has several limitations, related to both the methods and the findings, that should be considered when interpreting the results. These include heterogeneity in the implementation of the intervention across schools (e.g., differences in the number of curriculum activities implemented and in the amount of teacher training received), the fact that the reported outcomes reflect a short-term evaluation of *Irûmi*'s impact (approximately one month after the end of implementation), and the specific context of public schools in Paraguay, which may limit the generalizability of the findings. In addition, the intervention was relatively brief (approximately 36 h distributed over a three-month period). The study also lacks information on the mechanisms that may help explain the program's effects on girls' preferences and attitudes toward STEM-related toys, as well as outcome measures related to students' performance in core curriculum subjects where the intervention could plausibly have had spillover effects, such as mathematics, science, or language.

Future research should address several questions that could complement our results. First, it is important to assess the medium- and long-term impact of *Irûmi* on CT skills and other non-cognitive outcomes to analyze if its positive effects last beyond the program. Second, the program's impact should be evaluated in different contexts. Third, it would be useful to study how CT skills relate to student learning in other subjects like mathematics, science, and language. Finally, further research is needed to understand the mechanisms behind the program's effect on girls' preferences and attitudes toward STEM toys, which could help design interventions to reduce gender biases in teaching CT.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11423-026-10675-2.

Acknowledgements The authors are thankful to the Inter-American Development Bank for funding the research presented in this paper through its Korea Poverty Reduction Fund (KPR). We would also like to thank the Ministry of Education and Science of Paraguay (MEC) and the Multidisciplinary Organization for Support to Teachers and Students (OMAPA) for their support in implementing the program. The opinions expressed are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.

Funding The research leading to these results received funding from the Government of South Korea through a grant from the Korea Poverty Reduction Fund of the Inter-American Development Bank (IDB).

Data availability The data that support this study's findings are available from Innovations for Poverty Action (IPA), but restrictions apply to their availability because they include personal information regarding second-grade children whose parents authorized them to participate in the study and so are not publicly available. The data are, however, available from the authors upon reasonable request and with the permission of the Inter-American Development Bank (IDB) and the Ministry of Education and Science of Paraguay (MEC).

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose that are relevant to the content of this article.

Role of the authors The authors of this manuscript were involved in the design of the intervention and in the evaluation of its impact, as well as in the preparation of the manuscript. The implementation of the curriculum in treated schools was carried out by OMAPA in collaboration with the MEC, while the administration of the evaluation instruments was conducted by externally hired data collectors trained by Innovations for Poverty Action (IPA).

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Ardies, J., De Maeyer, S., & Gijbels, D. (2013). Reconstructing the pupils attitude towards technology- survey. *Design and Technology Education: An International Journal, 18*(1), 8–19. https://openjournals.ljmu.ac.uk/DATE/article/view/1660/843

Atkins, N. E., & Vasu, E. S. (2000). Measuring knowledge of technology usage and stages of concern about computing: A study of middle school teachers. *Journal of Technology and Teacher Education*, *8*(4), 279–302. https://www.learntechlib.org/primary/p/8038/

Bakala, E., Gerosa, A., Hourcade, J. P., & Tejera, G. (2021). Preschool children, robots, and computational thinking: A systematic review. *International Journal of Child-Computer Interaction*. https://doi.org/10.1016/j.ijcci.2021.100337

Bati, K. (2022). A systematic literature review regarding computational thinking and programming in early childhood education. *Education and Information Technologies, 27*, 2059–2082. https://doi.org/10.1007/s10639-021-10700-2

Bers, M. U., Gónzalez-Gónzalez, C., & Armas-Torres, M. B. (2019). Coding as a playground: Promoting positive learning experiences in childhood classrooms. *Computers & Education, 138*, 130–145. https://doi.org/10.1016/j.compedu.2019.04.013

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics, 119*(1), 249–275.

Bocconi, S., Chioccariello, A., Dettori, G., Ferrari, A., & Engelhardt, K. (2016). Developing computational thinking in compulsory education – Implications for policy and practice. *European Commission, Joint Research Centre*. https://doi.org/10.2791/792158

Bocconi, S., Chioccariello, A., Kampylis, P., Dagienė, V., Wastiau, P., Engelhardt, K., Earp, J., Horvath, M. A., Jasutė, E., Malagoli, C., Masiulionytė-Dagienė, V., & Stupurienė, G. (2022). *Reviewing computational thinking in compulsory education: State of play and practices from computing education*. Publications Office of the European Union. https://doi.org/10.2760/126955

Bruhn, M., & McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics, 1*(4), 200–232. https://doi.org/10.1257/app.1.4.200

Cherney, I., & London, K. (2006). Gender-linked differences in the toys, television shows, computer games, and outdoor activities of 5- to 13-year-old children. *Sex Roles, 54*, 717–726. https://doi.org/10.1007/s11199-006-9037-8

Chiazzese, G., Arrigo, M., Chifari, A., Lonati, V., & Tosto, C. (2019). Educational robotics in primary school: Measuring the development of computational thinking skills with the Bebras tasks. *Informatics*. https://doi.org/10.3390/informatics6040043

Christensen, R., & Knezek, G. (1996). *Constructing the Teachers' Attitudes toward Computers (TAC) questionnaire* [Paper presentation]. Annual Meeting of the Southwest Educational Research Association, New Orleans. https://eric.ed.gov/?id=ED398244

Critten, V., Hagon, H., & Messer, D. (2022). Can pre-school children learn programming and coding through guided play activities? A case study in computational thinking. *Early Childhood Education Journal, 50*, 969–981. https://doi.org/10.1007/s10643-021-01236-8

Curasma, R. P., & Curasma, H. P. (2020). Computational thinking in school education in South America: Systematic review of the literature. In *2020 IEEE XXVII International Conference on Electronics, Electrical Engineering and Computing (INTERCON)* (pp. 1–4). IEEE. https://doi.org/10.1109/INTERCON50315.2020.9220200

Duflo, E., Glennerster, R., & Kremer, M. (2007). Using Randomization in Development Economics Research: A Toolkit. In T. P. Schultz, & J.A. Strauss (Eds.), *Handbook of Development Economics*, *4*, 3895-3962. https://doi.org/10.1016/S1573-4471(07)04061-2

Flannery, L. P., & Bers, M. U. (2013). Let's dance the "robot hokey-pokey!" *Journal of Research on Technology in Education, 46*(1), 81–101. https://doi.org/10.1080/15391523.2013.10782614

García, J. L., Heckman, J. J., Ermini Leaf, D., & Prados, M. J. (2020). Quantifying the life-cycle benefits of an influential early childhood program. *Journal of Political Economy*. https://doi.org/10.1086/705718

Gertler, P., Martínez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. (2016). *Impact evaluation in practice, second edition*. Inter-American Development Bank and World Bank.

Gomoll, A., Hmelo-Silver, C. E., Šabanović, S., & Matthew, F. (2016). Dragons, ladybugs, and softballs: Girls' STEM engagement with human-centered robotics. *Journal of Science Education and Technology, 25*, 899–914. https://doi.org/10.1007/s10956-016-9647-z

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature, 47*(1), 5–86. https://doi.org/10.1257/jel.47.1.5

Kong, S.-C., Lai, M., & Sun, D. (2020). Teacher development in computational thinking: Design and learning outcomes of programming concepts, practices and pedagogy. *Computers & Education, 148*, Article 103801. https://doi.org/10.1016/j.compedu.2020.103872

Lo que dijo el ministro de Educación sobre la Jornada Escolar Extendida. (2024). *ABC Color.* https://www.abc.com.py/nacionales/2024/03/07/lo-que-dijo-el-ministro-de-educacion-sobre-la-jornada-escolar-extendida/

Lockwood, J., & Mooney, A. (2018). Developing computational thinking test using bebras problems. In A. Piotrkowicz, R. Dent-Spargo, S. Dennerlein, I. Koren, P. Antoniou, P. Bailey, T. Treasure-Jones, I. Fronza, & C. Pahl (Eds.), *Joint proceedings of the CC-TEL 2018 and TACKLE 2018 workshops*. CEUR workshop proceedings. https://ceur-ws.org/Vol-2190/

Master, A., Cheryan, S., Moscatelli, A., & Meltzoff, A. (2017). Programming experience promotes higher STEM motivation among first-grade girls. *Journal of Experimental Child Psychology, 160*, 92–106. https://doi.org/10.1016/j.jecp.2017.03.013

Mateo-Berganza, M., & Iribarren, M.L. (2016). ¿Cuáles son las ventajas de la jornada escolar extendida? *Enfoque Educación*. https://blogs.iadb.org/educacion/es/paraguay-jornada-escolar-extendida/

MEC entregó más de 650 notebooks a instituciones educativas de Central. (2021). *Agencia IP*. https://www.ip.gov.py/ip/mec-entrego-mas-de-650-notebooks-a-instituciones-educativas-de-central/

Mikusa, M. E. (2015). *The effect of technology self-efficacy and personal engagement on students' and teachers' attitudes toward technology use in education* [Doctoral dissertation, Appalachian State University, United States]. https://libres.uncg.edu/ir/asu/listing.aspx?id=18800

Nieminen, P. (2022). Application of standardized regression coefficient in meta-analysis. *BioMedInformatics, 2*(3), 434–458. https://doi.org/10.3390/biomedinformatics2030028

Quiroz-Vallejo, D. A., Carmona-Mesa, J. A., Castrillón-Yepes, A., & Villa-Ochoa, J. A. (2021). Integración del pensamiento computacional en la educación primaria y secundaria en Latinoamérica: Una revisión sistemática de literatura. *Revista De Educación a Distancia (RED)*. https://doi.org/10.6018/red.485321

Román-González, M., Pérez-González, J.-C., & Jiménez-Fernández, C. (2017). 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

Sadler, P., Sonnert, G., Coyle, H., Cook-Smith, N., & Miller-Friedmann, J. (2013). The influence of teachers' knowledge on student learning in middle school physical science classrooms. *American Educational Research Journal (AERJ), 50*, 1020–1049. https://doi.org/10.3102/0002831213477680

Seo, S. (2022). When female (male) robot is talking to me: effect of service robots' gender and anthropomorphism on customer satisfaction. *International Journal of Hospitality Management*. https://doi.org/10.1016/j.ijhm.2022.103166

Spinner, L., Cameron, L., & Calogero, R. (2018). Peer toy play as a gateway to children's gender flexibility: The effect of (counter)stereotypic portrayals of peers in children's magazines. *Sex Roles, 79*(5), 314–328. https://doi.org/10.1007/s11199-017-0883-3

Su, J., & Yang, W. (2023). A systematic review of integrating computational thinking in early childhood education. *Computers and Education Open*. https://doi.org/10.1016/j.caeo.2023.100122

Sullivan, A., & Bers, M. U. (2016). Robotics in the early childhood classroom: Learning outcomes from an 8-week robotics curriculum in pre-kindergarten through second grade. *International Journal of Technology and Design Education, 26*, 3–20. https://doi.org/10.1007/s10798-015-9304-5

Sullivan, A., & Bers, M. U. (2019). Investigating the use of robotics to increase girls' interest in engineering during early elementary school. *International Journal of Technology and Design Education, 29*, 1033–1051. https://doi.org/10.1007/s10798-018-9483-y

Svenningsson, J., Höst, G., Hultén, M., & Hallström, J. (2022). Students' attitudes toward technology: Exploring the relationship among affective, cognitive and behavioral components of the attitude construct. *International Journal of Technology and Design Education, 32*, 1531–2155. https://doi.org/10.1007/s10798-021-09657-7

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

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2025). *What you need to know about early childhood care and education.* https://www.unesco.org/en/early-childhood-education/need-know?hub=70242

Voogt, J., Fisser, P., Good, J., Mishra, P., & Yadav, A. (2015). Computational thinking in compulsory education: Towards an agenda for research and practice. *Education and Information Technologies, 20*, 715–728. https://doi.org/10.1007/s10639-015-9412-6

Weisgram, E., Fulcher, M., & Dinella, L. (2014). Pink gives girls permission: Exploring the roles of explicit gender labels and gender-typed colors on preschool children's toy preferences. *Journal of Applied Developmental Psychology, 35*(5), 401–409. https://doi.org/10.1016/j.appdev.2014.06.004

Wing, J.M. (2011). Research notebook: Computational thinking—What and why? *The Link Magazine, Spring*. Carnegie Mellon University, Pittsburgh. https://www.cs.cmu.edu/link/research-notebook-computational-thinking-what-and-why

Wing, J. M. (2006). 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*, 3717–3725. https://doi.org/10.1098/rsta.2008.0118

Yang, W., Tsz Kit Ng, D., & Gao, H. (2022). Robot programming versus block play in early childhood education: Effects on computational thinking, sequencing ability, and self-regulation. *British Journal of Educational Technology, 53*(6), 1817–2184. https://doi.org/10.1111/bjet.13215

Zhang, Y., Liang, Y., Tian, X., & Yu, X. (2024). The effects of unplugged programming activities on K-9 students' computational thinking: Meta-analysis. *Educational Technology Research and Development, 72*, 1331–1356. https://doi.org/10.1007/s11423-023-10339-5

Emma Näslund-Hadley is a Principal Education Specialist at the Inter-American Development Bank (IDB) in Washington DC. Emma's research spans pre-primary through secondary education, focusing on discovering learning processes that promote children's development of socioemotional and cognitive skills, primarily in mathematics and science.

Juan Manuel Hernández-Agramonte is the Senior Director of the Embedded Labs Program at Innovations for Poverty Action (IPA). His research focuses on generating evidence in education policy and early childhood development, as well as evaluating strategies to assist governments in effectively utilizing data and evidence for decision-making.

Humberto Santos is an independent researcher and external consultant for the Inter-American Development Bank's Education Division. His work focuses on school funding, school choice and vouchers, intercultural bilingual education, and socioemotional learning. He holds a master's degree in economics from the Universidad de Chile.

Pablo Zoido is a Lead Specialist in the Education Division of the Inter-American Development Bank (IDB). He previously worked as the technical leader for the OECD's PISA for Development project and on governance indicators at the World Bank and Latin American Economic Outlooks at the OECD Development Centre.