Built to Last

Advancing Education, Equity and Economic Development through Literacy and Numeracy

The Case of Latin America and the Caribbean

EDITED BY
Emma Näslund-Hadley
Ximena Dueñas

IDB

CONTENTS

LIST OF FIGURES

CONTRIBUTORS

■ **Elena Arias Ortiz,** a national of Costa Rica, is a Lead Education Specialist in the Education Division of the IDB Group in Washington, DC. Elena joined the Bank in November 2011 through the Young Professionals Program and has been part of the Education Division since then. Before, she worked as a consultant for the World Bank, UNDP, and the European Commission. Her work focuses on higher education, digital transformation, and innovation. Elena has contributed to high-level policy dialogue and the development of education programs in Colombia, Jamaica, Peru, Suriname, Uruguay, and Venezuela. She has been published in peer-reviewed international journals and has co-authored several books and book chapters focused on economics of education. Elena holds a master's degree in Economic Analysis and a Ph.D. in Economics, both from the Université libre de Bruxelles (ULB).

■ **Dante Contreras,** a national of Chile, is Professor of Economics at the University of Chile and holds a Ph.D. in Economics from the University of California. He has held several senior academic and policy positions, including Executive Director at the World Bank, Vice President of Chile's National Copper Corporation (CODELCO), officer at the United Nations Development Programme (UNDP) and Director of the Department of Economics at the University of Chile. He has also advised the Inter-American Development Bank, the World Bank, the OECD, and several Chilean government agencies. His research focuses on the economics of education, inequality, and social mobility, examining how institutions, public policies, and large-scale shocks influence human capital accumulation.

■ **Rafael de Hoyos,** a national of Mexico, is the Principal Economic Advisor at the Social Sector Department of the IDB Group in Washington, DC., where he leads the analytical agenda including education, health, social protection, labor, migration, gender, and poverty, with a focus on evidence-based policymaking. Previously, he served at the World Bank as Human Development Program Leader for the European Union and as Lead Economist for Education. He is also a founding partner of Xaber, an NGO that promotes the use of evidence in education policy across Latin America. Rafael has advised governments in Latin America and Europe, published extensively in peer-reviewed journals, and served as chief adviser to Mexico's Deputy Minister of Education. He holds a master's degree from the University of Sussex and a Ph.D. in Economics from the University of Cambridge.

■ **Ximena Dueñas,** a national of Colombia, is a Senior Education Specialist in the Education Division of the IDB Group. Her areas of interest are early literacy, assessment, and digital transformation. Before joining the Education Division, she was the director of the Colombian Institute for the Education Evaluation (ICFES), and previously a professor in the Department of Economics at Icesi University in Cali, Colombia. Ximena has a master's degree in International Affairs and a Ph.D. in Economics and Education from Columbia University.

■ **Silvana Godoy Mateus,** a national of Colombia, is an economist and education specialist with extensive experience in educational assessment, learning measurement, and public policy in Latin America. She has held senior technical and leadership positions at Colombia's Institute for Educational Evaluation (ICFES), where she served as Director of Evaluation and Subdirector of Analysis and Dissemination. In these positions, she led large-scale national and international assessments and contributed to the use of evidence for education policy and system improvement. Silvana holds a bachelor's degree in Economics and International Business and a master's degree in Economics from Icesi University in Cali, Colombia, as well as a Master of Arts in Liberal Studies from The New School.

■ **Andrea Paula Goldin,** a national of Argentina, is a neuroscientist and researcher at CONICET in Argentina. She has received international awards, given lectures in several countries, and authored dozens of scientific publications. Her research in educational neuroscience and behavioral sciences seeks to bring the most up-to-date scientific evidence to educators and public policymakers across Latin America. Andrea is an active science communicator, a member of international scientific associations, author of the award-winning book *Neurociencia en la escuela (Neuroscience at School)*, and of the series *Repensar la currícula del futuro (Rethinking the Curriculum of the Future),* written for UNESCO. In addition, she serves as a professor, columnist, and consultant for public and private institutions in different countries.

■ **Fernanda Marquez Ragonesi,** a national of Argentina, is an economist with a bachelor's degree from the National University of Tucumán and is pursuing a master's degree at the National University of La Plata. She previously worked with the Education Division at IDB, contributing to research on educational trajectories and foundational skills in Latin America and the Caribbean. She currently serves as a Sectoral Analysis Officer at the Development Bank of Latin America and the Caribbean (CAF). Her interests include education, gender, and development economics.

■ **Mercedes Mateo-Berganza,** a national of Spain, is the Education Division Chief at the IDB Group, where she leads an ambitious initiative to rethink education and strengthen learning ecosystems to prepare citizens with 21st-century skills. She coordinates the research, design, and execution of innovative education projects, working across areas of international development and social policy with a strong emphasis on inequality. Her contributions also include institutional reform, female labor participation, early childhood education and childcare policies, soft skills development, and social cohesion. She holds a Ph.D. in Political Science from the University of Louvain (Belgium). In 2004, she was a researcher at the Belgian Fund for Scientific Research (FNRS), continuing as an honorary researcher until 2007, and from 2002 to 2004 was a Marie-Curie Postdoctoral Fellow at the Robert Schuman Centre for Advanced Studies at the European University Institute.

■ **Emma Näslund-Hadley,** a national of Sweden and the United States, is a Principal Education Specialist in the Education Division of the IDB Group in Washington, DC, where she directs and collaborates in the design and execution of a wide range of education sector projects and reforms throughout Latin America and the Caribbean. 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. Previously she held positions with the European Parliament and the United Nations. Emma has a master's degree in international economics and finance from Linköping University and a master's degree in public affairs from Princeton University.

■ **Pablo Zoido,** a national of Spain, is a Lead Education Specialist at the IDB Group in Washington, DC. He heads initiatives on education reform, learning assessment, digital transformation, and skills development across Latin America and the Caribbean. Before joining the IDB, he worked at the OECD, where he oversaw technical work as part of the PISA for Development program. His research and work focus on improving learning outcomes through evidence-based policies, education technology, and stronger education management systems. Pablo holds a master's degree in International Relations from Johns Hopkins University and a master's degree in Business Research from Stanford University.

PREFACE

Emma Näslund-Hadley & Ximena Duenas Herrera

Every child who enters a classroom carries with them an implicit promise: schooling will give them the tools to understand the world, to participate in it, and to keep learning throughout their life. At the center of that promise are two skills so fundamental that everything else in education depends on them: the ability to read and the ability to reason with numbers.

Foundational skills refer to the literacy and numeracy competencies that form the basis of all subsequent learning, economic participation, and social inclusion. Reading and mathematics are not simple subjects on a school timetable; they are the cognitive infrastructure through which children make sense of information, develop reasoning and analytical capacity, and acquire the autonomy to keep learning throughout their lives.

These skills are cumulative by nature: each stage of development builds on what came before, which means that gaps left unaddressed in the early grades don't only persist, but compound over time. A child who cannot read fluently by the end of primary school is less likely to succeed in secondary education; an adolescent who leaves school without basic numeracy enters the labor market at a permanent disadvantage; an adult without foundational competencies faces barriers to productivity, civic participation, and lifelong learning that are increasingly difficult and costly to overcome. Foundational skills are, in this sense, not one investment among many. They are the condition on which all other investments in human development depend.

This publication is about that promise and what it will take to keep it in Latin America and the Caribbean.

The evidence assembled in these pages is both sobering and encouraging. It is sobering because decades of investment in expanding school access have not translated into the learning gains the region needs. Too many children leave the early grades unable to read a simple text or solve a basic arithmetic problem. Too many adolescents progress through secondary school without consolidating the foundations they were meant to build years earlier. And too many adults enter the labor market without the competencies needed to fully participate and prosper in it. The learning crisis documented in this monograph is not marginal. It is structural, emerges early, and if current approaches continue, largely self-perpetuating.

At the same time, the evidence is also encouraging because it clearly shows that this outcome is not inevitable. Foundational skills can be taught. Well-designed instruction produces meaningful, measurable learning gains across all ages and all contexts, including the most resource-constrained settings in the region. The problem is not that effective approaches are lacking. It is that those approaches are not yet reaching classrooms at scale, with fidelity, and sustained over time. That is the gap this publication sets out to close.

Built to Last: Advancing Education, Equity and Economic Development through Literacy and Numeracy brings together an introduction and seven chapters by researchers and practitioners in the region to trace foundational skills from their earliest roots in brain development through primary and secondary education and into adult learning. They also outline what has worked and what has not into down-to-earth guidance for policymakers, program designers, and education leaders. This work draws on systematic reviews of close to 280 rigorous impact evaluations, original neuroscience, large-scale assessment data, and qualitative fieldwork with some of the region's most experienced implementing organizations. The volume opens with an introduction that draws these threads together into a forward-looking policy agenda, translating what the chapters establish about what works, for whom, and under what conditions, into the strategic priorities and institutional commitments that the next decade of foundational skills reform in Latin America and the Caribbean will require.

The publication is organized in three parts

Part I: Why Foundational Skills Matter establishes the case for urgent action. The first chapter examines the long-term returns to foundational skills, including their links to labor market participation, earnings, social mobility, and economic growth, making the case that literacy and numeracy are not only concerns of the education sector but basic investments in human development and productivity. The second chapter documents the current state of learning across the region using evidence from three major international assessments. The picture that emerges is blunt: foundational learning is alarmingly low in the region, progress has been too slow, and the gaps are deepest for the most vulnerable children, including students from low-income households, rural areas, indigenous communities, and under-resourced public schools. Deficits emerge as early as third grade and widen as students advance, producing a funnel effect in which only around one in five students completes secondary school with a foundational competency at its minimum.

Part II: How Foundational Skills Are Built looks at the evidence on what works and how it works. It opens with a chapter grounding the instructional evidence in neuroscience, examining how the brain learns to read and compute, why reading does not emerge naturally through exposure alone, and what the science of learning tells us about the conditions under which foundational skills develop most effectively. Subsequent chapters present systematic syntheses of rigorous impact evaluations in mathematics and literacy, identifying which pedagogical approaches, instructional structures, delivery modes, and program durations produce the strongest learning gains, and for which skills. A consistent picture emerges: explicit, structured, sequenced instruction outperforms more diffuse approaches, and the dominant models in the region fall short of what the evidence recommends. The final chapter in this part extends the evidence to tutoring and remedial education: the targeted interventions that offer students who have fallen behind a genuine second chance, identifying the design features that determine whether remedial support closes learning gaps or merely marks time.

Part III: From Evidence to Scale addresses the hardest problem of all: the gap between what works in controlled conditions and what actually happens when programs are scaled across real school systems. It applies the lens of implementation to ask why programs that work as pilots so often lose potency at scale. The answer lies not in the program design itself, but in the conditions those programs encounter in practice, including fragile political will, broken distribution chains, inadequate initial teacher education, and the deeply relational work of shifting classroom practice that no training event alone can accomplish.

Taken together, these chapters make an argument that is as simple to state as it is demanding to act on. The region already has the evidence it needs. What it requires now is the political will, the institutional capacity, and the resolve to implement these foundational skills the way the evidence says they are best taught: systematically, explicitly, and without waiting for another generation of children to fall irreversibly behind.

This publication is addressed to the decision- and policy-makers who have the power to change that, including ministers of education, senior officials, program leaders, and the researchers and practitioners who support them. It is offered in the conviction that the gap between what we know and what we do is not fixed. It can be closed. And closing it is among the highest-return investments any society in the region can make.

A special thank you to Haydee Alonzo, Damián Gonzo, Ariana Grossi, Beatriz Guimaraes, Katelyn Hepworth, Greta Olivares, José David Pietro, Agustina Thailinger, and Juan Sebastián Vallejo, whose work on the systematic reviews and steadfast support made this publication possible.

INTRODUCTION

BUILT TO LAST: WHY READING AND MATH FOUNDATIONS STILL DECIDE WHO THRIVES

Mercedes Mateo-Berganza

Somewhere in Latin America and the Caribbean (LAC), at this very moment, a seven-year-old is looking at a page and not understanding what is on it. She is not failing to try. She is not lacking in intelligence. She simply was never given the specific and systematic instruction that turns those black marks on the page into meaning. Multiply her by the tens of millions of children across the region who share her situation, and you have more than a statistic: you have the central, most neglected and yet solvable problem in Latin American and Caribbean development.

This book began with a wager that may sound almost too simple for a region that debates trade policy, monetary frameworks, and industrial strategy with great sophistication: that the surest way to raise productivity, reduce inequality, and prepare the region for whatever the twenty-first century has in store is to make sure every child can read a sentence and solve a numerical problem. Not eventually. Not approximately. By third grade.

This introduction lays out why that wager is not simple at all, and why getting it right is the most consequential task the region can set for itself today.

Fix the Foundations or the 21st Century will Fix Itself Without Us

For more than two decades, LAC´s income per capita has hovered at roughly a quarter of that of the United States, while East Asia climbed from half to nearly three-quarters. The region is not converging. It is treading water in a current that moves.

The explanation on offer is not new, but it is more precise than it used to be. It is not geography, nor resource endowment, nor even, primarily, years of schooling. It is learning.

Eric Hanushek and Ludger Woessmann’s finding that cognitive skills, not diplomas, explain three-quarters of cross-country variations in growth reframes an entire policy conversation: the region has been investing in the container and neglecting the content [1]. Chapter 1’s estimate that closing the typical PISA gap between developing and advanced economies could nearly quadruple long-run GDP is not a rhetorical flourish; it is close to the most cost-effective growth strategy available to any government in the region, precisely because so much of the potential is still sitting on the table, unclaimed.

> IN AN ECONOMY BEING REORGANIZED AROUND MACHINES THAT CAN PRODUCE FLUENT ANSWERS ON DEMAND, THE ABILITY TO READ CRITICALLY AND REASON QUANTITATIVELY IS NOT BECOMING OBSOLETE. **IT IS BECOMING THE LAST DURABLE FORM OF HUMAN COMPARATIVE ADVANTAGE.**

Now add a new variable to the equation: generative artificial intelligence. It is tempting to read the arrival of tools that can write, calculate, and reason on command as evidence that foundational skills matter less than they used to: that if a machine can decode a paragraph or solve for x, perhaps children no longer need to. The evidence assembled for this book points in the opposite direction.

A randomized trial of university students who used ChatGPT to study a new subject found they retained roughly eleven percentage points less information 45 days later than peers who studied without it. The largest drops occurred in highly technical material, precisely the content where students felt most tempted and justified to offload their learning [2].

The tool did not fail the students; it did exactly what it was designed to do, which was to remove the productive struggle that makes learning stick. Artificial intelligence is not a substitute for foundational skills. But it is a test of how strong or weak they already are. Every technology that promises to think for us handsomely rewards those who can already think without it and strongly penalizes those who cannot yet tell a good answer from a mediocre one.

In an economy being reorganized around machines that can produce fluent answers on demand, the ability to read critically and reason quantitatively is not becoming obsolete. It is becoming the last durable form of human comparative advantage. In other words, foundational skills are not one item on a long list of contemporary challenges facing the region. They are the precondition for addressing them.

Why Reading and Math Are Some of the Most Human Things We Do

To understand why foundational skills matter so much, we must look at how the brain processes them—a topic covered in Chapter 3.

Let's start with an odd fact: reading is barely six thousand years old, far too recent for evolution to have built a dedicated reading organ. And yet, given the right instruction, nearly every human brain can learn to do it. But learning to read is not natural in the way learning to speak is. Unlike oral language, for which humans are biologically prewired and which children absorb simply through exposure, reading has no innate circuitry behind it: it demands explicit, systematic instruction and sustained, repeated practice, or it does not happen at all. This is not merely an interesting footnote in neuroscience. How we teach children to read matters more than we think, especially in a world saturated with screens and artificial intelligence. A skill this effortful to construct in the first place is also unusually easy to leave half-built.

The trick is a kind of neural improvisation. As a child learns to read, a patch of the brain's visual cortex (one that evolved for something else entirely, likely for recognizing faces and objects) is recruited and retrained into what neuroscientists call the visual word form area, its connections rewired to the language regions that process meaning. The brain does not grow a new organ for reading; it renovates an old one, borrowing circuits that were never designed for the job. Reading, in this sense, is not something the brain does naturally, the way it processes a spoken sentence or a human face. It is something the brain must be taught to do, deliberately and explicitly, because nothing in our evolutionary history prepared it for the task.

Mathematics tells a subtler story. Unlike reading, it rests on something genuinely innate: newborns, and many animals, can already perceive and compare small quantities. But that primitive sense of "more" and "fewer" is a long way from solving an equation. On top of this ancient scaffolding, every child must build an entirely invented system (symbols, place value, the number line, arithmetic) that no brain arrives with pre-installed. Reading and mathematics turn out to be two versions of the same deeper story: **a species whose greatest evolutionary trick was not any single hardwired skill, but the capacity to take a brain built for one purpose and repurpose it, generation after generation, to run the software of civilization.** Chapter 1 makes the case at the level of economies: it is the accumulation and transmission of ideas, not machines or minerals, that produced sustained growth wherever it has occurred. Chapter 3 shows where that transmission actually happens: inside a single developing brain, one deliberately taught connection at a time.

Literacy and numeracy are the interface between an ancient brain and a modern world. That is what makes these two skills distinctively, irreducibly human, and why leaving their construction to chance is not a minor pedagogical oversight. It is a decision to leave the species' signature capability half-developed in millions of children.

Warning Signs

If foundational skills are this important, the region should be watching its trend lines with some anxiety.

It should start outside LAC: new generations are reading less. Data from the United States' National Assessment of Educational Progress Long-Term Trend series, visualized in a widely circulated *Financial Times* chart, show a startling reversal: in 1984, roughly 35 percent of American teenagers read almost every day for pleasure and only 15 percent said they hardly ever read. By 2023, those lines had crossed to about 14 percent reading daily, and up to nearly 47 percent who almost never read at all [3]. A generation that once read as a matter of habit has largely stopped to do so, and the decline runs alongside the very years in which pocket-sized rectangles of instant, algorithmically curated content spread everywhere.

Mathematics offers no consolation. OECD's 2022 PISA assessment recorded the steepest drop in mathematics performance since the test began a quarter of a century ago, roughly three times larger than any previous three-year change, pulling average scores across wealthy economies down to levels last seen at the turn of the century [4]. One in four 15-year-olds across OECD countries is now a low performer in mathematics, but the decline has not been evenly shared: almost everywhere it is students already struggling the most who slid the furthest, widening a gap that was already too wide.

LAC is not immune to these trends. If anything, Chapter 2 shows the region experiencing a more severe version of both. In PISA 2022, 71.2 percent of the region's students scored below the minimum proficiency level in mathematics, while only 0.4 percent reached the highest band of performance: a talent pipeline for innovation that is, for practical purposes, almost empty. In reading, an average of 55 percent of the region's 15-year-olds cannot identify the main idea of a moderately long text, even after the least prepared teenagers have already been filtered out of the sample by dropping out earlier. And the story does not begin in secondary school: by sixth grade, an average of 70 percent of the region's children already fall below minimum reading proficiency, and 83 percent below minimum proficiency in mathematics, up from 45 and 49 percent respectively in third grade. These are gaps that open early and widen with every year a child remains unsupported. The region has spent two decades getting children into classrooms, but it has made far less progress teaching them, once there, to read and calculate.

These are not two unrelated crises unfolding in parallel. They compound each other. A generation reading less and doing worse in math is precisely the generation now gaining access to tools that can read and calculate for it, on command, at the tap of a screen. Weak foundations and abundant shortcuts are, taken together, a far more dangerous combination than either trend on its own.

Turning a Challenge into an Opportunity

This book turns from diagnosis to something even more useful: a rare degree of consensus about what actually works. Chapters 4, 5, and 6 draw on syntheses of more than 270 rigorous impact evaluations of mathematics, literacy, and tutoring programs from around the world, and the patterns that emerge are consistent enough to act on.

In **mathematics**, explicit, highly structured teaching (clear goals, sequenced explanations, guided practice, and constant checks for understanding) produces roughly three times the learning gains of looser, moderately structured approaches, and this holds regardless of a country's income level.

In **literacy**, code-based instruction that systematically teaches the relationship between letters and sounds outperforms meaning-based approaches, with an advantage in writing and spelling five times larger. And "balanced" approaches that try to do a bit of everything (the dominant model across Latin America and the Caribbean today) underperform both approaches.

When students have already fallen behind, **remedial and tutoring programs** that make a clear curricular choice -either staying on grade level or explicitly rebuilding foundational skills, rather than blending the two- consistently beat programs that try to do both at once. And well-trained non-specialist tutors in charge of small groups between two and six students can rival or beat one-on-one instruction from a certified teacher, at a fraction of the cost.

None of this requires the region to invent anything new. It just requires it to stop instructional approaches that the global evidence, taken together, rule out, and to adopt, at scale, the small number of practices that have proven to be reliable. That is a far more tractable problem than closing a productivity gap with East Asia through industrial policy alone, and it happens to be a precondition for succeeding at that larger task, too.

Make or Break: Why Implementation is Crucial

There is, however, a trap built into all this good news, and Chapter 7 identifies it: knowing what works is not the same as making it happen. Esther Duflo's distinction between the scientist, who discovers what is true; the engineer, who designs a program around it; and the plumber, who has to install that program in a real building, with its cracked pipes, unexpected pressure, and joints that do not quite fit, captures something the rest of this book cannot fully teach on its own [5]. LAC has invested in knowing what works and designing programs around it (the science and engineering part of it), while treating the plumbing, the actual installation in classrooms with rotating teachers, delayed textbooks, and outgoing governments, as a mere administrative detail.

There is a fitting parallel here, one that Chapter 3's neuroscience makes almost impossible to miss. A child does not know she can read until she decodes a sentence she has never seen before; abstractly understanding the rules of phonics, or having watched someone else read fluently, proves nothing. The skill exists only in the act of using it, in material that was not rehearsed in advance. Policy is no different. A ministry can design a literacy program grounded in flawless evidence, train its architects to recite the science of reading from memory, and still not know whether the program works until it reaches an actual classroom it was not built for: one with a teacher who was never taught how to teach reading, or a delivery truck with textbooks that has not yet reached a rural school.

Comprehension, in a child or in a country's political leaders, is never established by the quality of the instruction someone received in theory. It is earned only by performance, under real and unchosen conditions. This is precisely why the 17 implementation interviews behind Chapter 7 find that every one of the ten most common barriers to literacy programs sits not in program design, but in the terrain and the people who must carry it out. This is also why the four most effective countermeasures identified are in the realm of plumbing: anticipating the school calendar before signing a contract, building alternate supply routes before the rains cut off a road, treating a tutor as a peer rather than an inspector, and returning assessment results to a teacher before her students have moved to the next grade.

The region does not merely need better designs. It needs more plumbers: people willing to enter the building, find where the pipes are actually blocked, and make the unglamorous adjustments that no pilot, however rigorous, can fully anticipate.

The Foundation Is the Structure

Return, for a moment, to the seven-year-old looking at a page she cannot yet read. Everything this book has to say about GDP convergence and Gini coefficients, about code-based phonics and small-group tutoring, or about contractual delays and rural supply chains exists for her sake, and for the tens of millions of children like her across the region.

If the next decade is to look different from the last two, the region's leaders will need to treat foundational learning with the same rigor they bring to monetary policy or trade negotiations: teach by evidence, not by habit. Because the phonics that unlock reading and the modeled reasoning that does so for mathematics are no longer a matter of opinion. Explicit and structured instruction matched to a student's skill and grade; targeting support for struggling students, choosing one clear track for every struggling learner, and letting technology extend good teaching rather than substitute for it, are key elements for moving forward. Another key input is the discipline of an engineer who never forgets she is also a plumber who needs to protect a non-negotiable core of sequence, materials, and coaching, while investing as much in winning teachers' trust as in perfecting the design, because that is where good programs often fail.

However, none of this will happen as long as literacy and numeracy are treated as a school problem rather than an economic one. Poor foundational learning skills are the root of the region's twin failures of low productivity and high inequality, which should make them a topic for finance ministries and experts in growth strategies, as much as for ministries of education. Evidence shows that **among the highest-return social investments any government can make, one that very nearly pays for itself, is investing in building core foundational learning skills among the entire population.**

What the region cannot afford anymore is to keep discounting the cost of doing nothing: every year a classroom goes untouched by this evidence, a child's window narrows, a generation's inequality is locked in a little further, and the next round of policy loses the chance to learn from having tried. The evidence exists; the returns are proven; what remains is the will to build on it, one classroom, one decade and one government at a time.

The evidence assembled in the chapters that follow says something rare in development economics: that the region already knows, with real confidence, what the classroom needs to look like. It knows that closing the learning gap would do more for the region's long-run prosperity than several other policies combined. It knows that the very technologies now arriving in our lifetime will reward a student's literacy and numeracy more, not less, than they rewarded her parents. And it knows that the difference between a program that changes a student's life and one that merely looks good in a policy memo will be decided not in an office in a ministry, but in the details of what happens between the moment a textbook is printed and the moment it reaches the hands of schoolchildren.

None of this is decided in the abstract, and none of it is decided once. It will be decided the way the seven-year-old girl mentioned at the beginning will learn to read and count: not by anyone's intention, however evidence-based, but by what actually happens on the ground. That is the argument this book makes, chapter by chapter, and the reason it is called Built to Last. A region's future is not built on the foreign investments it hopes to attract or the technologies it hopes to adopt. It is built, one classroom at a time, on whether a child can count and read the sentence in front of her. And once she can, everything that follows might change the fortunes of the region.

REFERENCES

[1] Hanushek, E.A. and Woessmann, L., The Role of Cognitive Skills in Economic Development. Journal of Economic Literature, 2008, 46(3): p. 607–668.

[2] Barcaui, A., ChatGPT as a Cognitive Crutch: Evidence from a Randomized Controlled Trial on Knowledge Retention. Social Sciences & Humanities Open, 2025, 12: 102287.

[3] Burn-Murdoch, J., The Decline of Reading. Financial Times, 2024 (drawing on the U.S. National Assessment of Educational Progress Long-Term Trend Reading Assessments).

[4] OECD, PISA 2022 Results (Volume I): The State of Learning and Equity in Education. Paris: OECD Publishing, 2023.

[5] Duflo, E., The Economist as Plumber. NBER Working Paper No. 23213, 2017. Cambridge, MA: National Bureau of Economic Research.

Part I

Why Foundational Skills Matter

IT IS NOT ONE INVESTMENT AMONG MANY. IT IS THE ONE ALL OTHER INVESTMENTS DEPEND.

CHAPTER 1

LOOKING FOR A HIGH RETURN ON SOCIAL INVESTMENT? BET ON TEACHING FOUNDATIONAL SKILLS TO REDUCE INEQUALITY AND DRIVE GROWTH IN LATIN AMERICA AND THE CARIBBEAN

Dante Contreras & Rafael de Hoyos

Latin America and the Caribbean (LAC) occupy a paradoxical position in the global economy. The region is endowed with abundant natural resources, a large working age population, and increasing access to global markets and technologies. Yet, for decades, it has been unable to translate these endowments into sustained convergence toward the living standards of advanced economies. Two structural challenges stand out: persistently low labor productivity, and therefore low economic growth, and stubbornly high-income inequality.

The evidence on productivity is striking.

Figure 1.1 shows the evolution of Gross Domestic Product (GDP) per capita in LAC relative to the United States, the European Union, and East Asia over the period 2000-2023. While East Asian countries have achieved a remarkable convergence —rising from roughly 50 percent to over 70 percent of US GDP per capita [1]— LAC has remained stagnant, hovering around 25 percent throughout the first quarter of our century. The contrast is not cyclical but structural: the region remains trapped in a low-productivity equilibrium [2].

LAC is the most unequal region in the world, and income inequality is highly persistent. The Gini coefficient, a common measure of inequality, declines from 0.53 to 0.45 in LAC between 2000 and 2023, a notable improvement.

Still, it remains far above the levels observed in East Asia (around 0.34) and the European Union (around 0.31). Inequality in the region is not only high in absolute terms but also persistent, suggesting limited social mobility, with advantages and disadvantages carried across generations.

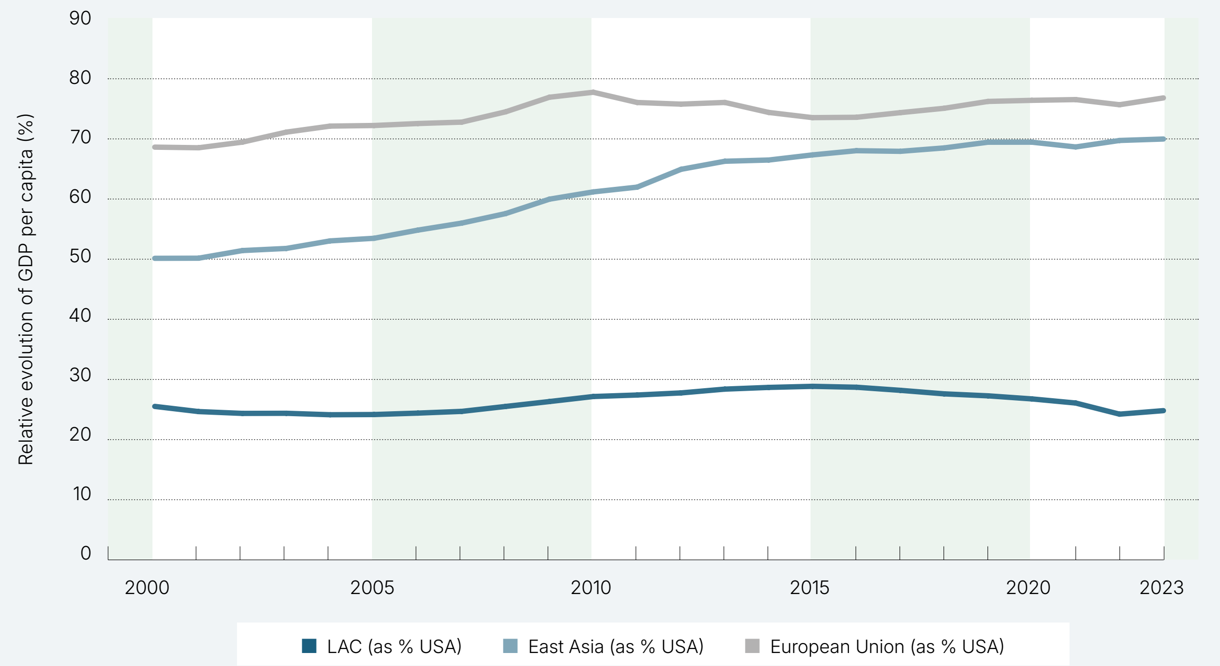
Both challenges —low productivity and high inequality —share a common structural root: the failure to build sufficient, high-quality human capital.
This leads to a lack of skilled workers needed for innovation and productivity growth, and it perpetuates the socioeconomic disadvantages of low-income families by denying their children the skills needed to rise above the circumstances of their birth. Understanding this link is the first step toward designing policies that address both challenges simultaneously.

WHAT DOES HISTORY TELL US ABOUT THE ROLE OF FOUNDATIONAL SKILLS IN ECONOMIC GROWTH?

The relationship between knowledge, human capital, and economic prosperity is one of the most robust findings in economic history. As the economist Oded Galor shows in an extensive analysis of humanity's economic trajectory [3], the transition from centuries of stagnant living standards (the so-called Malthusian trap, where population growth outpaced agricultural production, leading to an ever-repeating cycle that led back to subsistence-level economic standards), to sustained growth was triggered, above all, by the accumulation and diffusion of ideas.

The Industrial Revolution was not primarily a story of new machines; it was a story of the human capacity to generate, transmit, and apply new knowledge.
The countries that industrialized earliest were those that had already invested in education and human capital formation, enabling their populations to innovate, adapt to, and adopt new technologies.

FIGURE 1.1: FOREVER TRAPPED?
Productivity in LAC remains stagnant while GDP per capita in East Asia is catching up relative to the United States (2000-2023)



Source: Authors' elaboration based on data from the International Monetary Fund.

A culture that valued practical knowledge was the hidden engine behind Europe's economic takeoff. The European Enlightenment, for example, created a "culture of growth" in which knowledge, experimentation, and innovation became closely linked to material progress [4]. The spread of scientific thinking and the exchange of ideas through scholarly networks such as the "Republic of Letters" helped create the conditions for the Industrial Revolution. More broadly, historical evidence suggests that sustained growth emerged not simply from technological discoveries, but from societies' ability to diffuse and apply knowledge productively.

In a systematic survey of competing theories of the Great Divergence, Mark Koyama and Jared Rubin synthesize the extensive literature on geography, institutions, culture, colonial history, and human capital to explain why sustained economic growth emerged in Western Europe and North America, when it happened, and why it spread, albeit unevenly, thereafter [5]. Their central finding is that sustained growth requires a set of enabling conditions that must operate jointly. For example, political and legal institutions that protected property rights and constrained arbitrary rule were essential for fostering innovation-driven incentives. Cultural orientations that valued practical knowledge and experimentation fueled the underlying drive for innovative endeavors. But crucially, the authors argue that human capital does not just correlate with growth; it constitutes the very mechanism that generates and sustains it.

Societies that failed to invest in developing their populations' cognitive capabilities could not generate the innovators, adopters, and skilled workers that growth required, regardless of their geographic endowments or natural resources. History's verdict seems clear: it is people, not places, that drive prosperity.

This historical perspective finds rigorous empirical support in the modern growth literature. In a series of landmark studies, Eric Hanushek and Ludger Woessmann show that what matters for long-term economic growth is not merely the years of schooling achieved by the population, but the cognitive skills —the actual knowledge and reasoning capacity— that individuals acquire [6, 7]. Using internationally comparable test scores as a measure of human capital quality, they show that differences in cognitive skills explain between 75 and 79 percent of the cross-country variation in economic growth. Once cognitive skills are controlled for, years of schooling no longer predict growth. The implication is profound: it is not schooling per se, but learning, that drives prosperity.

The finding that education generates large economic and social returns is consistent with the broader literature. Global estimates put the private wage return at 8.8 percent per year of schooling, rising to 11 percent in LAC [8], while human capital has been shown to account for roughly 45 percent of global growth over the past four decades, with an outsized share concentrated among the poorest households [9, 10].

Drawing on the 'distance-to-frontier' framework, which measures the gap between a country's performance and the best practices (the 'frontier'), one research shows that the returns to different levels of education depend on a country's stage of development [11]. For economies far from the technological frontier, as most LAC countries are, strong foundational and secondary education is a critical investment, since it facilitates adopting existing technologies and imitating best practices. As countries approach the frontier, higher and graduate education become increasingly important in creating and expanding new knowledge. But crucially, education and innovation are complementary at all stages: education raises the capacity to adopt and create technological advances, while innovation raises the returns to education. Countries that fail to invest early in foundational skills get trapped in a low-education, low-innovation equilibrium.

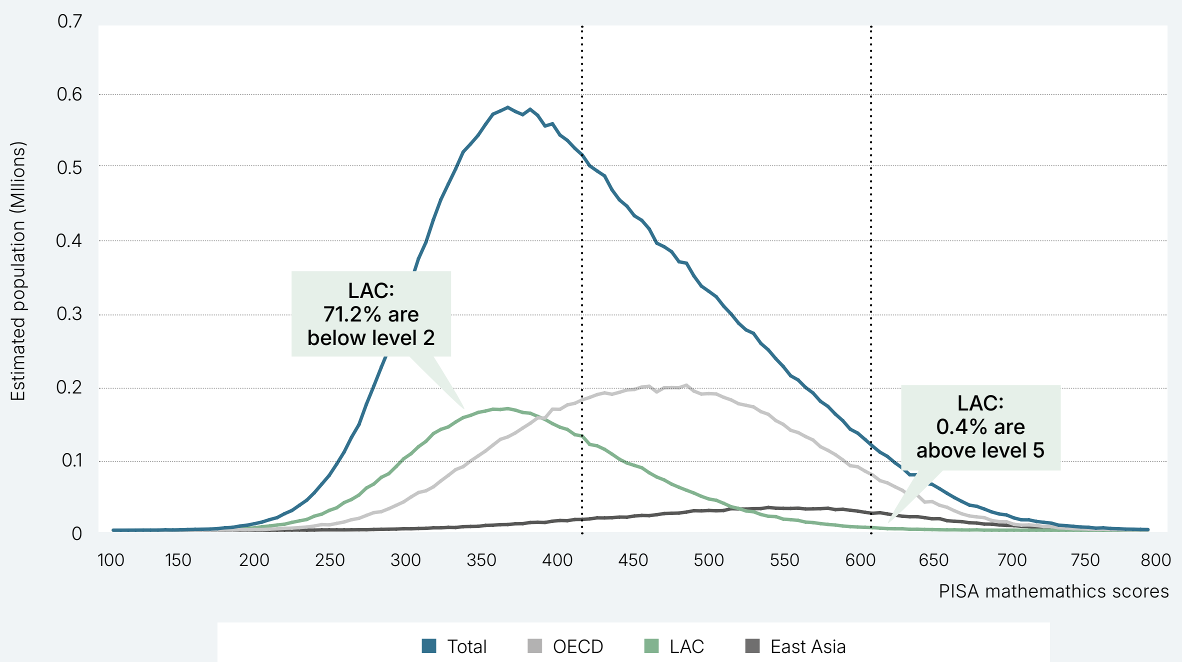
Countries with sustained growth have expanded learning opportunities across their entire population. The United States achieved technological leadership through a combination of early and universal expansion of secondary and higher education, a world-class research university system, and institutional environments that rewarded productivity and innovation. China's economic miracle combined strong foundational education with pro-market reforms that raised the returns to skills. Finland's transformation into one of the most innovative economies in the world was underpinned by an inclusive, high-quality education system that systematically expanded the pool of potential innovators by eliminating socioeconomic barriers to educational attainment [12].

The implication for LAC is clear: growth depends on whether education systems generate real learning. The central question no longer is exclusively how much and how well the region's countries invest in education, but whether their education systems produce the foundational skills required for productivity, innovation, and social mobility.

The evidence shows that the region has a 'learning crisis' and does not generate the foundational skills among its population needed to sustain productivity growth and expand social mobility.
Figure 1.2 shows the distribution of PISA 2022 mathematics scores across LAC, and East Asian countries, as well as the global total. The picture is bleak. Whereas the OECD distribution peaks around 450-500 points, East Asia's distribution is shifted considerably to the right, and the LAC, OECD, and East Asian distribution is concentrated at the low end of the scale [13]. Fully 71.2 percent of LAC students score below PISA Level 2, the threshold defined by the OECD as the minimum level of proficiency needed to participate effectively in modern economies and societies. At the other extreme, only 0.4 percent of LAC students reach Level 5 or above —the threshold of excellence at which students can engage in complex reasoning, generalization, and creative problem-solving.

These learning deficits have consequences for both growth and inequality. At the macroeconomic level, it constrains innovation and productivity growth, and at the microeconomic level, it perpetuates income inequality across generations.

FIGURE 1.2: NUMBERS THAT DON'T ADD.
Most students in LAC do not reach minimum proficiency in mathematics (Distribution of PISA 2022 mathematics scores by region).



Note: Vertical dashed lines indicate Level 2 (insufficient, score 420) and Level 5 (excellent, score 607). The density below each curve represents the share of students in each score.
Source: Authors' elaboration based on PISA 2022 (OECD), using population weights.

HOW DO FOUNDATIONAL LEARNING DEFICITS CONSTRAIN INNOVATION?

■ LAC is producing too few high-performing students to sustain innovation-led growth.
The near absence of students performing at the highest proficiency levels has direct macroeconomic implications. Low foundational skills constitute a binding constraint on innovation in the region: they limit firms' and workers' capacity to adopt and effectively use new technologies, which in turn holds back productivity growth. The very thin right tail of the LAC PISA distribution reinforces this constraint, signaling a limited pipeline of individuals equipped with the advanced quantitative and problem-solving skills needed to absorb and adapt frontier technologies. Evidence shows that it is not only average skills, but also the share of the population with advanced cognitive skills, that drives innovation and long-run economic growth [14].

■ Learning gaps usually emerge early and shape a country's future innovative capacity.
This has been documented. Examining U.S. patent data, a team of researchers from several institutions identified the so-called "Lost Einsteins" phenomenon: individuals with exceptional inventive potential who never realize it due to socioeconomic, gender, and early academic barriers [15].

Their findings show that for every 1,000 children who excel in third-grade math, men go on to register 6.0 patents, whereas women register 2.4. In contrast, among children with low third-grade math scores, these figures free-fall to 0.5 and 0.1 patents per 1,000, respectively. Mathematical proficiency acquired in elementary school is not just a predictor of labor-market success; it is the seed that enables inventive capacity and, aggregated across a generation, a country's capacity to innovate. For LAC, a region where 71.2 percent of students fail to reach even the basic mathematics proficiency threshold, this evidence implies that the region is systematically foreclosing the innovative potential of most of its population before they even reach secondary school.

■ In Latin America and the Caribbean, high inequality and failing schools amplify the 'Lost Einsteins' phenomenon.
The region's talent pool is trimmed at both ends, and this dual failure places LAC at a structural disadvantage relative to economies that have raised both the floor and the ceiling of their skill distributions.

■ Learning opportunities in the region remain deeply unequal.
Figure 1.3 shows the distribution of PISA 2022 mathematics performance broken down by household resources. Among students from households with limited resources, defined as those below the median on PISA's Index of Economic, Social and Cultural Status (ESCS), 75.3 percent score below Level 2, and only 0.2 percent reach Level 5 or above. Among students from households with resources at or above the OECD median, the situation is better but still alarming: 41.6 percent score below Level 2, and 1.9 percent reach Level 5 or above.

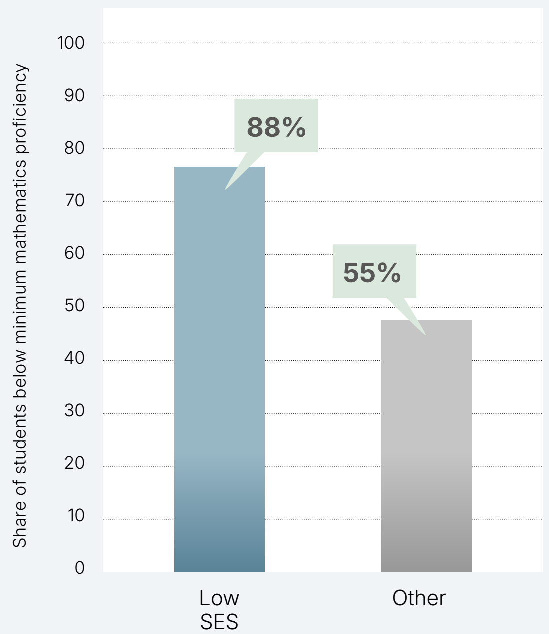
■ Unequal learning outcomes are one of the key mechanisms through which inequality is passed down across generations.
A child born into a low-income household in LAC will likely attend a lower-quality school, receive less learning support at home, and graduate without the cognitive skills necessary to obtain a wage premium in the modern workforce. Without those skills, they will earn low wages as an adult, and their own children will possibly face the same cycle. Hence, the learning gap is not merely a symptom of inequality; it is the mechanism that perpetuates it.

■ Large learning gaps effectively exclude many low-income students from the knowledge economy.
The World Bank's World Development Report 2018 documents the magnitude of this problem across the region.

For instance, in Uruguay, the proportion of sixth-grade students from the bottom income quintile who are classified as 'not proficient' in mathematics is five times higher than that of students from the top quintile. In Honduras, half of all students from the poorest quintile are at the very lowest reading level, compared to only 7 percent from the top quintile [9]. These are not marginal differences. It represents an almost complete exclusion of low-income children from the knowledge economy, condemning them to a lifetime of low productivity and low income.

Failing to develop talent across the population leads to a massive misallocation of talent, and the cost is not merely distributive but an economy-wide efficiency loss.
When access to quality education is determined by family income rather than individual ability, a large fraction of potential innovators and high-skilled workers is lost to low-productivity activities, leaving a society's most valuable asset untapped [11]. This human capital deficit is most pronounced in deeply unequal societies, which suffer the greatest disparity between inherent population potential and actualized skills.

FIGURE 1.3: POOR = FEWER BASIC SKILLS?
Students from low-income households are far less likely to acquire the skills needed for economic mobility (distribution of PISA 2022 mathematics performance in LAC by socioeconomic resources of the household)



Note: "Low SES" refers to students with low socioeconomic status, defined as those below the OECD median on PISA's Index of Economic, Social and Cultural Status (ESCS); "Other" refers to students at or above the OECD median. Bars show the share of 15-year-old students in LAC scoring below Level 2 (minimum proficiency) and at Level 5 or above (top proficiency) in mathematics. Source: Authors' elaboration based on the Inter-American Development Bank's Centro de Información para la Mejora de los Aprendizajes (CIMA), using PISA 2022 (OECD) data.

INVESTING IN AN EDUCATION MODEL THAT GENERATES LEARNING IS PROFITABLE

Investing in foundational skills is not only a matter of social justice; it is also one of the soundest economic investments any government can make. The returns to education —at a private, social, and fiscal level— are among the highest observed for any public policy intervention, particularly over the long run.

Education generates large and sustained returns for individuals and societies.
These returns have historically outperformed financial markets: between 1966 and 2015, the average return on stocks and bonds in the United States was 2.4 percent, compared to an average private return to education of 10.5 percent. By any measure, education is a sound investment [16].

The social rate of return to education, meaning the private wage return plus the broader external benefits of schooling such as improved health, reduced crime, greater social cohesion, and higher civic participation, averages 13.3 percent globally per additional year of schooling.

The benefits of education are particularly important for lower-income populations.
Distributional analysis shows that educational investments have historically been progressive [10]. Gethin's

estimates imply that without these educational gains, global inequality would be roughly 20 percent higher, underscoring the redistributive power of well-targeted investment in a region as unequal as LAC.

DO INVESTMENTS IN FOUNDATIONAL SKILLS PAY FOR THEMSELVES?

Yes, in fact, they generate substantial fiscal returns. The math is straightforward a country that invests in improving the mathematics and literacy skills of its schoolchildren raises the productivity in the long run and, therefore, the taxable income of a new and skilled workforce. If the present value of those additional tax revenues, discounted at an appropriate social rate, exceeds the cost of the investment, the program pays for itself.

■ **Evidence suggests that, for well-designed interventions capable of improving foundational skills, this is actually the case.**
Data from cross-country analyses show that sustained increases in education spending, especially quality education, boost long-term economic performance. Using data from 107 countries for the period 1970 to 2019, a group of economists from the Universidad de Chile and the Inter-American Development Bank estimate that a sustained 10 percent increase in per-capita primary or secondary education expenditure is associated with a 0.19 percent increase in long-run economic output [17]. Accumulated across generations of workers, these gains translate into higher taxable income, greater consumption, and lower expenditure on poverty-related social programs.

■ **Remarkably, this research shows that the 'quality of learning' acts as a fiscal multiplier.**
A country that closes the typical PISA score gap between developing and advanced economies (approximately 30 percent) can almost quadruple its long-term GDP and therefore its fiscal resources. This is a critical insight for LAC policymakers: the fiscal return to education spending depends not on how much is spent, but on how much learning is generated per dollar spent. A school system that absorbs resources without generating cognitive skills improvement produces negligible fiscal returns; one that translates expenditure into genuine learning gains generates returns that, over the long run, more than cover the public investment.

■ **Some of the most cost-effective public investments are those that improve foundational learning.**
Applying a unified framework to evaluate government programs across the United States, one study finds that targeting early childhood education and K-12 learning for disadvantaged populations consistently rank among the most efficient uses of public funds across the entire portfolio of government expenditure. [18] The mechanism is straightforward: an intervention that raises a low-income child's cognitive skills in early childhood generates returns compounded over an entire working life, while its costs are incurred only once.

■ **In other words, programs that maximize learning per dollar spent generate the largest returns.**
A synthesis of cost-effectiveness from randomized evaluations around the world, compiled by the MIT-based Abdul Latif Jameel Poverty Action Lab (J-PAL), reaches a consistent conclusion [19]. The common thread across the most cost-effective programs is that they are explicitly designed to maximize learning per dollar spent, rather than to expand enrollment or schoolroom inputs. So, access without quality education generates returns well below the potential. Given that the learning crisis described above is precisely a quality problem —millions of children are enrolled in schools but not really learning— this finding is actionable evidence for LAC's policymakers.

Over the long run, investments in foundational skills that improve learning do not represent a net cost for governments. If policymakers give sufficient weight to the future —applying a social discount rate that reflects the horizon over which educational investments pay off— spending on quality primary and secondary education that raises cognitive skills is not a budgetary burden. It is, according to the evidence, the most financially sound investment a country can make: one that raises productivity, expands the future tax base, reduces inequality, and largely pays for itself.

LEARNING AND ACQUIRING FOUNDATIONAL SKILLS: THE KEYS TO GROWTH, EQUITY, AND SOCIAL MOBILITY

Latin America and the Caribbean's low productivity and high inequality share a common root: insufficient learning. The evidence reviewed throughout this chapter —from economic history and growth theory to PISA data and the economics of education— points to a consistent conclusion: countries grow when they develop the foundational skills of their populations. Some key aspects reviewed in this chapter that could provide inputs and insights for the region's policymakers include:

- **LAC is not generating the foundational skills needed for productivity, growth and social mobility.**
More than 70 percent of students in the region fail to reach minimum proficiency in mathematics, and only 0.4 percent meet the threshold of excellence associated with innovation and complex problem-solving. At the other end of the distribution, the vast majority of low-income students are condemned by inadequate learning to a lifetime of low wages and limited opportunity, perpetuating the inequality that is already the region's most debilitating structural feature.

- **Investing in foundational skills is one of the soundest economic investments governments can make.**
Private returns to education consistently exceed those of the financial markets. Social returns are higher still. Over the long run, fiscal returns are positive and substantial: for every dollar invested in quality education, the state recovers multiple dollars through higher tax revenues and lower expenditures on poverty, crime, and social protection. Given sufficient weight to the future, learning investment pays for itself. The keyword is 'learning'; it is not schooling per se but genuine cognitive skill formation that generates these returns.

- **Every year that a child in the region spends in school without acquiring foundational skills is a year of lost human potential — not just for that child, but for the entire society.**
The conclusion is therefore forceful: today, thanks to the accumulated evidence of history and decades of careful empirical research, it is clear what is needed to ignite growth, eliminate poverty, and promote social mobility in LAC. Focusing on learning is key. Every generation of low-income students that emerges from the education system without the skills to participate productively in the economy is another generation in which inequality is reproduced and locked in.

- **The greatest economic waste in a highly unequal society is the failure to develop the potential of its people — its only truly irreplaceable asset.**
LAC's talent is not the problem. The region's children and young people are as capable as any in the world. The problem is a system that consistently fails to discover, nurture, and develop that talent. Addressing this failure by building education systems that generate real learning for every child, regardless of the circumstances of birth, is not only the most compelling moral imperative facing the region but, by the evidence reviewed here, also the most profitable economic investment available to its governments. It is also, by the evidence reviewed here, the most profitable economic investment available to its governments.

REFERENCES

[1] Hanushek, E. A. and Woessmann, L. (2016). Knowledge capital, growth, and the East Asian miracle. Science, 351(6271), 344-345.

[2] Hanushek, E. A. and Woessmann, L. (2012b). Schooling, educational achievement, and the Latin American growth puzzle. Journal of Development Economics, 99(2), 497-512.

[3] Galor, O. (2022). The Journey of Humanity: The Origins of Wealth and Inequality. Dutton, New York.

[4] Mokyr, J. (2016). A Culture of Growth: The Origins of the Modern Economy. Princeton University Press, Princeton, NJ.

[5] Koyama, M. and Rubin, J. (2022). How the World Became Rich: The Historical Origins of Economic Growth. Polity Press, Cambridge.

[6] Hanushek, E. A. and Woessmann, L. (2008). The role of cognitive skills in economic development. Journal of Economic Literature, 46(3), 607-668.

[7] Hanushek, E. A. and Woessmann, L. (2015a). The Knowledge Capital of Nations: Education and the Economics of Growth. MIT Press, Cambridge, MA.

[8] Psacharopoulos, G. and Patrinos, H. A. (2018). Returns to investment in education: A decennial review of the global literature. Education Economics, 26(5), 445-458.

[9] World Bank (2018). World Development Report 2018: Learning to Realize Education's Promise. The World Bank, Washington, DC.

[10] Gethin, A. (2024). Distributional growth accounting: Education and the reduction of global poverty, 1980-2019. The Quarterly Journal of Economics, 140(4), 2571-2638.

[11] Aghion, P., Almas, I., and Meghir, C. (2026). Human capital and development. Working Paper 34602, National Bureau of Economic Research (NBER).

[12] Aghion, P., Akcigit, U., Hyytinen, A., and Toivanen, O. (2023). Parental education and invention: The Finnish enigma. International Economic Review, 64(2), 471-498.

[13] OECD (2021). Education at a Glance 2021: OECD Indicators. OECD Publishing, Paris.

[14] Hanushek, E. A. and Woessmann, L. (2012a). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. Journal of Economic Growth, 17(4), 267-321.

[15] Bell, A., Chetty, R., Jaravel, X., Petkova, N., and Van Reenen, J. (2019). Who becomes an inventor in America? The importance of exposure effects. The Quarterly Journal of Economics, 134(2), 647-713.

[16] Psacharopoulos, G. (2024). Returns to education: A brief history and an assessment. Education Economics, 32(5), 561-565.

[17] Contreras, D., Galindo, A. J., and Lepe, I. (2026). Education Spending and Economic Growth: Short- and Long-Term Effects in a Global Panel. Forthcoming. Review of Development Economics.

[18] Hendren, N. and Sprung-Keyser, B. (2020). A unified welfare analysis of government policies. The Quarterly Journal of Economics, 135(3), 1209-1318.

[19] J-PAL (2024). Cost-effectiveness and welfare analysis. Abdul Latif Jameel Poverty Action Lab (J-PAL). Available at: https://www.povertyactionlab.org/page/cost-effectiveness-and-welfare-analysis

CHAPTER

2

CRACKS IN THE FOUNDATION: LITERACY AND NUMERACY IN LATIN AMERICA AND THE CARIBBEAN

Elena Arias Ortiz & Fernanda Marquez Ragonesi

Foundational skills are among the highest-return investments a society can make. As Chapter 1 documents, literacy and numeracy shape not only school success but also later productivity, employment, civic participation, and intergenerational mobility. The returns are especially large when those skills are built early in life, a promise that has made foundational learning a central goal of education policy worldwide.

Literacy and numeracy are not simply school subjects; they are the basic capabilities that allow children and teenagers to understand information, solve problems, and progressively build more advanced skills such as reasoning, analysis, and learning autonomy. Skills acquired in childhood shape later educational achievements, labor market outcomes, and social inclusion, while deficits at early stages tend to persist over time and become increasingly difficult to remediate [1].

This chapter deals with a direct question: how much of that promise is actually being realized in Latin America and the Caribbean? Despite major gains in school access over recent decades, the region continues to face a profound learning challenge. Too many children leave the early grades without mastering basic reading and mathematics, too many teenagers progress through school without consolidating those skills, and too many adults reach the labor market without the foundational competencies needed to fully participate and prosper [2].

HOW STRONG ARE FOUNDATIONAL SKILLS ACROSS THE LIFE CYCLE?

Drawing on evidence from three major learning assessments: UNESCO'S ERCE (Regional Comparative and Explanatory Study), OECD's PISA (Programme for International Student Assessment), and PIAAC (Programme for the International Assessment of Adult Competencies), this chapter traces foundational skills across the life cycle.

It examines the current state of literacy and numeracy from primary education through adulthood; reviews trends over time; and analyzes disparities by socioeconomic background, gender, territory, and school type (For a full description of assessment methodologies and country coverage (see Box A2 in the Annex).
There are four central findings:

1. **Foundational learning levels in the region remain critically low,** especially in mathematics, and deficits appear early in the education cycle.
2. **Progress has been limited and uneven:** most countries show little improvement over time, suggesting that education systems have not been able to significantly reduce early gaps.
3. **Learning outcomes are deeply unequal,** with students from poorer households and rural areas consistently facing lower achievement.
4. **Low learning is closely linked to broken educational trajectories:** many students move through the system without mastering basic competencies in foundational skills, while the most vulnerable are also the most likely to fall behind or leave school before completion.

In all available assessments, a large share of children, teenagers, and adults do not reach a minimum proficiency in foundational skills. The evidence is clear and highlights a troubling pattern: learning deficits emerge early in primary school and persist into adulthood.

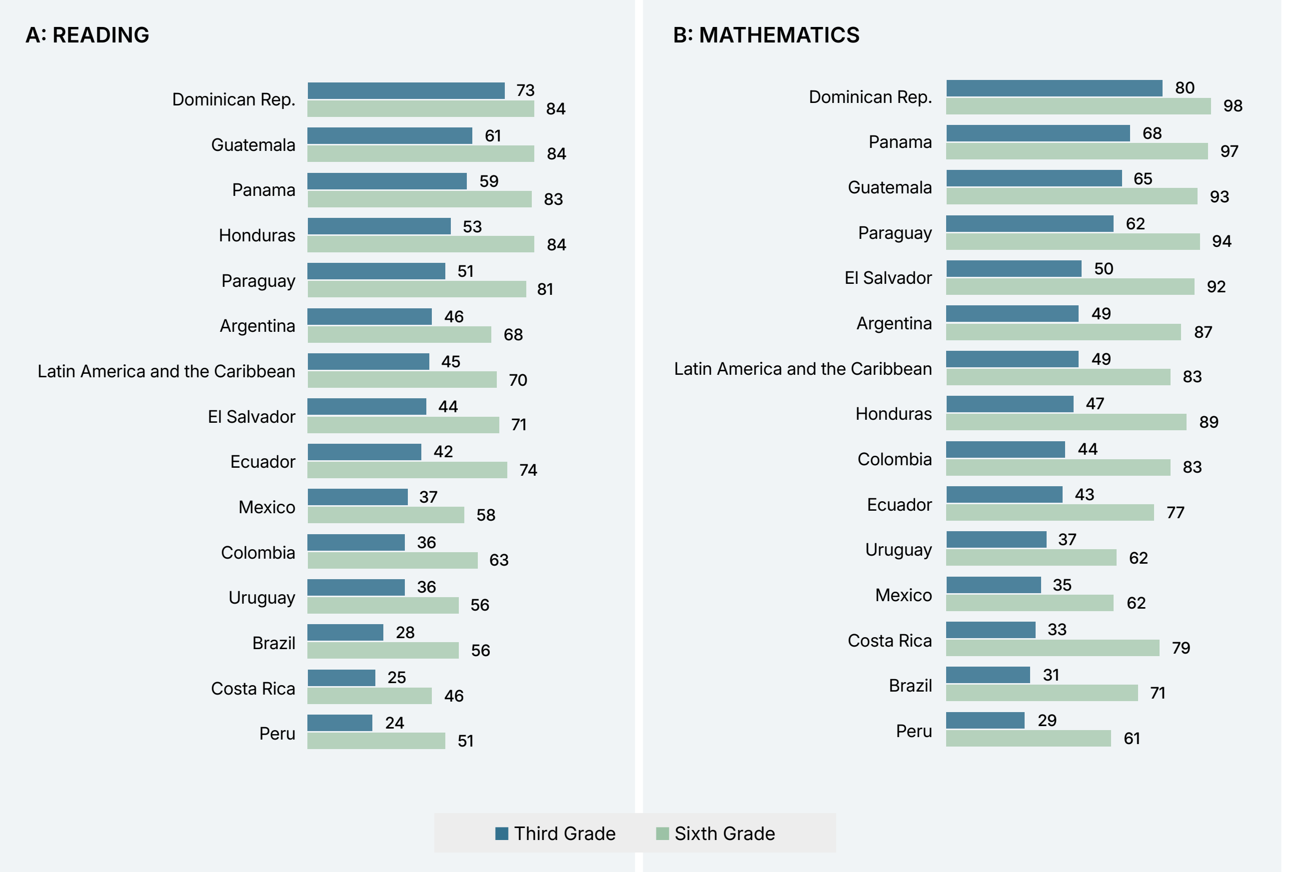
What's the Current State of Foundational Skills in Primary School?

■ **Learning deficits in reading emerge early and deepen as students progress through primary school.**
By third grade, a large share of students across Latin America and the Caribbean already perform below minimum proficiency in reading. That is, they cannot yet locate explicit information in a short text or draw simple inferences from it (see Box A2 in the Annex for the definition of proficiency levels). The highest incidence of low performance is observed in the Dominican Republic, where roughly 73 percent of third-grade students fall below the minimum benchmark in reading, followed by Guatemala. At the other end, Brazil, Costa Rica, and Peru show comparatively better results, where roughly one in four students scores below the minimum (see Figure 2.1 A).

■ **By sixth grade, the situation worsens markedly.**
On average, the share of low-performing students in reading rises from about 45 percent in third grade to around 70 percent in sixth grade. The most severe deficits are found in the Dominican Republic, Guatemala, Honduras, and Paraguay, where more than four out of five students do not achieve the expected reading proficiency for their grade. In contrast, Peru, Costa Rica, Brazil, and Uruguay record relatively stronger performance, but even in these countries, the proportion of students below minimum proficiency remains high.

■ **The results in mathematics are even more alarming.**
By third grade, nearly half of the students have already failed to reach the minimum proficiency level. The Dominican Republic records the highest share, about 80 percent below the minimum benchmark, followed by

FIGURE 2.1: HOW MANY CHILDREN ARE NOT LEARNING THE BASICS?
Percentage of students not reaching minimum proficiency in reading and mathematics in primary school, ERCE 2019



Source: CIMA (2025) based on ERCE 2019 (LLECE-UNESCO).

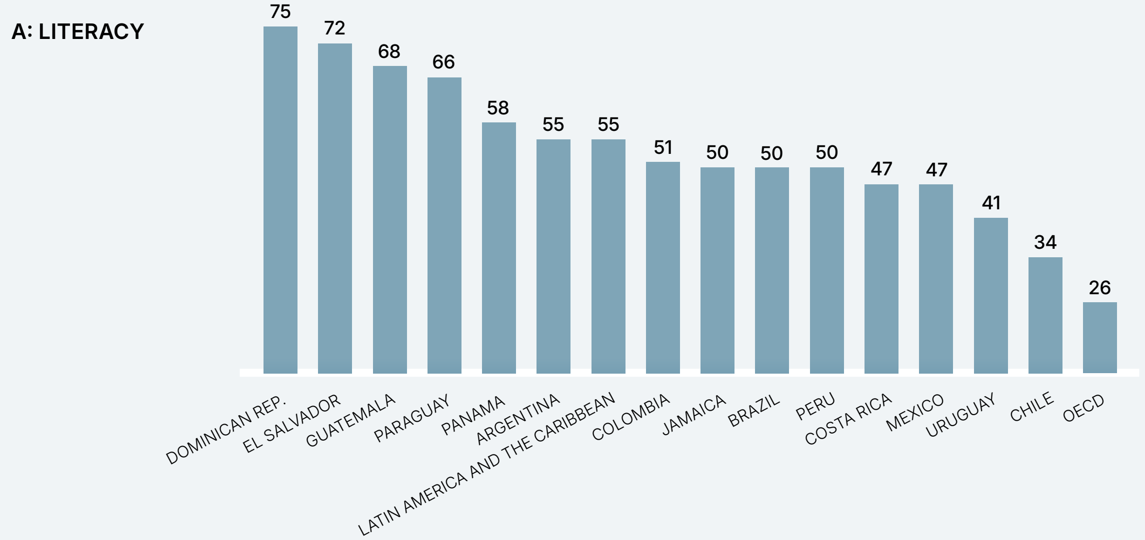
Panama and Guatemala. By sixth grade, the regional average share of students below minimum proficiency rises to 83 percent, indicating that mathematics learning gaps not only persist but widen substantially as students move through primary school (see Figure 2.1 B).

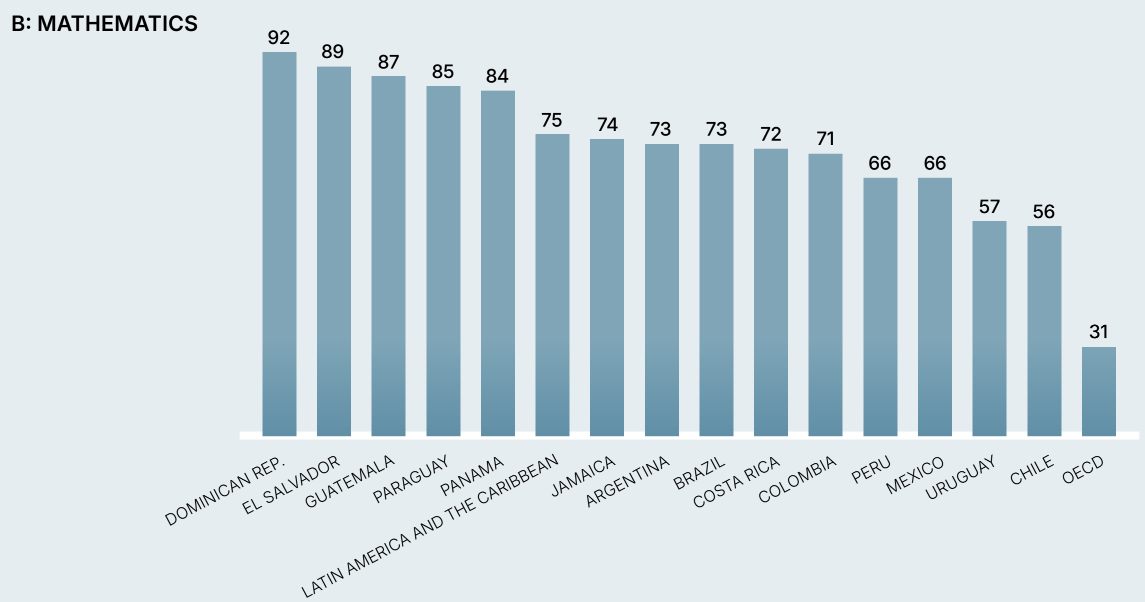
Progress in schooling has not yet been fully matched by progress in learning.
Although primary education coverage in Latin America and the Caribbean is nearly universal, with net attendance rates above 95 percent in most countries [2], these results make clear that getting children into school is not enough to ensure they are acquiring basic skills.

What's the Current State of Foundational Skills in Secondary School?

As students enter secondary education, learning outcomes remain at critically low levels. PISA assesses 15-year-olds enrolled in school who have completed at least six years of education, which means the most vulnerable teenagers, those who have dropped out or fallen far behind, are already excluded. In 2022, 89 percent of 15-year-olds in the 38 OECD countries were eligible to participate in PISA, compared to 71 percent in Latin America.[1] Even within this more advantaged group, an average of 55 percent of students in the region perform below the baseline level in reading. In other words, more than half cannot identify the main idea in a moderately long text (see Figure 2.2 A).

FIGURE 2.2: DO ADOLESCENTS FARE ANY BETTER?
Percentage of students not reaching minimum proficiency in literacy and mathematics in secondary school, PISA 2022.





Source: CIMA (2025) based on PISA 2022 (OECD).

1. Four Latin American countries are OECD members: Mexico (since 1994), Chile (2010), Colombia (2020), and Costa Rica (2021). PISA results are not directly comparable with ERCE results: each assessment defines minimum proficiency differently, so the lower shares of low performers in secondary school relative to sixth grade should not be read as an improvement between primary and secondary school (see Annex).

Box 2.1: What Do We Know About Foundational Skills in the English-Speaking Caribbean?

Countries in the English-speaking Caribbean do not consistently participate in large international learning assessments such as PISA. Instead, the region relies primarily on assessments administered by the Caribbean Examinations Council (CXC). Available evidence points to substantial foundational learning challenges. According to a 2025 World Bank analysis [3], only about 4.9 percent of students taking the Caribbean Secondary Education Certificate (CSEC) passed five or more subjects including mathematics and English. Trinidad and Tobago ranked 53rd in reading and 59th in mathematics out of 79 countries in the PISA 2015 evaluation, while Jamaica ranked 61st in reading and 62nd in mathematics in PISA 2022. These results suggest that students in the Caribbean face even larger learning gaps than those observed in most other Latin American countries [4].

The picture in mathematics is even more bleak.
An average of 75 percent of students in the region perform below the baseline level. The largest shares of low performance are found in the Dominican Republic (92 percent), El Salvador (89 percent), Guatemala (87 percent), Paraguay (85 percent), and Panama (84 percent). The lowest rates in the region are recorded in Chile (56 percent) and Uruguay (57 percent). Yet even in these better-performing countries, more than half of students cannot apply basic mathematical concepts to simple real-life situations. The gap relative to the OECD is stark: 26 percent of students in OECD countries perform below baseline in reading and 31 percent in mathematics, compared with regional averages of 55 and 75 percent, a difference of roughly 29 and 44 percentage points, respectively (see Figure 2.2 B).

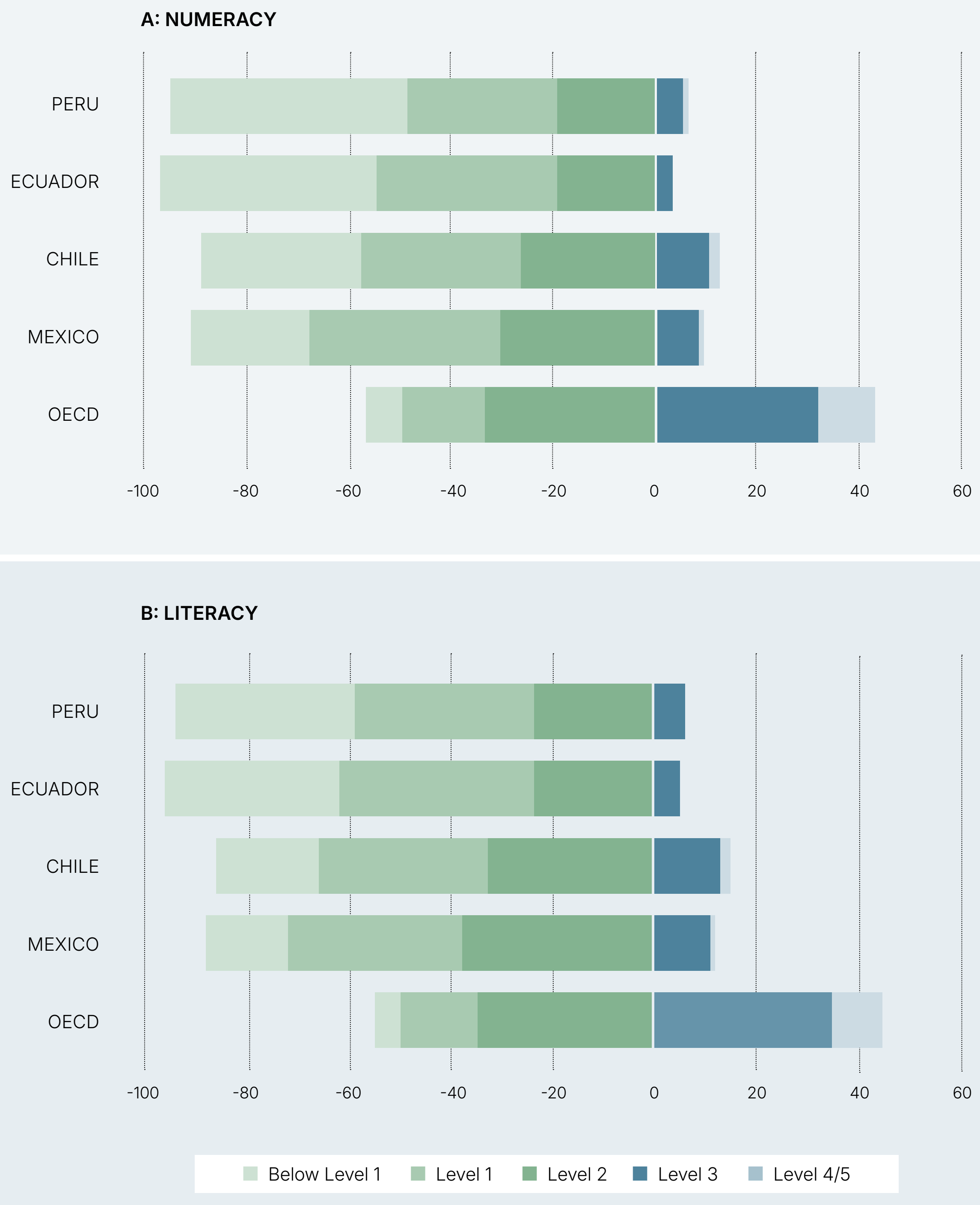
What's the Current State of Foundational Skills in Adults?

Weak learning in children and teenagers translates directly into low competencies in adulthood. Data from the first cycle of PIAAC, collected in four Latin American countries (Chile in 2014–2015; Ecuador, Mexico, and Peru in 2017), show that an average of 92 percent of adults do not reach a minimum proficiency in numeracy. The share of adults below this benchmark is particularly high in Ecuador (96 percent), Peru (94 percent), and Mexico (91 percent). Chile shows a slightly better but still strikingly low figure at 88 percent. Although 56 percent of adults in OECD countries also do not reach baseline numeracy proficiency, there is still a 36-percentage-point gap relative to LAC countries.

In all four participating countries (Chile, Ecuador, Mexico, and Peru), more than half of adults perform at the lowest levels of the numeracy scale (Level 1 or below; see Figure 2.3 A), meaning they can only carry out simple tasks: counting, sorting, or basic arithmetic in familiar contexts. A similar pattern holds in literacy. A large proportion of adults do not reach baseline literacy proficiency in Ecuador (94 percent), Peru (93 percent), Mexico (88 percent), and Chile (85 percent). These figures reflect the cumulative effects of weak foundational education. Adults with tertiary education in these countries often perform at levels comparable to, or even below, adults with only upper secondary education in other emerging countries [5] [6].

Taken together, the data points to a consistent and troubling trajectory: foundational learning deficits emerge early in primary school, remain widespread through secondary education, and continue into adulthood, underscoring the structural gap in the region's learning challenge.

FIGURE 2.3: DO LEARNING GAPS FOLLOW STUDENTS INTO ADULT LIFE?
Percentage of adults scoring at each proficiency level in numeracy and literacy, Survey of Adult Skills, PIAAC first cycle (Chile 2014–2015; Ecuador, Mexico, and Peru 2017)



Source: Survey of Adults Skills, PIAAC first cycle (OECD).

ARE LEARNING OUTCOMES IMPROVING OVER TIME?

Expanding school access was supposed to translate into increased learning. For most countries in the region, that has not been the case. Despite two decades of near-universal primary enrollment, the evidence tells a sobering story: progress has been limited and uneven, and the gap with OECD countries remains wide.

Despite two decades of near-universal primary enrollment, the evidence tells a sobering story: progress has been limited and uneven, and the gap with OECD countries remains wide.

Have primary school outcomes improved?

The comparison between TERCE 2013 and ERCE 2019 points to a troubling pattern of stagnation. In those six years, most countries recorded only limited changes in the share of students performing below minimum proficiency. In sixth-grade mathematics, the regional average remained virtually unchanged: around 82 percent in 2013 and 83 percent in 2019. Peru shows the clearest improvement, reducing the share of low-performing students from about 76 percent to 61 percent. Brazil and Ecuador also made some progress. By contrast, Argentina shows a deterioration, with its share of low performers rising from about 78 percent to 87 percent. In sixth-grade reading, the overall picture is similarly one of limited change: the regional average moved only from about 69 percent in 2013 to 70 percent in 2019. Even where there were gains, levels of low performance remained extremely high (see Figure A2.1 in the Annex).

How has secondary education performed over time?

■ **Between 2018 and 2022, outcomes worsened across most participating countries, particularly in mathematics.**
In mathematics, the share of low performers increased in most countries, with Costa Rica and Mexico experiencing the sharpest deterioration. In reading, changes were more moderate. Peru, Panama, and the Dominican Republic are the only countries where the share of low-performing students declined. Costa Rica, again, was the only country where low performance increased significantly.

PISA 2022 documented an unprecedented decline in OECD performance in both subjects: average OECD scores fell by about 10 points in reading and nearly 15 points in mathematics, which was three times larger than any previous consecutive change [7]. Latin America and the Caribbean were among the hardest-hit regions in terms of lost instructional time due to COVID-19 school closures (see Figure A2.2 in the Annex).

■ **The long-term picture is mixed, with only a small number of countries making sustained gains.**
Tracking trends from 2006 to 2022, only three countries show positive long-term trends in mathematics average test scores: Peru (+26 points), Colombia (+13), and Brazil (+9). In contrast, Mexico, Uruguay, and Costa Rica display declining trends. In reading, improvements are observed in Peru, Argentina, Colombia, Uruguay, and Brazil, while Costa Rica shows a pronounced decline. In mathematics, five countries show statistically significant increases in the proportion of low-performing students between 2006/2009 and 2022; Peru is the only country

to show a statistically significant improvement, reducing its share of low performers by 7 percentage points (see Figure 2.3 in the Annex).

WHO IS BEING LEFT BEHIND?

Low regional averages mask a deeper problem: foundational learning deficits are not distributed equally. Where a child is born and raised – into which family, community, and type of school it goes to– is a powerful predictor of whether they will learn to read and do basic mathematics. Gaps by socioeconomic status, gender, and territory emerge early in the education cycle and persist through secondary school and into adulthood.

How does family background shape learning?

■ **Socioeconomic status is the most consistent driver of learning inequality across the region.**
Large achievement gaps separate students from the lowest- and highest-income students across all participating countries and at all stages of schooling. In third-grade mathematics, gaps between the bottom and top income quintiles range from 13 to 51%. Guatemala (51%), Colombia (43), and El Salvador (42) have the most pronounced gaps. By sixth grade, achievement gaps in mathematics persist and, in some countries, even widen: Brazil (49 percentage points), Uruguay (42), and Peru (41) have the largest disparities (see Figure 2.4).

At age 15, socioeconomic differences in education continue in all countries. In mathematics, the widest gaps between the top and bottom quintiles are observed in Uruguay (51%), Peru (49), Colombia (45). These gaps consistently favor students in the top quintile, and in most countries that participated in both ERCE and PISA, the socioeconomic gap persists or widens as students advance through school (see Figure A2.4 in the Annex).

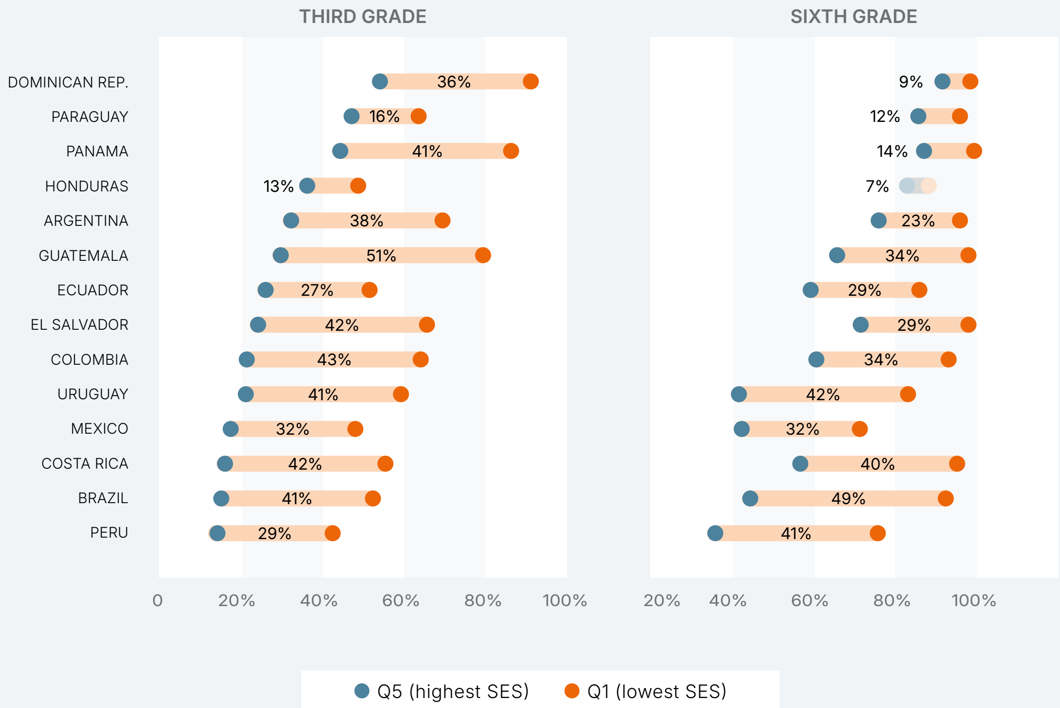
■ **The socioeconomic disparity persists into adulthood.**
In the four participating Latin American countries, a clear pattern emerges: adults whose parents have higher levels of education consistently achieve higher literacy proficiency, suggesting that early-life socioeconomic conditions continue to shape literacy outcomes well into the working life.

Do boys and girls learn differently?

■ **Gender differences are minor in primary school but widen significantly by secondary education.**
In early grades, boys and girls tend to perform at broadly similar levels in mathematics. Most differences are small and not statistically significant. In reading, a more consistent pattern emerges by the end of primary school: girls begin to show a clear advantage over boys across many countries. The largest gaps in favor of girls are observed in Peru (9 percentage points), Dominican Republic (8), Mexico (7) (see A2.5 in the Annex).

■ **By secondary school, a consistent two-directional pattern emerges: girls do better in reading, boys do better in mathematics.**
In mathematics, nine countries exhibit significant gaps in favor of boys. The largest are observed in Costa Rica (9 percentage points), Chile (9), and Peru (8). In reading, girls outperform boys in ten countries, with the gender gap particularly pronounced in Jamaica (16 percentage points), the Dominican Republic (11), and Panama (8) (see Figure A2.6 in the Annex).

FIGURE 2.4: DOES FAMILY INCOME DETERMINE WHO LEARNS?
Gap by socioeconomic status in percentage of students not reaching minimum proficiency in mathematics in primary school, ERCE 2019.



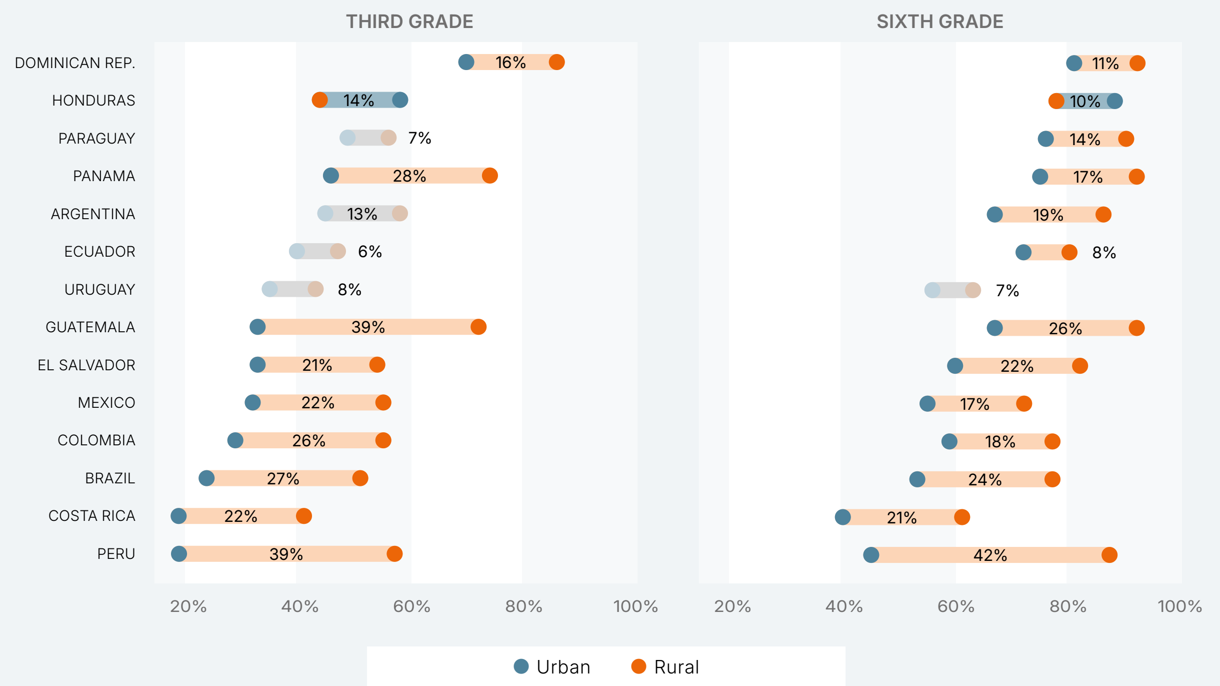
NOTE: Values show the gap in low performance by socioeconomic status, calculated as the low-performance rate among quintile 1 (poorest) students minus the rate among quintile 5 (richest) students. Orange values indicate higher low-performance rates among the poorest students. Values shown in a lighter shade indicate that the difference is not statistically significant.
Source: CIMA (2025) based on ERCE 2019 (LLECE-UNESCO).

Gender differences remain visible in adulthood, but the pattern shifts slightly.
In numeracy, men consistently achieve higher proficiency levels across participating countries. In literacy, gender gaps are smaller and statistically significant only in Chile and Peru (see Figure A2.7 in the Annex).

Do students' geographic background (rural vs. urban) influence their learning outcomes?

Territorial disparities emerge early and remain persistent throughout schooling.
Students living in urban areas consistently achieve better results than those in rural areas from the first years of primary school onward. In reading, by third grade, urban students outperform rural students in nearly all countries, with especially large gaps in Guatemala and Peru (39 percentage points each). By sixth grade, the pattern remains firmly in place, particularly in Peru (42 percentage points), Guatemala (26), and Brazil (24) (see Figure 2.5). These territorial disparities remain substantial in secondary education and exceed those observed in OECD countries by about 6 percentage points in mathematics and 13 percentage points in reading (see Figure A2.8 in the Annex).

FIGURE 2.5: DOES WHERE YOU LIVE DETERMINE WHAT YOU LEARN?
Gap by geographic area in percentage of students not reaching minimum proficiency in reading in primary school, ERCE 2019.



NOTE: Values show the gap in low performance by geographic area, calculated as the low-performance rate among rural students minus the rate among urban students. Orange values indicate higher low-performance rates among rural students; blue values indicate higher low-performance rates among urban students. Values shown in a lighter shade indicate that the difference is not statistically significant.
Source: CIMA (2025) based on ERCE 2019 (LLECE-UNESCO).

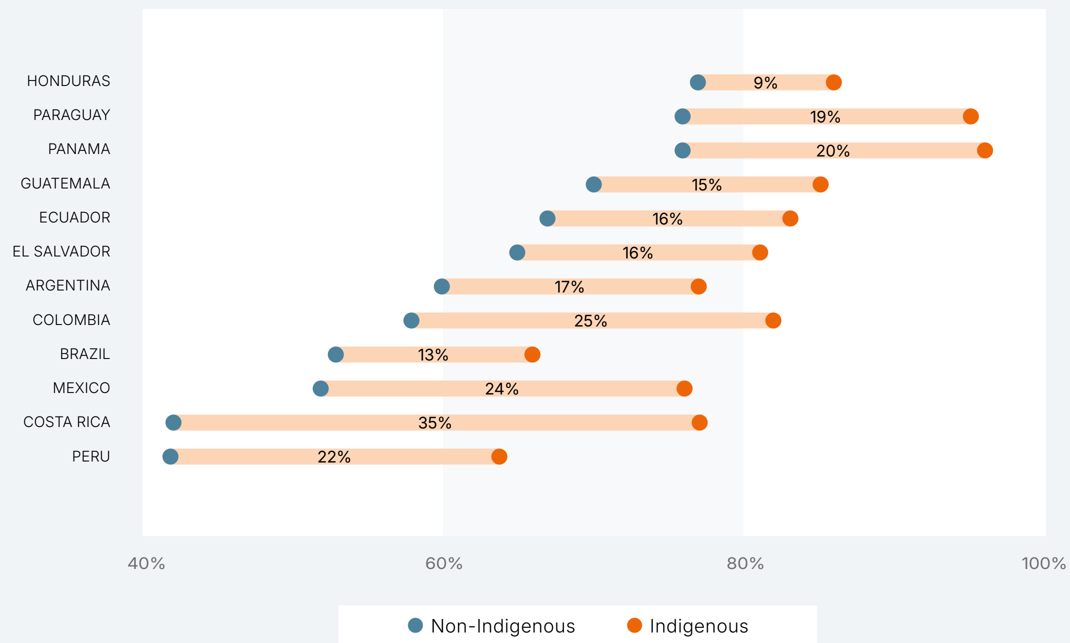
Private vs. Public Schools: does it make a difference?

Private schools consistently outperform public schools by having lower rates of poor performance. But this is fundamentally a story about resources and inequality, not about school type.
Private schools in the region generally offer better learning conditions, such as reliable access to libraries, the Internet, and qualified teachers. In primary education, the largest public-private gaps are observed in Guatemala (52 percentage points), Panama (43), El Salvador (37), and Colombia (24) (see Figure A2.9 in the Annex). These differences persist into secondary education. Roughly nine in ten students in private secondary schools have Internet access at school, compared with about seven in ten in public schools.

What are the learning gaps facing Indigenous students?

Indigenous students face the steepest barriers of all, compounding the effects of geography, poverty, and language.
More than 15.5 million Indigenous children and adolescents live in Latin America and the Caribbean [8]. In sixth-grade mathematics, Indigenous students show significantly higher rates of low performance in all countries. The

FIGURE 2.6: DO INDIGENOUS STUDENTS FACE THE STEEPEST BARRIERS?
Gap by Indigenous status in percentage of students not reaching minimum proficiency in reading in primary school, ERCE 2019



NOTE: Values show the gap in low performance by Indigenous status, calculated as the low-performance rate among Indigenous students minus the rate among non-Indigenous students. Orange values indicate higher low-performance rates among Indigenous students.
Source: CIMA (2025) based on ERCE 2019 (LLECE-UNESCO).

largest disparities are observed in Mexico (19 percentage points), Peru (17), and Costa Rica (14). In reading, gaps are larger than in mathematics in all countries. Costa Rica (35 percentage points), Colombia (25), and Mexico (24) exhibit some of the widest differences in low-performance rates between Indigenous and non-Indigenous students. These students often face the compounding effects of geographic isolation, poverty, and instruction in a language that is not their mother tongue (see Figure 2.6).

WHAT IS AT STAKE IF THE REGION FAILS TO ACT?

The evidence points to an obvious and structural learning crisis in Latin America and the Caribbean. Foundational skills are low, gaps start in early education, and they do not close over time. The challenge is especially severe in mathematics, but is consistent across subjects, levels, and countries: progress in access to schooling has not been matched by progress in learning.

The numbers likely underestimate the problem. The assessments only capture students who remain in school. Many of the most vulnerable children and teenagers have already fallen behind or dropped out. Weak foundational skills do not just affect later learning: they increase the risk of grade repetition, disengagement, and leaving school early. Poor learning is not only a consequence of inequality; it is one of the mechanisms through which inequality is produced and reproduced.

- **The "funnel effect" illustrates how far the region falls short of its own ambitions.**

Large numbers of students enter the system in the early years, but far fewer progress through secondary school with the skills they need. Only around one in five students completes secondary school with at least minimum competencies in reading, while roughly 65 percent finish without those basic skills, and another 13 percent leaves school before completion. Chile is the region's best performer, with 44 percent completing secondary with adequate skills; at the other extreme, only 7 percent do so in the Dominican Republic and Guatemala. These broken trajectories are not a marginal issue: they are a defining feature of the region's human capital challenge (see Figure A2.10 in the Annex).

The available evidence suggests that current approaches to teaching and learning in the early grades have not been sufficient to trigger a significant and broad shift in educational outcomes. Improving foundational learning will require action across the full education cycle, ensuring that primary school reliably builds basic literacy and numeracy, that secondary education retains and supports the most vulnerable students, and that learning opportunities extend into adulthood. This is not only a challenge for education systems but also a broader constraint on productivity, inclusion, and long-term development in the region.

REFERENCES

[1] Busso, M., J. Cristia, D. Hincapié, J. Messina, and L. Ripani (Eds.). Learning Better: Public Policy for Skills Development. 2017. Washington, DC: Inter-American Development Bank.

[2] Arias Ortiz, E., C. Giambruno and S. Karsaclian. The State of Education in Latin America and the Caribbean: A Geospatial Perspective, 2026. Washington, DC: Inter-American Development Bank.

[3] World Bank. The Caribbean's education system: What do declining pass rates reveal? 2025. World Bank Blog.

[4] Thailinger, A., C. Pecha, D. Beuermann, Arias Ortiz, C. Hobbs, and C. Piras. Gender Gaps in the English-Speaking Caribbean: Education, Skills, and Wages, 2023. Washington, DC: Inter-American Development Bank.

[5] OECD. Skills in Latin America: Insights from the Survey of Adult Skills (PIAAC), 2023. OECD Skills Studies. Paris: OECD Publishing.

[6] OECD. Survey of Adult Skills 2023: Chile, 2024. Paris: OECD Publishing.

[7] Arias Ortiz, E., Bos, M., Giambruno, C., and Zoido, P. PISA 2022 in Latin America and the Caribbean: How Much Did the Region Improve? 2023. Washington, DC: Inter-American Development Bank.

[8] UNICEF and FILAC. FILAC and UNICEF Join Forces to Promote the Rights of Indigenous Children in Latin America and the Caribbean, 2023. Panama City: UNICEF Latin America and the Caribbean. Press release.

Part II

How Foundational Skills are Built

THE ANSWER IS NOT IN ANY SINGLE STUDY. IT IS IN WHAT HUNDREDS OF STUDIES, READ TOGETHER, MAKE PLAIN.

CHAPTER

3

I THINK, THEREFORE I LEARN: THE NEUROCOGNITIVE FOUNDATIONS OF LEARNING TO READ AND DO MATHEMATICS

Andrea Paula Goldin

Reading is a culturally invented skill that reorganizes neural systems that had been serving other tasks. Mathematical reasoning, by contrast, combines innate abilities for basic magnitude processing with culturally constructed symbolic systems. Understanding this distinction allows us to reflect on how effective our education systems are (or, in certain cases, how ineffective they are), and also to gauge the sources of the difficulties we commonly observe in teaching and learning these skills.

This chapter presents the existing evidence on how neural, cognitive, and emotional factors contribute to the successful development of mathematical thinking and reading fluency.

The brain is at the core of learning.

If you are reading this book in the order set out in the table of contents, you can surely recall something you learned in the previous chapter. Or perhaps you can recall where you were when you read it, what you had done just before opening the book, or even what you ate, if you are one of those people who eat while they read. Information, in its various forms, is stored in your brain. That is why you can now find it.

The brain is the organ responsible for taking in, retaining, and retrieving knowledge, perceptions, concepts, and emotions.

You have stored not only what you found relevant in what you read; you may also remember what you connected that information to, how you perceived it, or whether some piece of information presented in Chapters 1 or 2 "struck" you emotionally.

In educational neuroscience, we conceptualize learning as brain changes induced by experience across a lifespan. Those changes are constrained and guided by genetics, but they depend on lived experience, and they unfold according to how each person's genetics has interacted with each life history.

A direct corollary is that we can only learn something, consciously or unconsciously, because the human brain contains the cerebral, biological, and physiological structures needed to process that information. It also means that practically all of us have the neural bases to master reading and mathematical reasoning successfully — but coming to handle these tasks requires scaffolding and progressive, explicit, and concrete instruction.

HOW DOES THE BRAIN LEARN?

Our actions arise from patterns of brain activity. A neuron is a cell that transmits information to another through synapses. When their activity becomes synchronized, we speak of neural networks, made up of thousands or millions of interconnected cells. It is from that brain activity that every cognitive, emotional, and behavioral action of an individual emerges.

■ **These networks are dynamic, and the brain reorganizes itself as it develops.**
Synaptic proliferation, strengthening, and pruning are the forces that sculpt neural networks. During a learning process, the three are complementary: proliferation and strengthening arise because the use of some networks has increased relative to others, and pruning makes it possible to discard networks that are underused. In this way, the networks that endure become stable and form a set of connections far more efficient than what existed before [1].

This capacity of the brain is called plasticity, and it underpins academic learning when we want information to be stored in a more organized way, allowing a more efficient later use that, over time, enables more complex activities (see Figure 3.1).

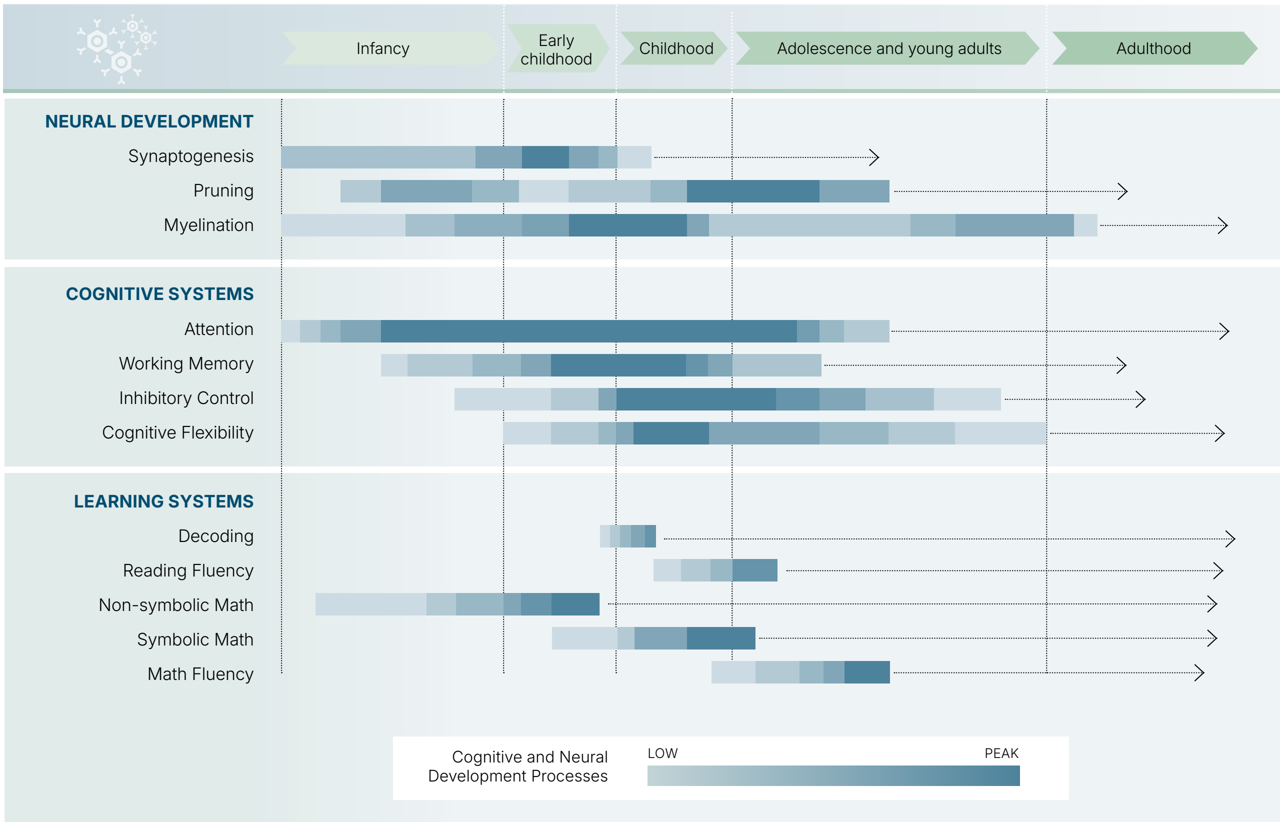
A paradigmatic example of brain reorganization as a product of experience is the acquisition of reading. As we learn to read, a region that recognizes words develops: the visual word form area (VWFA) [2]. This area, in turn, connects with areas that process language and enable its comprehension and production, such as Wernicke's and Broca's areas (see Figure 3.2).

LEARNING TO READ CHANGES OUR BRAIN

People who learn to read as adults undergo brain changes similar to those who learn as children. The VWFA increases its activity in response to written words and is less active in response to spoken words or to strings of letters that do not form a word. Acquiring reading also requires auditory and phonological processing. Once we know how to read and we see written sentences, activity increases in the language-processing networks. And if, in addition, we listen attentively to someone speaking to us, the VWFA is also activated (even with our eyes closed). In other words, no matter the age at which one begins: practically everyone can learn to read.

■ **That said, adults who learned to read as children generally read better than those who learned later.**
But research is still unclear as to whether that advantage is due to a decline in plasticity after childhood or simply to the lesser reading experience of those who learned to read as adults.

FIGURE 3.1: LEARNING SYSTEMS DEVELOP ACROSS THE ENTIRE LIFE COURSE
Neural, cognitive, and learning systems mature on staggered timelines, from infancy into adulthood



Note: Age ranges are defined as follows: infancy (0–2 years), early childhood (3–5 years), childhood (6–11 years), adolescence and young adults (12–25 years), and adulthood (26 years and above). These ranges are indicative and may vary by source.
Source: Developed by the author.

Is there a brain area for mathematics?

Less than a decade ago, what was called the number form area was identified in the inferior temporal cortex [3], but studies are still inconclusive [4]. What is clear is that from a very early age we possess a set of language-independent neurobiological abilities that, among other things, allow a degree of perception and discrimination of quantities [5]. The literature debates whether this is an innate basic capacity (which some call “number sense”) or whether what matters is how the neural, physiological basis is rapidly shaped by experience [6, 7]. Both positions share a point that is crucial for education and grounded on neural plasticity: it is essential to give people appropriate opportunities to develop these abilities.

Our nervous system is prepared to understand and estimate numerical quantities and to operate with them, but it needs exercise.

With appropriate demands, sustained practice, and the necessary support, we gradually manage to cope with and even master symbolic mathematics. As a person (whether in childhood, adolescence, or adulthood) exercises the understanding of quantities and magnitudes, they will come to operate more effectively and/or efficiently with them, which will translate into improved mathematical performance [8].

Are there foundational skills shared by reading and mathematics?

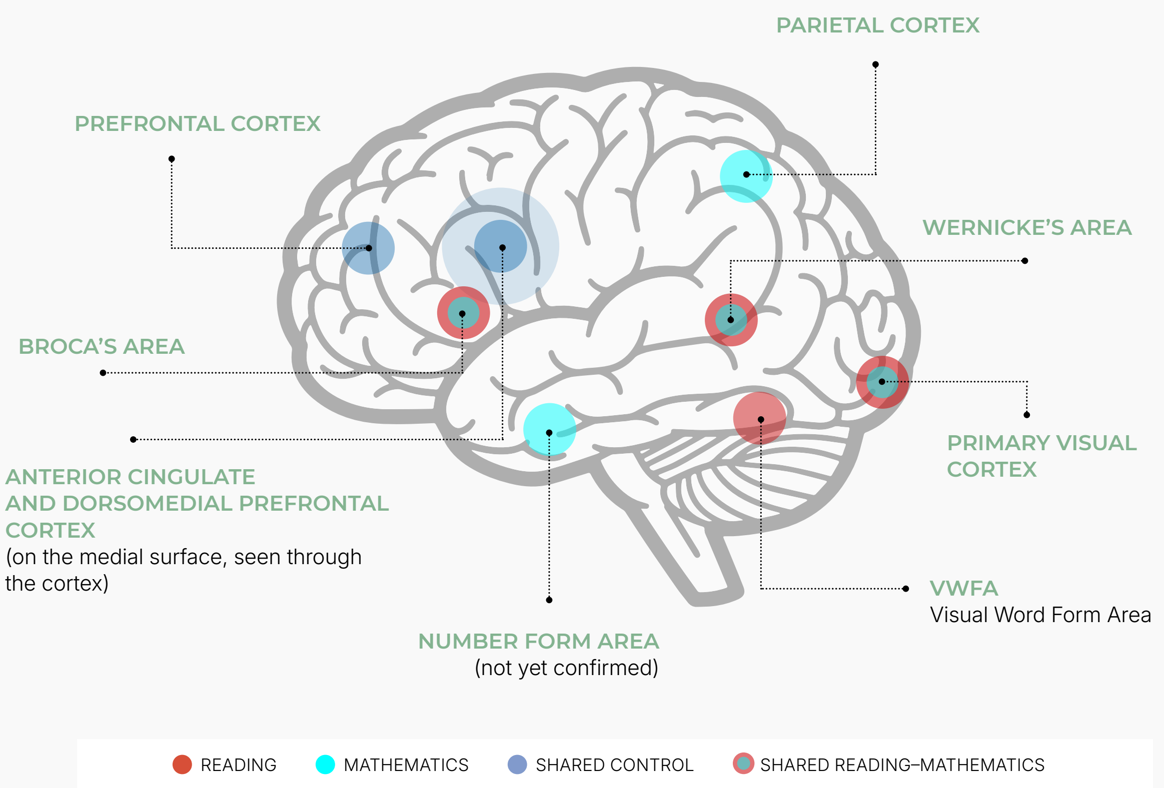
In learning and grasping ideas, certain cognitive and emotional skills act as mediators of academic success. They are foundational because they include the preconditions that enable more advanced cognitive capacities, such as a deep command of working with numbers and of comprehending and processing what is read. They are known as *executive and attentional functions*, or *self-regulation*. Their appropriate development allows us to maintain focus, retain information in memory, control actions and thoughts, and adapt to changing environments. They cut across all school subjects, but they are usually not obvious. Throughout life, the frontal cortex is specifically related to these skills.

In particular, the anterior cingulate cortex is crucial for cognitive control, monitoring one's own performance, and allocating mental effort.

This cortex, together with the dorsomedial prefrontal cortex, acts as a major controlling entity, influencing how much cognitive effort is devoted to each learning task. This motivational control could explain why focusing on mathematics or reading demands sustained mental effort the more complex or novel a task is for us, and why academic performance may depend on inter-individual differences in the willingness to devote cognitive effort to a task.

FIGURE 3.2: THE BRAIN REGIONS INVOLVED IN READING AND MATHEMATICS

Reading, mathematics, and the control networks shared by both are supported by distinct but overlapping regions



Source: Developed by the author.

Why are foundational skills a prerequisite for automatization?

The development of numerical skills (understanding the meaning of numbers as a representation of quantities and magnitudes) and arithmetic skills (understanding conceptually what an operation represents, and how —and how well— it is carried out) is much more than the product of a single brain region or a single cognitive mechanism. As in the case of reading, its development is complex and multidimensional [9]. Efficient and flexible processing of these capacities, in both children and adults, depends on communication between the brain networks specialized for numerical processing and reading and the frontal cortex.

■ **Both children and adults use their executive and attentional networks for tasks they have not yet automatized.**

The strategies children use require a great deal of cognitive control, but that need decreases as they become more efficient. Something similar happens with adults when they are not experts at a given task. The changes do not appear to be due to development per se, but rather to specific practical experience and to how difficult each task is for each person, regardless of age.

HOW DOES THE BRAIN LEARN TO READ?

Reading is a process that almost always begins in the eyes and is carried out with the brain. It requires understanding that certain visual symbols (letters) represent language, which we grasp when we decode those symbols and put them together to form words. Ideally, reading requires an understanding of what is read: being able to extract the ideas from the text and relate, contrast, and combine them with prior ideas. In this way, a text can challenge us, anger us, or give us goosebumps. As you read this book, you may find yourself doubting, emboldened, or surprised by something written here. You can do so only because you already know how to read well. To reach your level of reading expertise, you spent years exhaustively and progressively exercising all the skills mentioned at the beginning of this paragraph.

How do we move from decoding to fluency?

■ **An illiterate person must first learn that each letter (grapheme) has an associated sound (phoneme), and remember which sound is represented by which letter.**

They must learn to identify them unambiguously. As they manage this, they can detect the sounds in each word and begin to put sounds together to produce words. First simple ones — "mama" or "ball" — and then more complex ones (*"supercalifragilisticexpialidocious"*). Through phonological decoding, we can read any word, including those whose meaning we do not know or that we barely use. But it is a slow process. Learning to read fluently requires understanding written text as well as we process spoken language, which is what makes it possible to devote most of our processing capacity to reading comprehension.

■ **This may lead us to think, mistakenly, that the best approach is to teach reading directly with whole words.**

After all, why else would we have a visual word form area dedicated to that? The reasoning seems very logical, but it conceals a serious problem: the VWFA is a structure that we "recycle"; before we learned to read, it was busy with other things [2]. Transforming it and "filling" it with words takes a great deal of time and practice, because it requires rewiring and major modifications to the neural networks.

As we advance in reading practice and manage to decode the same written word many times over, we establish the connection between its "form," its sound, and —if we know it— its meaning. In this way we build a kind of *supply of written words*, thanks to which the next time we see a written word whose form we have already stored, we will be able to recognize it quickly, along with its meaning.

How to achieve good reading comprehension?

Decoding makes it possible to read any word. But once that has been achieved, attaining reading comprehension requires being able to read fluently and with very little cognitive effort, almost automatically. That way we can devote our attentional resources to understanding what the text says. Knowing in advance the words we are going to read and, ideally, understanding their meaning has an amplifying effect: it favors fluency, and it positively affects interest and motivation (nobody wants to read a text they do not understand or are not interested in) (see Figure 3.3).

■ **One key factor in achieving effective and efficient early reading performance is to foster vocabulary knowledge during childhood.**

This can be done through shared reading and storytelling, but also by talking to babies and young children, allowing them to interact with different adults, and introducing them to vocabulary and content relevant to their context. Without prior knowledge, comprehension becomes superficial; it is harder to relate what is read to other ideas or concepts, and the resulting learning is short-lived and, above all, deadly boring.

Box 3.1: Are all languages learned in the same way?

Latin America has a great advantage. Grapheme–phoneme correspondences are practically one-to-one in Spanish and in Portuguese, which is not the case in other languages, such as English (which we describe as opaque because it lacks the consistency that transparent orthographies have).

What do we know about dyslexia?

Dyslexia is a persistent difficulty in processing written words that is not due to poor teaching or to a lack of intelligence or motivation. Nor is it a general comprehension problem, because when the information is auditory, comprehension is intact [10]. The prevailing view is that the difficulties stem from an atypical phonological representation of the form of spoken words, which can impede phonological awareness (the ability to recognize and manipulate the individual sounds that make up a word).

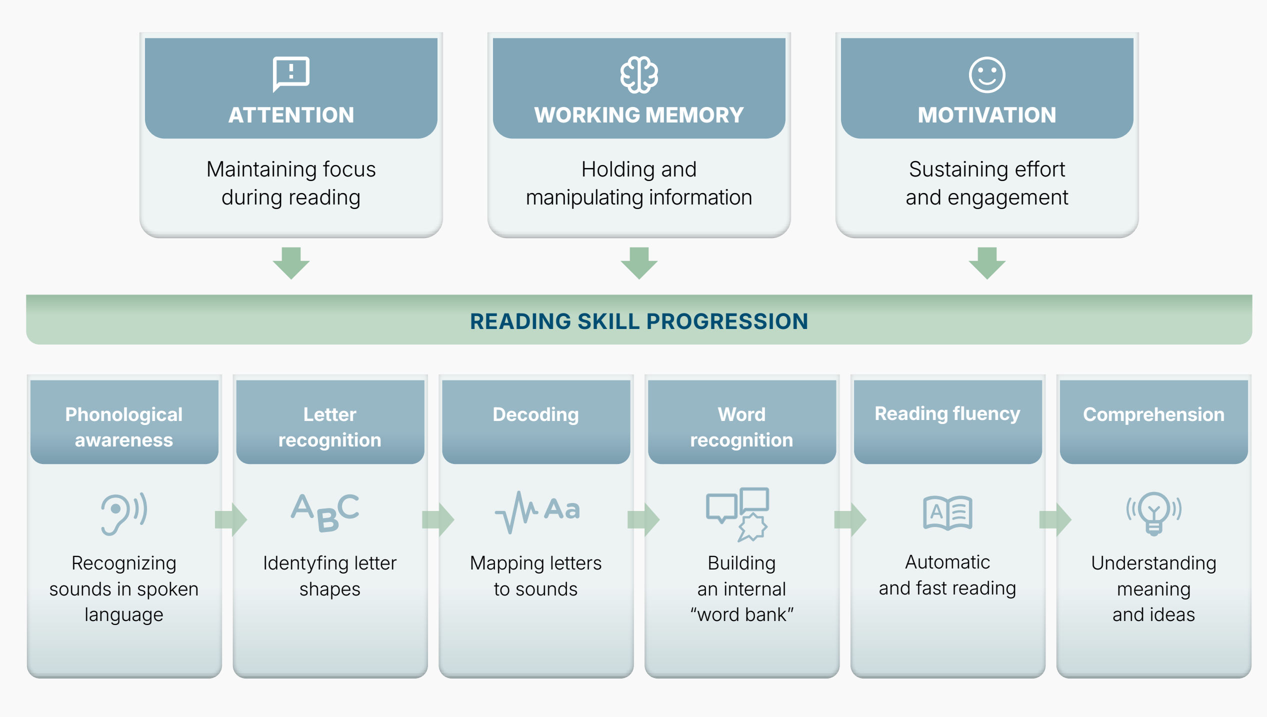
■ **Fostering phonological awareness is a widely used strategy for improving, at the behavioral level, the perception of rhythm in language, reading, and spelling.**

The empirical evidence indicates that the reading functioning of people with dyslexia can improve enormously with appropriate cognitive therapies [11], although there is still a big expert debate about these strategies as a way of generating protective neural mechanisms.

HOW DOES THE BRAIN LEARN MATHEMATICS?

The primitive brain mechanisms that allow estimation, comparison, and ordering of basic quantities are plastically modified by the teaching of symbolic number systems and formal mathematics. Learning mathematics in the first years of primary school depends in part on the number skills children arrive with, and those initial differences tend

FIGURE 3.3: READING IS BUILT IN STAGES, FROM SOUNDS TO MEANING
Progression runs from phonological awareness to fluent comprehension, supported throughout by attention, working memory, and motivation



to persist throughout their schooling [12] [13]. There is no age limit for improving one's command of numerical magnitudes. But attaining it requires appropriate, progressive, and systematic practice, with motivation and feedback (see Figure 3.4).

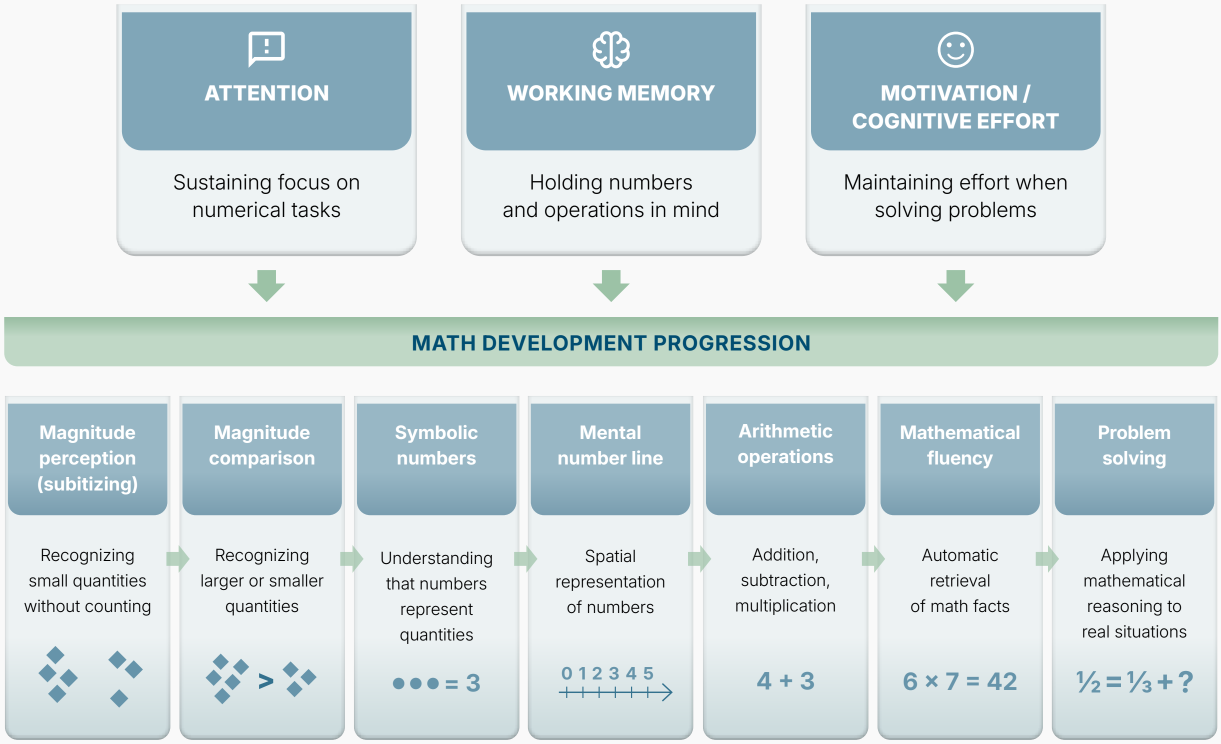
Four things are four things. But the number 4 is arbitrary; the way it is written could have been anything else. The unambiguous relationship between the formal symbol and the quantity can be acquired very early through playful activities with numerical content, such as dice or board games in which one has to count squares, compare sets of objects, or recognize shapes or quantities visually and/or by touch and name them.

Comparing and ordering quantities facilitates understanding the concept of linearity on which formal mathematics is based.

Linearity implies that, for example, the distance between 2 and 3 is the same as the distance between 5002 and 5003 (even though, proportionally, the first difference is highly relevant and the second insignificant). The internal spatial representation organized in number line format is a dynamic structure that begins with small magnitudes and that we progressively expand as we come to (re)cognize larger numbers and, later, negative numbers, fractions, and decimals. Representing numbers mentally in a linear way is associated with proficiency in formal mathematics and predicts academic performance.

Organizing numbers on a line requires spatial reasoning, and the development of this ability can also be encouraged early on through tasks that call for orienting oneself or moving in space —whether real or imagined— solving puzzles, building with blocks, folding origami, running, climbing, or hiding.

FIGURE 3.4: MATHEMATICS IS BUILT FROM INTUITIVE QUANTITY TO SYMBOLIC REASONING
Progression runs from magnitude perception to problem solving, supported throughout by attention, working memory, and effort



These examples of playful activities are not a suggestion to begin schooling before primary school; rather, they are about exposing children early on to interacting in every possible way with quantities, magnitudes, and their order, so that they connect non-symbolic with symbolic representations, take on the principle of linearity, and become familiar with the appropriate vocabulary.

How is mathematical fluency built?

At the cognitive level, solving problems through formal representations requires knowing what each symbol means and how it is handled. But, above all, it requires a great deal of practice. Plastic changes in the brain need time and exercise, which gradually stabilize the underlying neural networks until, ideally, a "eureka moment" is reached: we have understood how a problem is solved.

Just as building a store of words gives us rapid access to their meaning and lets us understand even long, complex sentences as we read, there is also fluency in mathematics [14]. Having the results of frequent operations stored in memory (such as the multiplication tables, or how much 4 + 3 is) allows us to reduce cognitive load when solving a problem, so that we can devote our attention and executive functions to thinking about the big picture rather than getting jammed on small calculations.

The only way to achieve fluency is through exercising progressively more difficult problems, neither too easy nor too frustrating. The starting point and the pace of progression will depend on each person, and on how much they have become familiar with the numbers and their use. Broadly speaking, it is necessary to begin with small numbers and simple operations and to make tasks progressively more complex, to gain a natural command of numbers and their operations and to generate the fluency needed to understand what lies behind the problem to be solved.

WHAT DOES ALL THIS IMPLY FOR TEACHING?

■ **Learning to read means recycling neural systems that evolved "for something else" and using them to understand a written text.**
But the invention of reading is a recent development in human history (some six thousand years ago), and there has not been enough evolutionary time for us to develop circuits, regions, or genes specific to this task. The consequence is that reading does not come to us naturally through mere exposure to a text (as it does with spoken language or with face recognition, for example): learning to read requires explicit instruction.

■ **Mathematical learning, by contrast, does build on an innate component.**
Although the literature debates the extent to which we are born with it, there are biological bases, which we share with other animals, that allow us to detect, discriminate, and operate with certain quantities. Layered on top of these capacities are invented components that require explicit instruction, such as symbolic number systems (we use the Arabic base-10 system, but others exist) and formal mathematics (performing algebraic calculations, geometry, notation).

Understanding these distinctions allows us not only to reflect on the effectiveness, or ineffectiveness, of our education systems, but also to gauge the sources of learning difficulties. Most do not arise from students' conditions or deficits; rather, they reflect inconsistencies in how instructions and materials are delivered relative to the stage of development each person is at. These can range from difficulty in mapping quantities and operations onto their symbols, or letters onto their sounds, to excessive cognitive, attentional, or emotional load. This is why instruction adjusted to each student's level of cognitive development is key, and why the teacher is an indispensable facilitator for the brain changes that underpin this learning.

■ **It is never too late to learn.**
The brain is modified by experience. The changes we can expect over time do not depend exclusively on genetics but, above all, on how genetics interacts with what happens to a person. This dynamic integration takes place throughout our development. We now know that, although the first two decades of life include sensitive periods of plastic reorganization in which it is "easier" to modify the brain and to learn, learning can be achieved across the entire life cycle —though usually with greater effort later on [15].

REFERENCES

[1] Lindenberger, U. and M. Lövdén, Brain plasticity in human lifespan development: the exploration–selection–refinement model. Annual Review of Developmental Psychology, 2019. 1(1): p. 197-222.

[2] Dehaene, S., L. Cohen, J. Morais, and R. Kolinsky, Illiterate to literate: behavioral and cerebral changes induced by reading acquisition. Nature Reviews Neuroscience, 2015. 16(4): p. 234-244.

[3] Yeo, D.J., E.D. Wilkey, and G.R. Price, The search for the number form area: A functional neuroimaging meta-analysis. Neuroscience & Biobehavioral Reviews, 2017. 78: p. 145-160.

[4] Merkley, R., B. Conrad, G. Price, and D. Ansari, Investigating the visual number form area: a replication study. Royal Society Open Science, 2019. 6(10): p. 182067.

[5] Izard, V., C. Sann, E.S. Spelke, and A. Streri, Newborn infants perceive abstract numbers. Proceedings of the National Academy of Sciences, 2009. 106(25): p. 10382-10385.

[6] Whitacre, I., B. Henning, and Ş. Atabaş, Disentangling the research literature on number sense: Three constructs, one name. Review of Educational Research, 2020. 90(1): p. 95-134.

[7] Andrews, P. and J. Sayers, Identifying opportunities for grade one children to acquire foundational number sense: Developing a framework for cross-cultural classroom analyses. Early Childhood Education Journal, 2015. 43(4).

[8] Siegler, R.S. and D.W. Braithwaite, Numerical development. Annual Review of Psychology, 2017. 68(1): p. 187-213.

[9] Vogel, S.E. and B. De Smedt, Developmental brain dynamics of numerical and arithmetic abilities. Science of Learning, 2021. 6(1): p. 22.

[10] Peterson, R.L., and B.F. Pennington. Developmental Dyslexia. 2015. Annual Review of Clinical Psychology, 11(1): 283–307.

[11] Goswami, U. Sensory Theories of Developmental Dyslexia: Three Challenges for Research. 2015. Nature Reviews Neuroscience, 16(1): 43–54.

[12] Jordan, N.C., D. Kaplan, C. Ramineni, and M.N. Locuniak. Early Math Matters: Kindergarten Number Competence and Later Mathematics Outcomes. 2009. Developmental Psychology, 45(3): 850.

[13] Silver, A.M., L. Elliott, A.D. Ribner, and M.E. Libertus. The Benefits of Math Activities Depend on the Skills Children Bring to the Table. 2024. Developmental Psychology, 60(2): 376.

[14] McNeil, N.M., N.C. Jordan, A.A. Viegut, and D. Ansari. What the Science of Learning Teaches Us About Arithmetic Fluency. 2025. Psychological Science in the Public Interest, 26(1): 10–57.

[15] Lipina, S.J., and M.S. Segretin. 6000 días más: evidencia neurocientífica acerca del impacto de la pobreza infantil [6,000 more days: neuroscientific evidence on the impact of child poverty]. 2015. Psicología Educativa, 21(2): 107–116.

CHAPTER 4

1+1=2. WHY? FOUNDATIONAL SKILLS FOR MATHEMATICAL LEARNING: WHAT WORKS, WHY IT WORKS, AND ESSENTIAL TEACHING GUIDELINES

Emma Näslund-Hadley

Mathematics can be taught in many ways, but not all approaches lead to the same learning outcomes. How learning is delivered, how it is organized, and how much guidance instruction provides are key elements. The question is which approaches make the greatest difference for learning.

Do students develop different mathematical skills when mathematics is taught within the classroom as a stand-alone subject than when it is reinforced through activities beyond it? What role does instructional structure play in shaping learning outcomes? Are some pedagogical approaches more effective than others for building math skills? How does the mode of delivery, whether face-to-face, virtual, or hybrid, shape learning outcomes? How does the length of an instructional lesson impact math learning outcomes?

Looking across studies makes it possible to identify consistent patterns in what works.
Answering these questions based on individual studies alone is difficult, as each study provides only a partial view of what works in one specific context. However, by synthesizing evidence across studies, it becomes possible to identify patterns that are not always visible in individual analyses [1].

This chapter draws on a systematic synthesis of 102 studies evaluating the impact of mathematics programs, policies, and teaching approaches, including 91 in pre-primary and primary education, and 11 in secondary education [2]. To synthesize the evidence and enable comparisons across studies, effect sizes are converted into standard deviations, allowing results from different programs, policies, and teaching approaches to be expressed on a common scale. A useful way to interpret these effect sizes is in terms of instructional time. As a rough benchmark, an effect of 0.40 standard deviations is approximately equivalent to about one additional year of learning, or roughly 35 to 40 weeks of instruction, while an effect of 0.10 corresponds to about 9 to 10 weeks [3]. Findings from secondary education are reported whenever the evidence-base permits.

The studies included in the synthesis vary not only in their findings, but also in the strength of the evidence they provide. All studies included in the review use a treatment and comparison group, allowing for causal estimation of program impacts. Of the 102 studies included, 58 are classified as high rigor and 44 as moderate rigor. High-rigor studies are primarily large randomized controlled trials (RCTs) with strong implementation quality, including clear and credible randomization procedures, balanced baseline characteristics between groups, limited risks of spillovers or contamination, and high levels of compliance and treatment fidelity. Moderate-rigor studies include smaller RCTs and quasi-experimental designs, which are generally more vulnerable to threats to internal validity and rely on stronger assumptions to attribute observed outcomes to the intervention. No low-rigor studies were included in the synthesis. For a complete description of the methodology, see online Methodological Annex.

Understanding what works is especially critical in contexts where learning gaps are largest.
Of the 102 studies included in this chapter, 60 were conducted in North America, 17 in Europe, and 19 in low- and middle-income countries, including 9 from Latin America and the Caribbean. The remaining studies were carried out in other high-income settings. Thus, roughly one in five studies are drawn from low- and middle-income contexts, providing important insights for identifying approaches that are relevant beyond high-income settings (see Figure A4.1 in the Annex). Even so, increases of 0.20 standard deviations are substantial, as they are equivalent to approximately 20 additional weeks of learning. In causal studies of education interventions that measure academic achievement, effects below 0.05 standard deviations are considered small, effects from 0.05 to less than 0.20 moderate, and effects of 0.20 or greater large [3].

Far from diminishing the relevance of the evidence, these findings make it especially important to understand which approaches are most feasible and effective in contexts where the need for stronger foundational skills is often greatest. For policymakers in low- and middle-income countries, the findings in this chapter are therefore highly relevant: they point not only to what works, but to the kinds of instructional approaches that may hold the greatest promise in practice.

WHERE IS MATHEMATICS LEARNING EFFECTIVE?

Mathematics is most often taught as a stand-alone subject, in its own class period and within clear disciplinary boundaries. Yet many education systems are increasingly looking beyond the traditional classroom to strengthen mathematical learning. Students may encounter math when it is woven into other subjects, explored through extracurricular activities, or reinforced through expanded school-day models. These learning approaches do not replace dedicated mathematics instruction, but complement it by broadening the settings, time, and contexts in which students encounter and use numerical skills.

The biggest gains happen in the math classroom, with expanded learning adding valuable support.
A clear pattern emerges when examining the evidence from the 91 studies on pre-primary and primary math education. Mathematics taught as an independent subject produces stronger learning gains overall than delivered through expanded mathematics learning approaches. On average, sessions that taught mathematics as a stand-alone subject generated a gain of 0.42 standard deviations, compared with 0.27 standard deviations for expanded

> How mathematics is taught not only matters for how much students learn, but also for what kind of learning they develop. Stand-alone mathematics instruction appears especially important for building procedural fluency and skills, **while expanded mathematics learning may be valuable for helping students make sense of ideas, connect the subject to broader experiences, and deepen their conceptual understanding.**

mathematics learning in pre-primary and primary education. Taken together, these findings suggest that expanded approaches cannot replace strong classroom instruction. Nevertheless, by providing additional time, settings, and opportunities for students to engage with mathematics, it can be a valuable complement. This pattern appears in the evidence from all countries, regardless of their income levels.

The way mathematics is taught shapes the kind of thinking students develop.

The most interesting differences emerge, however, when we look not only at how much students learn, but at what kinds of learning each approach seems to support best. In math education, two of the most important outcomes are procedural and conceptual understanding. A child with procedural understanding can carry out operations and apply rules accurately, for example, by memorizing a formula and correctly substituting numbers into it. A child with conceptual understanding, in contrast, grasps the underlying principles behind that same formula, understanding why it works and how it can be applied across different contexts [4].

Practice makes proficiency, and that happens in the classroom.

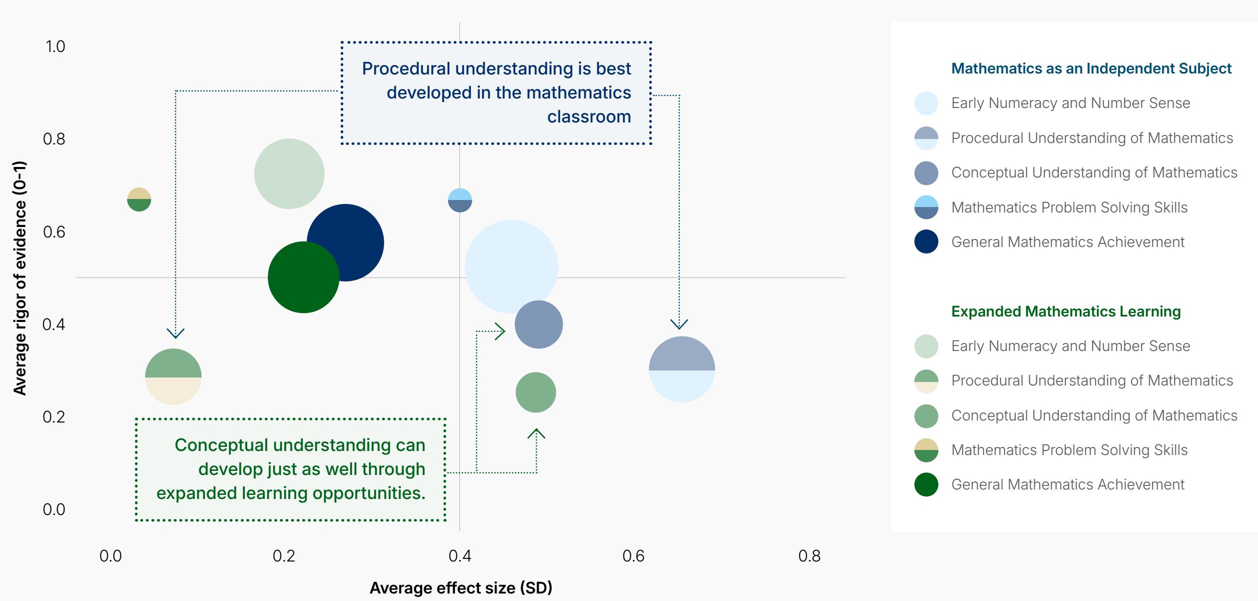
Mathematics taught as an independent subject appears especially important for building procedural understanding, where gains are far larger (0.65 standard deviations, compared with 0.07 in expanded settings). This likely reflects the fact that procedural learning depends heavily on repetition, sequencing, and sustained practice, features that are more commonly built into regular math instruction than into extracurricular or cross-curricular activities.

Conceptual understanding can be built both inside and beyond the classroom.

Expanded mathematics learning appears to be just as effective as classroom-based instruction in supporting conceptual understanding of why mathematical relationships and procedures work, not just how to apply them. In other words, students can develop an understanding of the meaning behind a formula, not only learn to crunch the numbers correctly [4]. Both approaches produce effects of 0.49 standard deviations. This suggests that conceptual understanding can be developed effectively across different learning settings. Because this type of understanding often takes time to develop, opportunities both inside and beyond the classroom may give students the time and space to explore, discuss, and apply mathematical ideas, helping them make connections and deepen their understanding of what those ideas mean.

These findings suggest that how mathematics is taught not only matters for how much students learn, but also for what kind of learning they develop. Stand-alone mathematics instruction appears especially important for building procedural fluency and skills, while expanded mathematics learning opportunities may be particularly valuable for helping students make sense of ideas, connect the subject to broader experiences, and deepen conceptual understanding. Their greatest value, therefore, may lie not in replacing classroom instruction, but in enriching and extending it (see Figure 4.1).

FIGURE 4.1: TRADITIONAL MATH INSTRUCTION VS EXPANDED TEACHING METHODS: WHAT WORKS BETTER?
Classroom instruction drives procedural understanding in pre-primary and primary education, while expanded learning supports conceptual understanding of mathematics



Note for Figures 4.1–4.5: Effect sizes are reported in standard deviations (SD). Following Kraft (2020), effects below 0.05 SD are considered small, effects between 0.05 and 0.19 SD are moderate, and effects of 0.20 SD or larger are large. Rigor refers to the methodological strength of the causal identification strategy. Higher rigor studies include large randomized controlled trials and stronger quasi experimental designs. No low rigor studies without a credible comparison group or causal identification strategy are included in the review. The diameter of each circle corresponds to the number of studies included in each category of the intervention characteristic. Source: Näslund-Hadley et al 2026.

DOES HOW WE DELIVER MATH CHANGE WHAT STUDENTS LEARN?

The use of virtual and hybrid learning has expanded rapidly in recent years, particularly following the COVID-19 pandemic, reshaping how education systems deliver instruction. Across Latin America and the Caribbean, there is a growing emphasis on integrating digital tools and hybrid models into regular teaching, both to increase flexibility and to expand access. In this chapter, virtual approaches refer to programs and teaching approaches delivered entirely through digital technologies, while hybrid approaches combine face-to-face instruction with online learning. The effectiveness of these approaches depends not only on access to technology, but on how well teachers are prepared to design and deliver instruction in hybrid environments [5]. The evidence on virtual and hybrid approaches is somewhat less robust than the evidence on face-to-face instruction, as it is based more often on smaller studies and less rigorous evaluation designs.

How we deliver math changes less how much students learn, and more what they learn.

The studies reviewed compare specific programs and teaching approaches delivered through different modalities, rather than entire education systems operating face-to-face or online. In pre-primary and primary education, face-to-face instruction generates gains of around 0.41 standard deviations, compared to 0.32 standard deviations for fully virtual and hybrid models. A similar pattern appears at the secondary level, where face-to-face instruction is associated with gains of 0.30 standard deviations, compared to 0.21 standard deviations for virtual and hybrid approaches. All of these effects are educationally meaningful, while the differences between them are comparatively modest [3]. Evidence from low- and middle-income countries remains limited, with only 19 studies examining delivery modality, but the pattern points in the same direction as in higher-income settings. Much larger differences emerge when looking at specific learning outcomes.

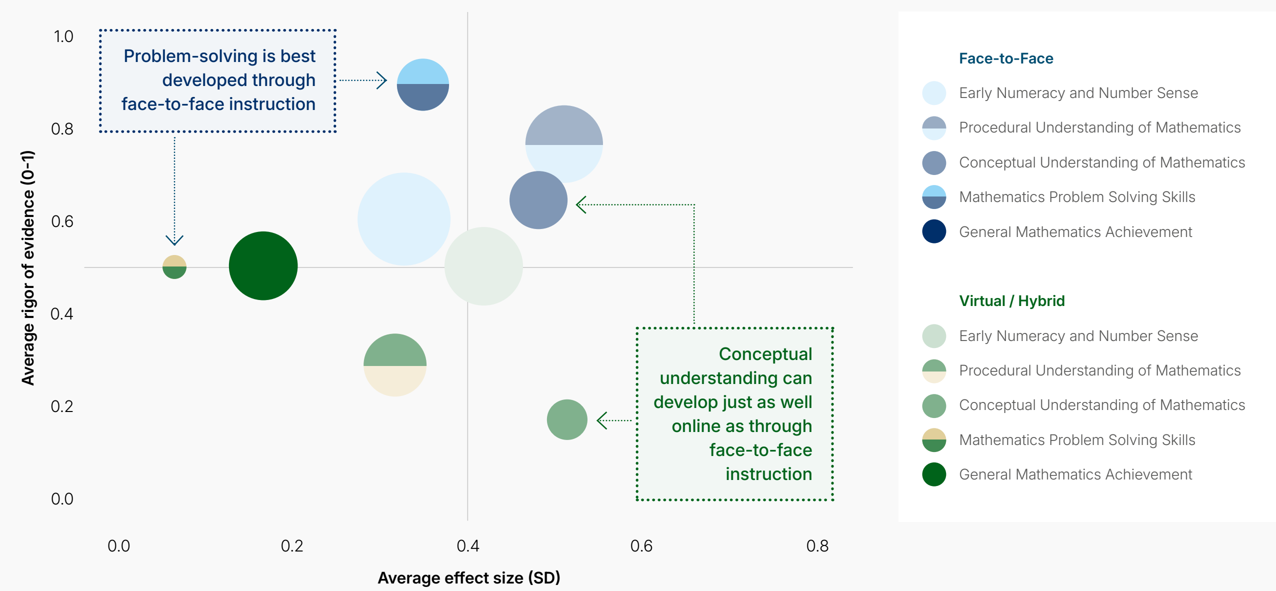
- **Problem solving thrives on interaction, which face-to-face instruction provides best.**

Face-to-face delivery appears to be particularly important for building mathematics problem-solving skills, where effect sizes reach 0.35 standard deviations, compared to just 0.06 for virtual and hybrid approaches combined at the preprimary and primary levels of education. This may reflect the fact that higher-order mathematics learning, such as problem solving, depends on sustained opportunities for teachers to observe student thinking, provide feedback, and address misconceptions in real time. It is also associated with stronger gains in general mathematical achievement (0.33 versus 0.17 standard deviations), suggesting that face-to-face instruction may be particularly well suited to supporting the combined set of skills that broader mathematics assessments tend to capture.

- **Conceptual understanding can develop just as well online as through face-to-face instruction.**

Students appear to build a similarly deep understanding of mathematical ideas whether instruction is delivered face-to-face or through virtual and hybrid approaches, with effect sizes of 0.49 and 0.51 standard deviations, respectively. This suggests that when digital and blended environments are well designed, they can provide students with meaningful opportunities to explore mathematical relationships, reflect on ideas, and build a deeper understanding of why mathematical procedures work (see Figure 4.2).

FIGURE 4.2: IN-CLASS TEACHING OR ONLINE/HYBRID MATH INSTRUCTION: WHAT WORKS BETTER?
In pre-primary and primary education, delivery modes are not interchangeable, they shape different learning outcomes



Source: Näslund-Hadley et al 2026.

Different delivery modes are not interchangeable, they shape different kinds of learning.
The implication of the overall analysis of these dozens of studies is not that face-to-face instruction is always superior, but that different delivery modes may be better suited depending on different learning goals. For educational decision-makers, this suggests using face-to-face instruction strategically for the kinds of mathematics learning that depend most on discussion, feedback, and close attention to student thinking, while ensuring that virtual and hybrid environments are designed to support these processes as effectively as possible.

How mathematics lessons are organized has a major impact on learning. Lessons with clear goals, carefully sequenced explanations, guided practice, and continuous feedback produce substantially stronger gains than less structured approaches.

HOW SHOULD MATHEMATICS LESSONS BE ORGANIZED TO SUPPORT LEARNING?

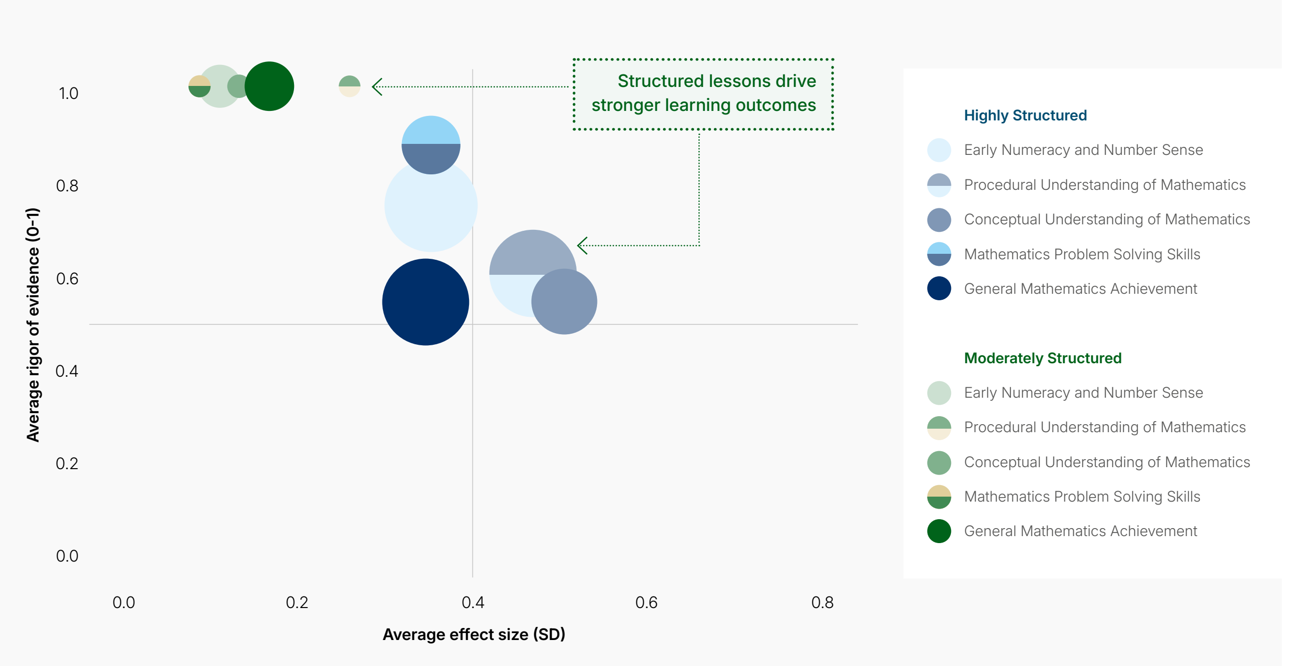
A useful way to think about instruction is in terms of how learning is structured in the classroom. Some lessons are highly structured. Advocates of this type of teaching argue that learning is most effective when it is carefully structured, setting clear goals, introducing content in carefully sequenced steps, and providing guided practice with continuous feedback that supports students as they build understanding [6]. Other approaches rely on less structured lessons, in which students are given more responsibility for exploring ideas, organizing information, and making sense of problems, with guidance provided more selectively or at later stages [7].

How lessons are organized has a major impact on how much students learn.
The synthesis of evidence presented in this chapter provides a clear and consistent message: structure matters for learning mathematics. Across educational levels, highly structured lessons are associated with substantially stronger learning gains than moderately structured approaches. In pre-primary and primary education, they generate gains of around 0.39 standard deviations, compared to 0.14 standard deviations for moderately structured lessons. A nearly identical pattern emerges at the secondary level, where highly structured lessons are associated with gains of 0.43 standard deviations, compared to 0.09 standard deviations for moderately structured lessons. The studies in both groups are similar in methodological rigor, suggesting that the larger effects associated with highly structured lessons are unlikely to reflect differences in study quality (see Figure 4.3).

Highly structured teaching improves mathematics learning across countries at all income levels.
While only nine studies from low- and middle-income countries analyze the level of instructional structure in mathematics education, the direction of the findings is the same across countries of different income levels. A well-known example from a low-income setting is the Primary Math and Reading (PRIMR) initiative in Kenya [8]. The program provided teachers of first and second graders with highly structured lesson guides built around sequenced steps, independent work, and continuous formative feedback, and produced gains of 0.37 standard deviations in procedural understanding and 0.33 standard deviations in conceptual understanding among Grade 2 students.

In Latin America, classroom structure often supports order, not learning.
The strong effectiveness of structured lessons across education levels suggests that the issue is not whether structure matters, but whether it is sufficiently put into practice. Teachers in Latin America self-report frequent

FIGURE 4.3: DO STUDENTS LEARN MORE IN HIGHLY STRUCTURED OR MODERATELY STRUCTURED CLASSES?
More structured lessons lead to substantially stronger mathematics learning in pre-primary and primary education



Source: Näslund-Hadley et al 2026.

use of core elements of structured mathematics instruction, but these reports should be interpreted with caution, especially when they are not corroborated by direct classroom observation. Since its first cycle in 2008, the Teaching and Learning International Survey (TALIS), the world's largest international survey of teachers and principals conducted by the Organization for Economic Cooperation and Development (OECD), has shown high reported use of practices such as stating lesson objectives and checking for understanding, with roughly 85 to 90 percent of teachers in the Latin American region reporting frequent use [9] [10].

One possible explanation for these high self-reported rates is that individuals with more limited knowledge or skill in a given area may overestimate their own performance [11]. Social desirability or differences in how teachers interpret these practices may also contribute. These self-reported data should be interpreted cautiously, however, as the available TALIS evidence for primary education covers only a limited number of Latin American education systems. Yet classroom videos and observation studies from the region point to a more sobering conclusion: the structure visible in many Latin American classrooms often seems to be aimed more at maintaining order and routines than at supporting the kind of sequenced explanation, guided practice, and sustained engagement that mathematical learning requires [12] [13].

Strong instruction combines clear objectives, step-by-step explanations, and continuous feedback. The implication is not that classrooms need more order for their own sake, but that they need stronger instructional structure. In mathematics, effective structure means organizing lessons around clear learning goals, carefully sequenced explanations, guided practice, and frequent checks for understanding. For policymakers and instructional leaders, this shifts the focus from whether teachers are "using structure" in a general sense to whether classroom routines are actually supporting mathematical thinking and learning.

WHICH TEACHING APPROACHES ARE MOST EFFECTIVE FOR IMPROVING MATHEMATICS LEARNING?

Pedagogy is often at the center of debates about how mathematics should be taught. Yet when looking across the evidence, not all approaches are equally represented in the research we examined. In the synthesis of the 59 studies that assessed pedagogy, only explicit pedagogy appears frequently enough to be examined on its own.

Teaching approaches that make thinking explicit lead to the strongest learning gains at the preprimary and primary levels of education.
Teaching approaches that make thinking explicit lead to the strongest learning gains at the pre-primary and primary levels of education. Looking across the evidence, the 36 evaluations of explicit pedagogy at the pre-primary and primary levels of education are associated with strong learning gains across multiple outcomes: 0.43 standard deviations, compared to 0.14 standard deviations for other pedagogical approaches.

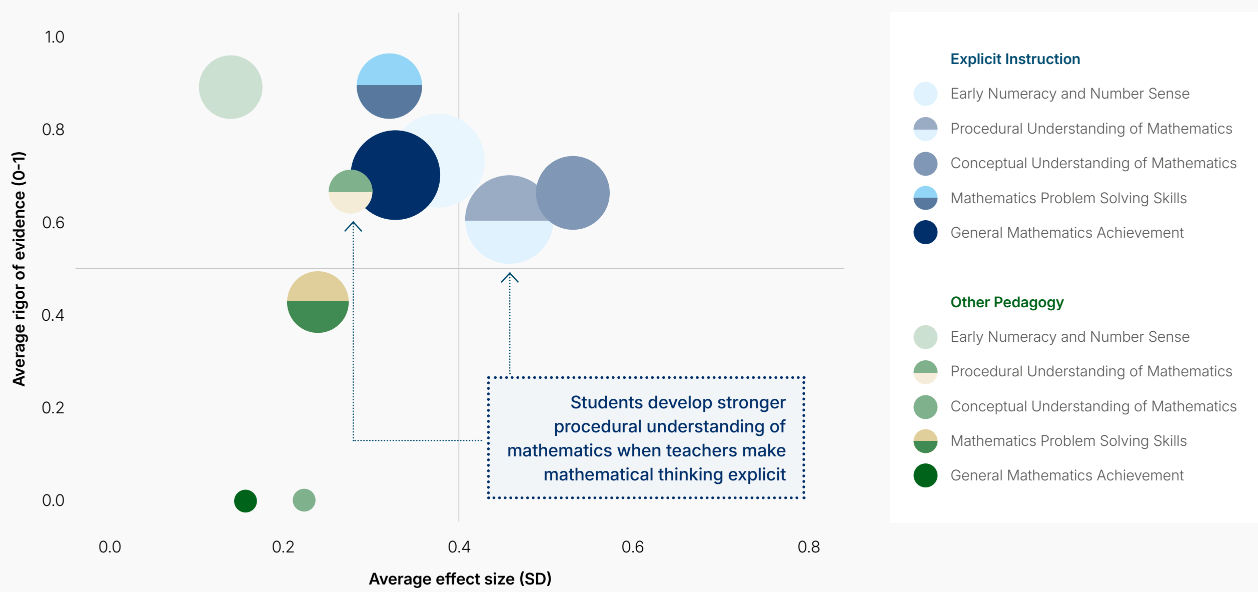
Explicit pedagogy makes mathematical thinking visible through clear explanations and guided practice.
Explicit pedagogy refers to instruction that makes the learning process visible through step-by-step explanations and guided practice that supports students as they build mathematical understanding [7]. Other approaches, while diverse, have a smaller and more fragmented evidence base, and are therefore considered together. Explicit pedagogy stands out most in developing conceptual understanding, where other approaches fall short. Effect sizes for explicit pedagogy are consistently substantial, particularly for procedural understanding, where gains reach 0.48 standard deviations, respectively, compared with 0.28 standard deviations for other pedagogical approaches [3]. For conceptual understanding, the evidence for other approaches is more limited and does not show comparable results. A similar pattern emerges for early numeracy, problem solving, and overall mathematics achievement, where explicit pedagogy is associated with larger and more consistent gains than other pedagogical approaches (see Figure 4.4).

The implication is not that explicit pedagogy is the only effective way to teach math at the pre-primary and primary levels of education.

But it clearly is an evidence-based pedagogical approach that improves learning. For policymakers and teaching leaders, this underscores the value of supporting schooling practices that make mathematical ideas and procedures explicit through clear explanation, modeling, guided practice, and attention to student reasoning.

Teachers report using explicit practices, but these are not consistently implemented in ways that support deep learning.
Much like their self-reported use of structured instruction, teachers in Latin America also report frequent use of several elements of explicit instruction. Across successive rounds of TALIS, primary education teachers in the region have consistently reported frequent use of practices such as making learning goals explicit, demonstrating procedures, and checking for understanding [9] [10]. Yet, once again, available classroom video evidence points to a more cautious conclusion. It reveals an important distinction between teacher-led instruction and effective explicit pedagogy. While lessons are often organized around demonstration and practice, they do not consistently

FIGURE 4.4: TIP FOR TEACHERS: BE EXPLICIT, EXPLAIN EACH STEP
Making thinking explicit drives stronger learning in pre-primary and primary education



Source: Näslund-Hadley et al 2026.

include the explicit explanation, reflection, and attention to student reasoning needed to make underlying concepts clear [13]. As a result, students may learn how to carry out procedures, but without fully understanding why they work or how to apply them across contexts.

Older students may benefit from a broader range of teaching approaches.
At the secondary level, however, nine high-rigor studies suggest a different pattern: other pedagogical approaches are associated with larger gains (0.44 standard deviations) than explicit pedagogy (0.26 standard deviations). This suggests that as students grow older and mathematical content becomes more complex, a broader range of instructional approaches may become effective, particularly once learners have developed the foundational skills needed to engage more independently with mathematical ideas [7].

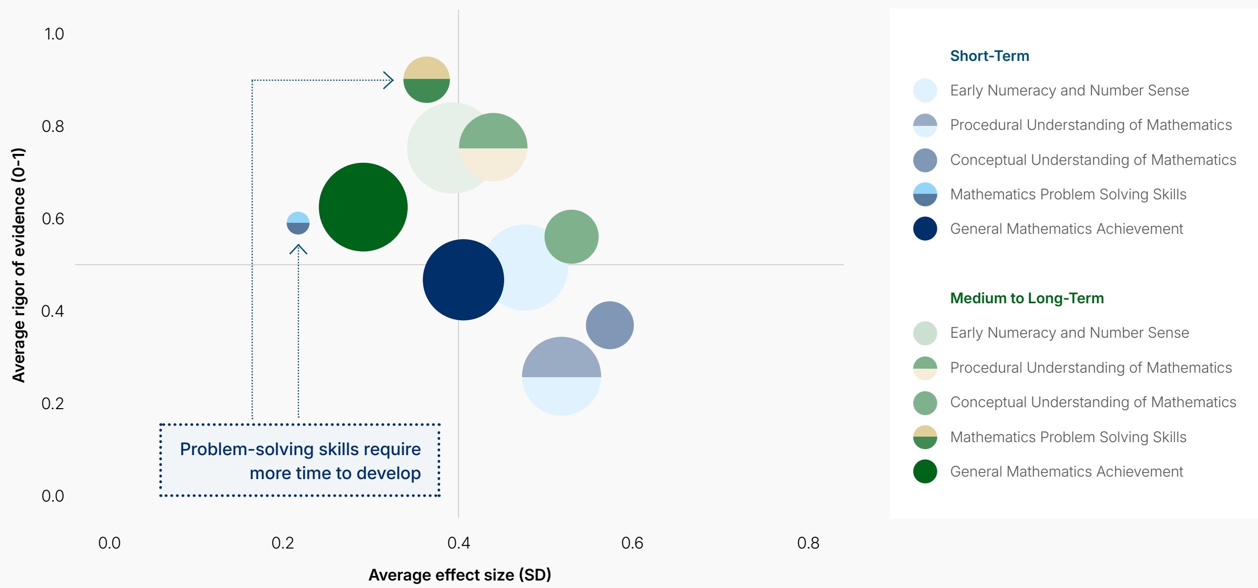
DO MORE COMPLEX MATHEMATICAL SKILLS REQUIRE MORE TIME TO DEVELOP?

Effects in education studies are often larger when learning is measured immediately after an intervention. Over time, these effects may become smaller, which is one reason short-term evaluations sometimes report larger impacts than studies that measure outcomes later [3]. Medium- and long-duration programs and teaching approaches, defined as those lasting three months or more, are associated with smaller gains than shorter interventions lasting less than three months in pre-primary and primary education (0.32 versus 0.46 standard deviations) as well as in secondary education (0.36 versus 0.54 standard deviations). This pattern is visible in the synthesis of evidence for early numeracy, procedural fluency, conceptual understanding, and general mathematics achievement. However, evidence from 13 studies suggests that mathematics problem-solving follows a different trajectory.

Problem-solving skills require sustained instruction to develop.

At the pre-primary and primary levels of education, programs and teaching approaches lasting three months or more are associated with larger gains in problem solving (0.31 standard deviations) than shorter interventions lasting up to three months (0.18 standard deviations). This suggests that problem-solving skills may require more sustained exposure, practice, and instructional support to fully develop. This is consistent with broader research in mathematics, which suggests that problem solving depends on a gradual development of strategic competence and adaptive reasoning over time [4]. The ESMATE program in El Salvador [14] offers a compelling illustration of how mathematical learning accumulates over time. Second-grade students who received a full year of instruction supported by carefully designed textbooks and teacher guides, together with teacher training and classroom observation, improved their reasoning skills by 0.29 standard deviations. Importantly, these gains did not fade once the intervention ended. Students continued to build on them during the following school year, highlighting how sustained instructional support can strengthen higher-order mathematical thinking over time (see Figure 4.5).

FIGURE 4.5: TOO FAST, TOO SLOW?
Different mathematics skills develop at different paces in pre-primary and primary education



Note: Intervention duration was coded as short (less than 3 months) and medium-long (3 or more months).
Source: Näslund-Hadley et al 2026.

WHAT DO THESE FINDINGS MEAN FOR POLICY AND PRACTICE?

■ **The evidence reviewed in this chapter shows that learning mathematics at a young age is shaped not only by what is taught but by how it is instructed, designed, and delivered.**

Different teaching approaches appear to back different kinds of learning. Traditional classroom instruction is especially important for building procedural fluency, while expanded learning opportunities, for example, through expanded school days, may help deepen conceptual understanding. Highly structured lessons and explicit pedagogy are consistently associated with stronger learning gains.

■ **Different instructional approaches support different types of mathematics learning.**

A particularly important finding for policymakers is that more complex mathematical skills, especially problem-solving, may require both more time and more direct teaching interaction. Face-to-face instruction appears especially important for supporting problem-solving skills, and longer-duration sessions are associated with stronger gains than shorter ones. This suggests that some of the most important goals of math education cannot be developed quickly or through lighter-touch delivery models alone (see Figure 4.6).

■ **Improving learning is not only about expanding access, but about improving how instruction is designed and delivered.**

The implications are clear: improving foundational numeracy requires moving beyond business-as-usual approaches and making greater use of evidence-based instructional practices. This includes strengthening structured classroom teaching to build core skills, while also ensuring that students have sufficient time and opportunities for reasoning. And develop instructional formats that encourage interaction and feedback when the goal is to acquire a deeper understanding and problem-solving ability. Improving numeracy, therefore, is not only a question of expanding access, but of aligning instructional design with the complexity of the skills students are expected to develop.

FIGURE 4.6: SO, HOW TO TEACH MATH?
Instructional approaches are associated with different mathematics learning outcomes

Instructional approach	Strongest learning outcomes	Key implication for policy and practice
WHERE TO TEACH		
Stand-alone mathematics instruction	Procedural understanding	Core mathematics skills require sustained classroom instruction and practice
Expanded learning opportunities	Conceptual understanding	Additional learning time can help deepen mathematical reasoning and connections
HOW TO DELIVER		
Face-to-face instruction	Problem solving skills Procedural and conceptual understanding	Higher order skills benefit from teacher interaction, feedback, and discussion
HOW TO ORGANIZE INSTRUCTION		
Structured lessons	Early numeracy and number sense Procedural and conceptual understanding Mathematical problem-solving skills General mathematics achievement	Clear sequencing, guided practice, and feedback strengthen learning
HOW TO TEACH		
Explicit pedagogy	Procedural understanding	Making mathematical thinking visible improves learning outcomes, particularly at the pre-primary and primary levels of education
HOW MUCH		
Longer duration support	Problem solving skills	Complex mathematical skills require sustained instructional support over time

Source: Developed by author based on the synthesis of 102 mathematics impact evaluations.

REFERENCES

[1] Davies, P., The relevance of systematic reviews to educational policy and practice. Oxford Review of Education, 2000, 26(3–4): p. 365–378. doi.org/10.1080/713688543

[2] Näslund-Hadley, E., Alonzo, H., Grossi, A., Yang, H.R., Bando, R. & Olivares, G. What Builds Mathematical Learning? Global Evidence from a Systematic Review of Education Interventions. 2026. Washington, DC: Inter-American Development Bank.

[3] Kraft, M.A., Interpreting effect sizes of education interventions. Educational Researcher, 2020, 49(4): p. 241–253. https://doi.org/10.3102/0013189X20912798

[4] National Research Council, Adding It Up: Helping Children Learn Mathematics. 2001, Washington, DC: The National Academies Press. https://doi.org/10.17226/9822

[5] Villegas-Reimers, E., Pogré, P., Freire, S., & Näslund-Hadley, E., Preparing teachers to deliver hybrid education: A framework for Latin America and the Caribbean. 2021, Washington, DC: Inter-American Development Bank. doi.org/10.18235/0005008

[6] Rosenshine, B., Principles of instruction: Research-based strategies that all teachers should know. 2012, American Educator, 36(1): p. 12–19. https://www.aft.org/sites/default/files/periodicals/Rosenshine.pdf

[7] Hiebert, J. & Grouws, D.A. The effect of classroom mathematics teaching on students' learning, in Second handbook of research on mathematics teaching and learning, F.K. Lester, Jr., Editor. 2007, Information Age Publishing: Charlotte, NC. p. 371–404.

[8] Piper, B., Ralaingita, W., Akach, L. & King, S. Improving procedural and conceptual mathematics outcomes: Evidence from a randomized controlled trial in Kenya. Journal of Development Effectiveness, 2016, 8(3): p. 404–422. https://doi.org/10.1080/19439342.2016.1149502.

[9] OECD, Creating Effective Teaching and Learning Environments: First Results from TALIS. 2009, Paris: OECD Publishing. https://doi.org/10.1787/9789264068780-en

[10] OECD, TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners. 2019, Paris: OECD Publishing. doi.org/10.1787/1d0bc92a-en

[11] Kruger, J. & Dunning, D. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. 1999, Journal of Personality and Social Psychology, 77(6): p. 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

[12] Bruns, B. & Luque, J. Great Teachers: How to Raise Student Learning in Latin America and the Caribbean. 2015, Washington, DC: World Bank. https://doi.org/10.1596/978-1-4648-0151-8

[13] Näslund-Hadley, E., Loera Varela, A. & Hepworth, K. A. What Goes On Inside Latin American Math and Science Classrooms: A Video Study of Teaching Practices. 2014, Global Education Review, 1(3): p. 110–128. https://ger.mercy.edu/index.php/ger/article/view/27

[14] Maruyama, T. & Kurosaki, T. "Developing Textbooks to Improve Math Learning in Primary Education: Empirical Evidence from El Salvador," Economic Development and Cultural Change, University of Chicago Press, 2024. 72(2), p. 833-866. https://doi.org/10.1086/721768

CHAPTER

5

"A-E-I-O-U... THAT'S HOW WORDS COME TRUE." FOUNDATIONS FOR LITERACY: WHAT WORKS, WHY IT WORKS, AND ESSENTIAL TEACHING COMPONENTS

Emma Näslund-Hadley & Ximena Duenas

Few debates in education have been as persistent, or as heated, as the question of how children learn to read. For decades, researchers, teachers, and policymakers have argued over whether reading instruction should begin with the mechanics of decoding: matching letters to sounds, sounding out words, building fluency from the ground up; or whether it should immerse children in rich texts and meaningful language, trusting that comprehension and motivation will carry the rest.

This divide, sometimes called the "reading wars," has shaped curricula, informed policy, and generated fierce disagreement. Yet the debate has often generated more heat than light, with strongly held positions on both sides outpacing the evidence, and children paying the price while adults argue.

This chapter takes stock of what the evidence actually shows. As with the mathematics programs and teaching approaches examined in Chapter 4, the core question is not whether literacy can be taught, but which method is most effective for all students to develop reading proficiency at the expected age level.

The synthesis behind this chapter is as wide in reach as the literacy challenge itself. This chapter draws on a systematic synthesis of 105 studies — 73 at the pre-primary and primary levels, 26 at the secondary level, and 6 of adult literacy education — evaluating the impact of literacy education across a wide range of contexts [1]. To enable comparisons across studies, effect sizes are expressed in standard deviations, placing results from different programs and settings on a common scale. As a rough benchmark, and as discussed in Chapter 4, an effect of 0.40 standard deviations corresponds to approximately one additional year of learning, or roughly 35 to 40 weeks of instruction, while an effect of 0.10 corresponds to about 9 to 10 weeks [2].

- **The studies included in this synthesis vary in their findings and in the strength of the evidence they provide.**

All use a treatment and comparison group, allowing for the estimation of causal impacts. Of the 105 studies included, 27 are classified as high rigor and 78 as moderate rigor. As explained in Chapter 4, high-rigor studies are primarily large randomized controlled trials with credible randomization, balanced baseline characteristics, and high treatment fidelity. Moderate-rigor studies include smaller RCTs and quasi-experimental designs, which rely on stronger assumptions to attribute observed outcomes to the intervention. No low-rigor studies were included (for a full description of the methodology (see online Methodological Annex).

- **The evidence is global - and so is its relevance.**

Of the 105 studies included in this chapter, 38 were conducted in North America, 32 in Europe, and 6 in other high-income countries. An additional 29 studies were conducted in low- and middle-income countries, including 11 from Latin America and the Caribbean (see Annex 5). Scale is another important consideration. The majority of these studies evaluate pilot-level programs, and impacts tend to diminish as coverage expands: from a pooled 0.47 standard deviations in samples of 400 students or fewer to 0.29 in larger samples. Average impacts were 0.44 standard deviations in low- and middle-income countries and 0.39 standard deviations in high-income settings, suggesting that well-designed literacy programs can produce substantial learning gains regardless of context. This finding is particularly encouraging given that the need for effective literacy instruction is often greatest in the settings where learning gaps are widest.

For policymakers, the following findings are especially relevant. They do not resolve the 'reading wars' with a simple verdict. But they do offer something more useful: a clear, evidence-based picture of which approaches produce the strongest literacy gains, across what ages, and under what conditions.

WHICH TEACHING APPROACHES ARE MOST EFFECTIVE FOR IMPROVING LITERACY SKILLS?

Much of the discussion on how to teach reading has focused on two main approaches: meaning-based and code-based. Meaning-based instruction, often called whole word or whole language instruction, focuses on focuses on understanding words and texts. It emphasizes vocabulary, comprehension, and the ability to derive meaning from written language, often encouraging students to use context and prior knowledge to interpret what they read [3]. For example, students may look at pictures with words or sentences and use the images and context to guess the meaning or listen to a story and discuss what is happening. Code-based instruction, also called phonics-based, focuses on helping students read words by understanding the relationship between letters (graphemes) and sounds (phonemes). It emphasizes phonemic awareness, phonics, and decoding, enabling students to sound out and recognize words accurately [4] [5]. For example, students may practice identifying letter sounds, blending sounds to read simple words, or breaking words into parts to decode them. A key distinction between the two approaches is not just what is taught, but how: code-based instruction teaches these relationships explicitly and systematically, while meaning-based approaches tend to assume children will absorb them gradually through exposure to texts [6].

In practice, however, this distinction is often less clear-cut. Across Latin America and most of the world, literacy instruction is frequently implemented through mixed or "balanced" approaches that combine elements of code-based and meaning-based instruction, rather than relying exclusively on a single method [7] [8] [9]. These approaches typically weave together components such as phonics, vocabulary, fluency, and comprehension, on the premise that exposing children to all dimensions of reading simultaneously will produce stronger readers.

■ **Code-based instruction leads to stronger literacy gains than meaning-based approaches.**

Of the 73 studies at the pre-primary and primary levels, 61 report on the pedagogical approach used: 16 use code-based instruction, 17 meaning-based, and 28 balanced approaches. Code-based instruction, compared to meaning-based instruction, provides somewhat stronger average literacy outcomes (0.54 vs. 0.46 standard deviations), with advantages across foundational reading skills (0.55 vs. 0.43) and reading comprehension (0.47 vs. 0.40).

The advantage is most pronounced for writing and spelling — one of the most central literacy skills — where code-based instruction produces gains of 0.60 standard deviations, five times the effect of meaning-based approaches (0.12 standard deviations). This is not a marginal difference: it reflects a fundamental mismatch between what writing and spelling demand and what meaning-based instruction provides. The skills most closely tied to mastering the alphabetic code are especially responsive to instruction that explicitly and systematically targets the relationship between sounds and letters.

The evidence showing code-based instruction leading to stronger literacy gains suggests that teaching approaches that directly support decoding, and the relationship between letters and sounds, play a particularly important role in building foundational reading skills in the early years. As explained at length in Chapter 3, this is not surprising: unlike spoken language, reading is a culturally acquired skill that does not emerge naturally through exposure alone. Learning to read requires the brain to recruit and reorganize neural systems that originally evolved for other purposes, a learning process that depends on explicit and systematic instruction.

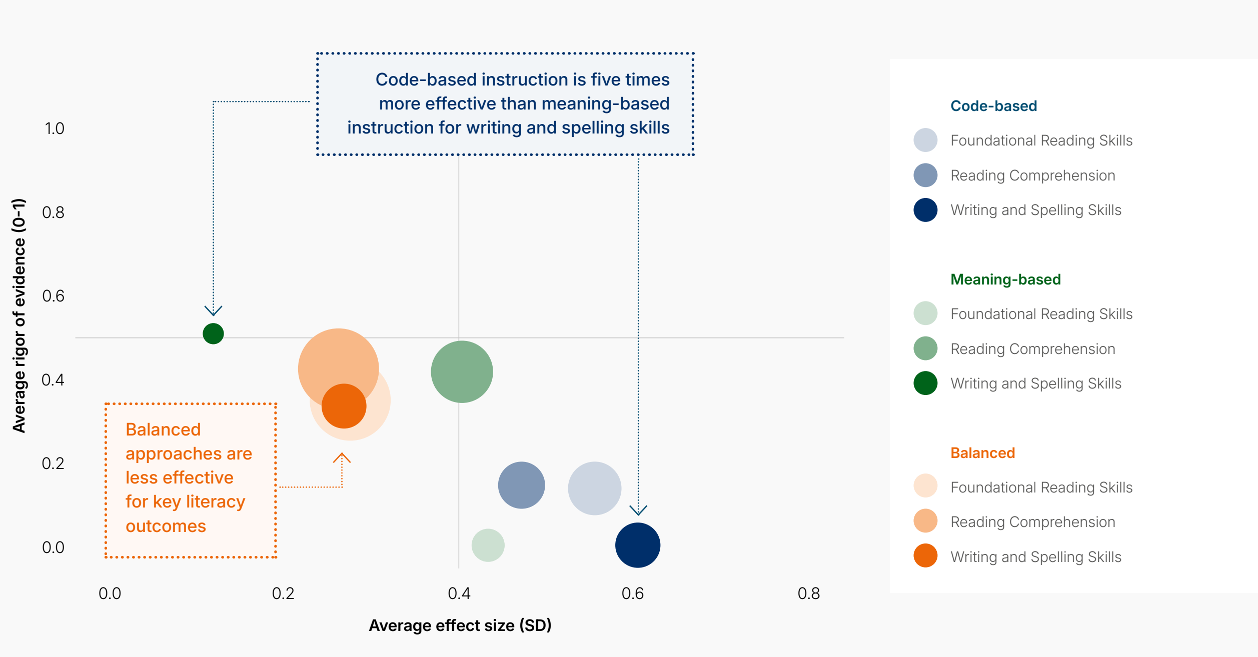
■ **More is not always better: combining instructional approaches does not necessarily lead to stronger literacy gains.**

Balanced approaches — the dominant model across Latin America and the Caribbean — produce average gains of 0.28 standard deviations in the early grades, significantly less than both code-based (0.54) and meaning-based (0.46) instruction on their own. Combining elements of code-based and meaning-based instruction does not produce the best of both worlds — it appears to dilute the impact of both approaches (see Figure 5.1).

■ **What works at the secondary level? The honest answer: we don't fully know.**

Of the 22 secondary studies that report on pedagogical approaches, none examine code-based instruction, 16 use meaning-based approaches, and 6 use balanced approaches. Meaning-based instruction produces gains of 0.45 standard deviations — nearly identical to its effect in the early grades (0.46) — and continues to outperform balanced approaches (0.20 standard deviations) by a similar margin as in pre-primary and primary education. However, secondary studies of meaning-based instruction measured only fluency, vocabulary, and reading comprehension outcomes; none assessed writing and spelling skills. Since writing and spelling are precisely where code-based instruction shows its strongest advantage in the early grades, the absence of these measures makes it impossible to draw firm conclusions about which pedagogical approach is most effective at the secondary level. This is a meaningful gap in the evidence base — and one that deserves attention from both researchers, policymakers and educators focused on secondary literacy.

FIGURE 5.1: BALANCE COMES AT A COST
Code-based instruction yields stronger literacy gains, while combined approaches show more modest effects in pre-primary and primary education



Note for Figures 5.1–5.5: Effect sizes are reported in standard deviations (SD). Following Kraft (2020), effects below 0.05 SD are considered small, effects between 0.05 and 0.19 SD are moderate, and effects of 0.20 SD or larger are large. Rigor refers to the methodological strength of the causal identification strategy. Higher rigor studies include large randomized controlled trials and stronger quasi experimental designs. No low rigor studies without a credible comparison group or causal identification strategy are included in the review. The size of the circles represents the number of studies.
Source: Näslund-Hadley et al 2026.

DOES LANGUAGE INFLUENCE WHICH TYPE OF INSTRUCTION TO USE?

Languages differ in their *orthographic transparency*, that is, how consistently letters map onto sounds. Spanish, for instance, has a highly predictable orthography, with just 27 letters corresponding to approximately 25 phonemes. Opaque languages like English and French, by contrast, have far more irregular grapheme-phoneme correspondences, making pronunciation much harder to predict from spelling alone. From this, some researchers have argued that learning to read in a transparent, orthographically regular system like Spanish may not require explicit phonemic awareness instruction to the same extent as English. The implication, drawn by some educators and curriculum designers across Latin America, is that systematic phonics may be less essential when the writing system is already highly regular [10], a view that has quietly shaped teacher preparation in the region.

But the evidence points in a clear direction and offers policymakers a signal worth heeding.
Code-based instruction proves effective regardless of the orthographic transparency of the language in which it is delivered — and the gains are, if anything, slightly larger in transparent languages (0.57 standard deviations) than in opaque ones (0.47 standard deviations). The very regularity that makes Spanish easier to decode may actually

> **While moderately structured approaches perform well for vocabulary, highly structured lessons produce stronger and more consistent gains across the full range of early literacy outcomes.**

amplify the returns to phonics: when each letter reliably maps to a single sound, systematic instruction in those correspondences gives children a near-complete set of tools to unlock any word they encounter. The evidence does not support the idea that Spanish-speaking children naturally absorb letter-sound relationships through exposure alone. Even in the phonetically regular orthographies prevalent across much of Latin America, many children struggle to learn to read, and experimental evidence shows that explicit phonics instruction can produce large gains in Spanish language instruction. In a randomized trial, struggling first-grade Spanish speakers in the United States who received an intensive phonics-based instruction showed improvements ranging from 0.34 to 1.06 standard deviations across Spanish literacy assessments [11].

Meaning-based instruction, on the other hand, shows particular promise beyond the primary schooling years.

Of the 39 meaning-based studies in the synthesis, 22 assess secondary and adult learners, and the results are encouraging: gains of 0.45 standard deviations at the secondary level and 0.83 standard deviations in adult education, particularly in vocabulary and reading comprehension. These figures, however, should be interpreted with care. The adult education estimate is based on just six studies, all of medium rigor. Adult literacy acquisition is also neurologically distinct from childhood reading — adults who did not learn to read as children show different patterns of neural reorganization and may respond differently to instruction, as discussed in Chapter 3. While 22 studies are a solid evidence base by the standards of most syntheses in education, the adult subsample remains thin to draw solid insights. These results nonetheless take on particular urgency. As shown in Chapter 2, more than one in two 15-year-olds in Latin America lack foundational literacy skills, and nine in ten adults across the four Latin American countries participating in OECD's PIAAC evaluation (Programme for the International Assessment of Adult Competencies) — Chile, Ecuador, Mexico, and Peru — do not reach a baseline proficiency. Reaching these learners will require not only sustained political commitment, but a much stronger and more targeted evidence base on what works for teaching reading and writing beyond the early years.

WHERE IS LITERACY INSTRUCTION MOST EFFECTIVE?

Reading and writing are commonly taught as separate subjects, with structured time set aside for literacy instruction within the school day. Yet education systems increasingly look beyond this traditional model, integrating literacy into other subjects, extending learning through after-school programs, or embedding it in broader language development activities. These approaches do not replace direct literacy instruction but seek to expand the contexts in which students encounter and practice reading and writing.

The classroom remains the most powerful setting for building foundational literacy skills.

Evidence from pre-primary and primary education shows that teaching literacy as an independent subject is associated with stronger average gains than expanded learning approaches, which are designed to complement rather than replace regular classroom instruction. At the pre-primary and primary levels of education, independent subject instruction generates an average effect of 0.51 standard deviations, while expanded learning settings — adding to what students receive in class — are associated with 0.38 standard deviations. While both produce meaningful gains, the evidence consistently points to dedicated in-class literacy instruction as the stronger overall driver of learning. Yet what each setting develops best differs considerably.

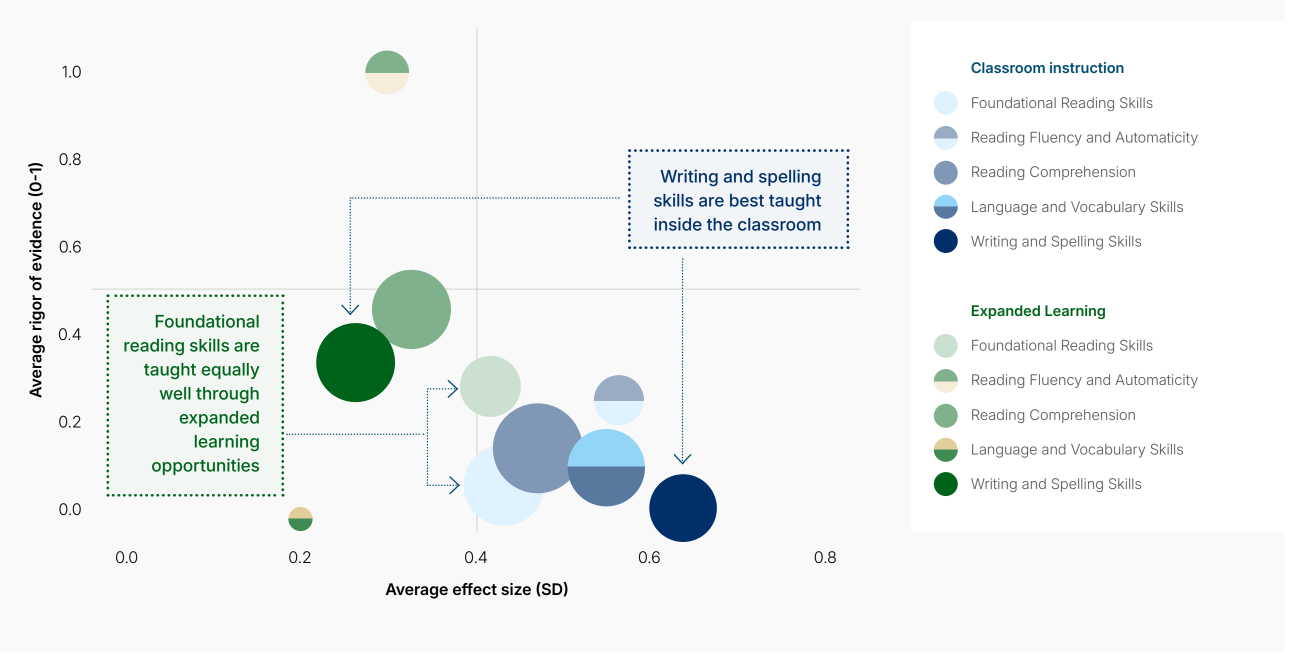
The classroom builds vocabulary and writing, while expanded settings reinforce reading and fluency. Independent subject instruction shows particular strength in writing and spelling skills (0.63 standard deviations) and language and vocabulary (0.54 standard deviations), suggesting it is especially well suited to building the structured, cumulative skills that literacy requires. Expanded learning settings, by contrast, show notable gains in foundational reading skills (0.43 standard deviations) and meaningful gains in reading fluency (0.30 standard deviations), suggesting they can reinforce decoding and fluency practice beyond the classroom (see Figure 5.2).

It is never too late to learn to read.

A particularly encouraging finding is that the benefits of dedicated literacy instruction are not confined to the early years. Independent subject instruction at the secondary level is associated with gains of 0.51 standard deviations, identical to those seen in pre-primary and primary education, suggesting that structured literacy teaching remains just as powerful for older learners. In adult literacy education, the evidence base is smaller but points to even larger effects (0.83 standard deviations), though these results should be interpreted with caution given the limited number of studies and their medium rigor. Together, these findings carry an encouraging message: it is never too late to learn to read, and never too late to teach. The gains achievable at the secondary and adult levels are not

FIGURE 5.2: WHERE YOU TEACH SHAPES WHAT YOU TEACH

Classroom instruction drives foundational literacy skills, while expanded learning reinforces reading and fluency in pre-primary and primary education



Source: Näslund-Hadley et al 2026.

consolation prizes, for they are comparable to what structured instruction produces in the early grades. For the millions of adolescents and adults in Latin America and the Caribbean who never mastered the basics, this is not a small finding.

DOES HOW WE DELIVER LITERACY INSTRUCTION CHANGE WHAT STUDENTS LEARN?

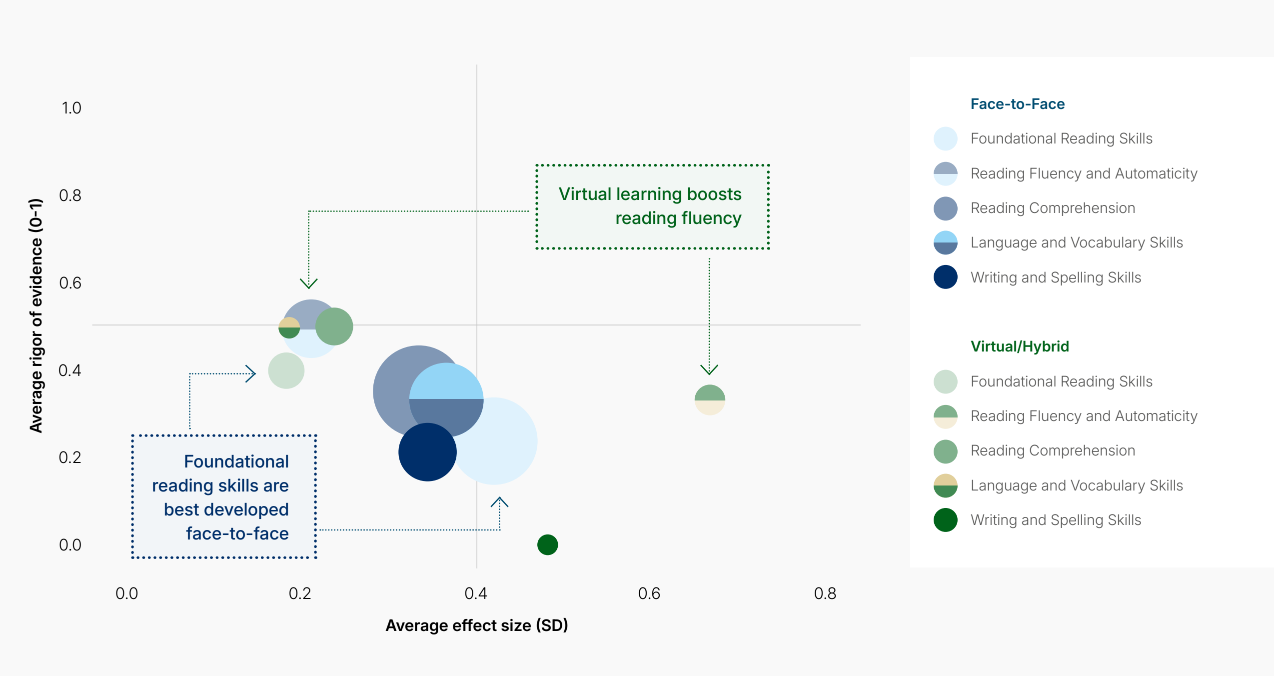
Learning to read and write has always been understood as a deeply interactive process, one that depends on a teacher hearing a child sound out words, catching where comprehension breaks down, and responding in the moment. Yet digital tools now offer new ways to support literacy development, from adaptive reading platforms to audio-assisted practice. Across Latin America and the Caribbean, these technologies are increasingly present in classrooms, raising a practical question for educators and policymakers: Does the method of instruction shape how much and what kind of literacy students develop? In this chapter, virtual approaches refer to instruction delivered entirely through digital technologies, while hybrid approaches combine face-to-face and online learning.

Overall, face-to-face instruction remains the stronger driver of literacy learning.
There are things a skilled teacher does in the moment — hearing where a child stumbles on a word, catching the exact point where comprehension breaks down, adjusting pace and explanation on the fly — that no screen can fully replicate. The evidence reflects this. In pre-primary and primary education, face-to-face instruction is associated with average gains of 0.40 standard deviations, compared with 0.23 standard deviations for hybrid approaches and 0.34 standard deviations for fully virtual models. This pattern holds at the secondary level as well, where face-to-face instruction generates gains of 0.41 standard deviations, compared with 0.23 standard deviations for virtual or hybrid approaches. In adult literacy education, face-to-face instruction produces even larger gains (0.90 standard deviations), compared with 0.73 standard deviations for virtual or hybrid delivery. Nevertheless, both estimates should be interpreted with caution, given the small number of studies and medium methodological rigor.

Looking at specific outcomes in pre-primary and primary education, highly structured teaching shows its clearest edge in writing and spelling skills.

Don't abandon the classroom...
but don't underestimate the screen either.
When looking at specific outcomes in pre-primary and primary education, face-to-face instruction shows particular strength in developing foundational reading skills (0.42 standard deviations), reading comprehension (0.33 standard deviations), and literacy-specific attitudes (0.51 standard deviations). By contrast, virtual and hybrid approaches show a surprising advantage in reading fluency and automaticity (0.67 standard deviations versus 0.21 for face-to-face), suggesting that well-designed digital tools, such as adaptive reading platforms and audio-supported practice, may be especially effective at building the kind of repeated, paced practice that fluency development requires (see Figure 5.3).

FIGURE 5.3: HOW YOU DELIVER LITERACY INSTRUCTION CHANGES WHAT STUDENTS LEARN
Face-to-face instruction drives foundational reading skills, while digital tools show an edge in fluency in pre-primary and primary education



Source: Näslund-Hadley et al 2026.

The ways literacy is taught are not interchangeable, but they develop different literacy skills. Face-to-face instruction is not always superior, but the way literacy is taught shapes learning outcomes. For policymakers and instructional designers, this suggests deploying face-to-face instruction strategically for the skills that depend most on teacher feedback, discussion, and close observation of student reading, while ensuring that virtual and hybrid environments are purposefully designed to support specific literacy outcomes.

HOW SHOULD LITERACY LESSONS BE ORGANIZED TO SUPPORT LEARNING?

Literacy instruction, like all teaching, varies enormously in how it is structured and presented. Some lessons follow a clear and deliberate sequence: the teacher sets an explicit goal, introduces a skill step by step, models it, guides students through practice, and checks for understanding before moving on [12]. Others give students more latitude to explore texts, develop understanding at their own pace, and engage with reading and writing in less prescribed ways. The question is not which style feels more natural or culturally familiar, but which produces stronger literacy outcomes [13].

The evidence shows that structure matters in the early literacy classroom, although its impact is not quite as powerful as in mathematics. Of the 105 studies analyzed, 61 examine the level of instructional structure in pre-

For Latin America and the Caribbean, the stakes are too high for gradualism. The evidence is clear; what the region needs now is the will to teach reading the way the evidence says it works, systematically, sequentially, and without waiting for children to fall irreversibly behind.

primary and primary literacy education, and a clear and consistent pattern emerges: highly structured lessons outperform moderately structured ones. Highly structured literacy instruction is associated with gains of 0.41 standard deviations, compared with 0.37 standard deviations for moderately structured approaches.

Evidence from Kenya illustrates what that structure adds in practice: a randomized trial found that teacher coaching and professional development did not translate into meaningful literacy gains (0.20 standard deviations), but adding structured lesson plans to the package raised average effects to 0.83 standard deviations — at an additional cost of just $0.16 per pupil [14].

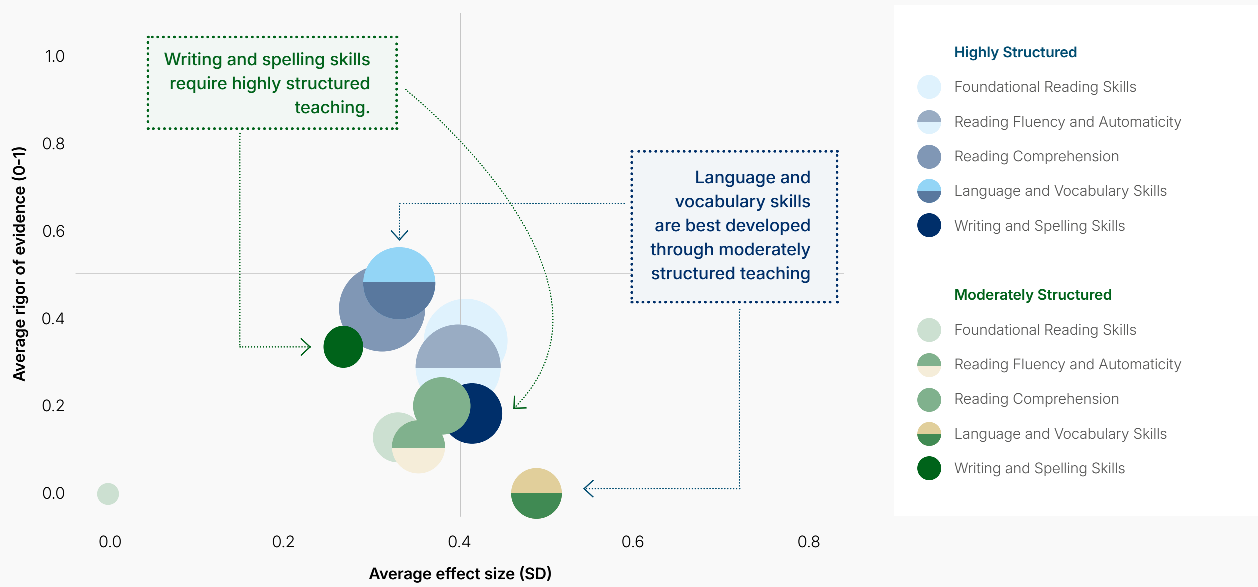
A Latin American example points in the same direction. The 'Let's All Learn to Read' program in Colombia combined a carefully structured weekly curriculum — including phonological awareness, vocabulary, oral language, and storybook reading — with intensive in-class coaching and complementary materials for teachers and students. Students in treated schools gained 0.39 standard deviations in overall literacy at the end of first grade — and those gains held through second and third grade, even after the program ended [15].

■ **Structured instruction provides the greatest advantage for foundational literacy development.**
Looking at specific outcomes in pre-primary and primary education, highly structured teaching shows its clearest edge in writing and spelling skills (0.42 standard deviations compared with 0.27 for moderately structured lessons) and foundational reading skills (0.40 versus 0.35 standard deviations). These are precisely the skills that depend most on carefully sequenced, step-by-step instruction — clear goals, deliberate progression from simple to complex, guided practice, and frequent checks for understanding — rather than open-ended exploration of texts (see Figure 5.4).

■ **This picture shifts, however, when it comes to language and vocabulary development.**
In this context, moderately structured lessons — where teachers guide conversation and exploration rather than following a fixed sequence — are associated with stronger gains (0.49 standard deviations) than highly structured ones (0.33 standard deviations), based on 7 and 21 studies respectively. This makes intuitive sense: building a rich vocabulary requires exposure to a varied language, discussions, and contexts, not just step-by-step instruction. However, while moderately structured approaches perform well for vocabulary, highly structured lessons produce stronger and more consistent gains across the full range of early literacy outcomes. For policymakers, this consistency matters: a structured approach is more predictable and less likely to leave gaps in children's foundational skills.

■ **As literacy skills mature, the advantages of the structured approach fade... but the gains do not.**
In secondary and adult literacy education, both methods of teaching produce similar average gains — a pattern that likely reflects the cumulative linguistic knowledge that older learners bring to the classroom. Over time, readers build up layers of syntactic, semantic, grammatical, and pragmatic understanding through exposure and

FIGURE 5.4: STRUCTURE MATTERS
Highly structured lessons drive stronger and more consistent literacy gains in pre-primary and primary education



Source: Näslund-Hadley et al 2026.

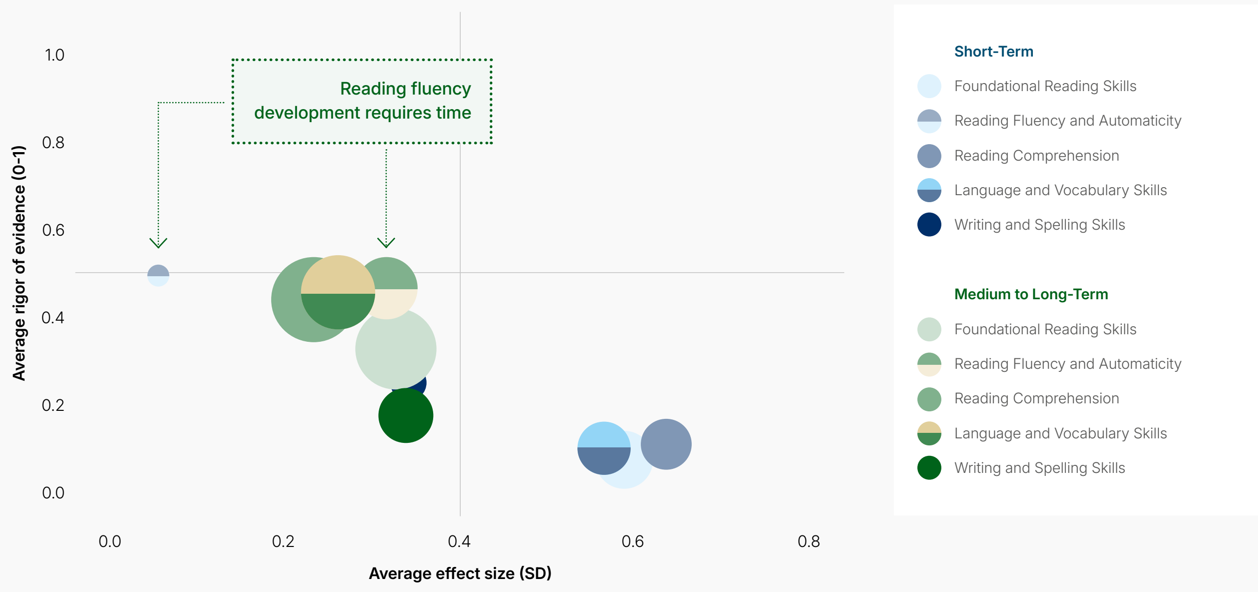
experience, which gradually enables more autonomous engagement with texts. It is this accumulated knowledge — not maturity itself — that may allow older learners to benefit from a wider range of instructional approaches. This is consistent with broader learning science research suggesting that instructional guidance is particularly important when learners are novices, whereas learners with richer knowledge structures can draw on them to support learning even in less structured environments [16].

DO SOME LITERACY SKILLS REQUIRE MORE TIME TO DEVELOP THAN OTHERS?

At first glance, the evidence on program duration seems to defy common sense. Short programs appear to outperform longer ones — but this reflects what short programs tend to measure, not what they actually build. Short programs focus on a narrow set of concrete skills that are easier to capture in assessments administered close to the end of instruction. Across the evidence base, short programs (up to three months) are associated with average gains three times as high as long-duration programs — but that gap closes when looking at specific outcomes like writing and spelling, where medium and long-term programs perform equally well.

But reading fluency and ‘automaticity’ seem to follow a different pattern. Unlike foundational reading skills, which respond quickly to short and intensive instruction, fluency development requires sustained exposure and practice over time. In pre-primary and primary education, medium and long-duration programs are associated

FIGURE 5.5: FLUENCY CANNOT BE RUSHED
Short programs drive rapid gains in foundational reading, while sustained instruction builds fluency skills in pre-primary and primary education



Source: Näslund-Hadley et al 2026.

with gains of 0.31 standard deviations in reading fluency and automaticity, compared with just 0.05 standard deviations for short-duration programs. This is not simply a matter of needing more practice — it reflects how the brain builds fluency. As Chapter 3 explains, fluency represents the gradual automatization of decoding: the brain shifts from slow, effortful processing of individual letters and sounds to a fast, automatic recognition of whole words, a transition that requires repeated exposure over time and cannot be rushed by intensive short-term instruction. Short programs can build decoding skills, but they cannot complete the neural consolidation that fluency requires (see Figure 5.5). This pattern is consistent with evidence from intervention studies with older struggling readers. One example is a rigorous randomized trial conducted in the United States that found that fourth and fifth graders with significant reading difficulties who received two years of the Simple View of Reading — a framework holding that literacy depends on two components: decoding and listening comprehension — showed substantially greater fluency gains than those who received only one year, while gains in word reading were similar across both groups [17].

But fluency does not arrive on a schedule. Children do not become confident readers simply because they have been promoted to the next grade. A child who struggles to read smoothly at the end of second grade has not “missed her window” — she is still on the path, just not yet at the same point as her classmates. The challenge for education systems and teachers is to know where each child actually stands, and not simply to grade them as they walk out the classroom door. That requires clear decisions about what to teach first and in what order, to assess each child’s progress over time, and to know how to adjust when a child is falling behind.

WHAT DO THESE FINDINGS MEAN FOR POLICY AND PRACTICE?

Literacy instruction is not a single intervention, but a set of design choices — about how literacy is taught, in what setting, and for how long — each of which shapes what students learn. The findings, summarized in Figure 5.6, do not resolve every question, but they offer a consistent and actionable picture for policymakers.

- **Make code-based instruction the foundation.**

Code-based instruction outperforms meaning-based approaches in the early grades, with a fivefold advantage in writing and spelling, yet meaning-based and balanced models remain dominant across Latin America and the Caribbean. Closing that gap requires deliberate policy action: curriculum alignment, teacher preparation, structured materials, coaching, and assessment tools that track decoding, fluency, and comprehension over time.

- **Balance is not a strategy: combining approaches dilutes both.**

Mixed approaches do not automatically produce stronger gains. Instructional design should follow the evidence on which combinations work, not the intuition that more comprehensive means more effective.

- **Protect classroom time, and complement it strategically.**

Dedicated, face-to-face literacy instruction outperforms expanded settings and virtual models for most outcomes. But digital tools show a real edge in fluency, and expanded settings reinforce foundational reading: use each modality for what it does best.

- **Structure matters, especially in the early years.**

Highly structured lessons consistently outperform moderately structured ones for foundational reading and writing. The question for instructional leaders is not whether teachers report using structure, but whether that structure is actually building skills, step by step.

- **Literacy development cannot stop at primary school.**

Gains at the secondary and adult levels are comparable to those in the early grades, but the near-absence of evidence on code-based instruction beyond primary school makes building that evidence base a research and policy priority.

For Latin America and the Caribbean, the stakes are too high for gradualism. The evidence is clear; what the region needs now is the will to teach reading the way the evidence says it works, systematically, sequentially, and without waiting for children to fall irreversibly behind.

FIGURE 5.6: SO HOW TO TEACH READING & WRITING?
Different teaching strategies lead to varying literacy outcomes

Instructional approach	Strongest learning outcomes	Key implication for policy and practice
HOW TO TEACH		
Code-based Instruction	Foundational reading skills; writing and spelling	Decoding is the foundation; explicit, systematic phonics instruction produces the strongest gains in the early years
Meaning-based Instruction	Vocabulary; reading comprehension; secondary and adult literacy	Meaning-based approaches show effects beyond primary education, but there is no evidence for code-based instruction at those levels
Balanced Approaches		Combining code-based and meaning-based instruction does not produce the best of both worlds. In fact, balanced approaches underperform both single approaches.
HOW TO ORGANIZE INSTRUCTION		
Highly structured lessons	Foundational reading skills; writing and spelling	Structure matters most in the early years for skills that depend on sequenced, explicit instruction
Longer duration programs	Reading fluency and automaticity	Fluency cannot be built quickly —it requires sustained practice and consolidation over time
WHERE TO TEACH		
Stand-alone instruction	Reading comprehension; language and vocabulary; writing and spelling; reading fluency and automaticity	Dedicated literacy instruction outperforms expanded learning for the skills that require sustained, structured teaching
Expanded learning settings	Foundational reading skills	Expanded settings can reinforce foundational reading beyond the classroom but cannot replace dedicated literacy instruction
HOW TO DELIVER		
Face-to-face instruction	Foundational reading skills; Reading comprehension	Face-to-face teaching outperforms virtual and hybrid models for most literacy outcomes
Digital tools	Reading fluency and automaticity	Well-designed digital tools show an edge in fluency —deploy them strategically

Source: Developed by authors based on the synthesis of 105 literacy impact evaluations across the world.

REFERENCES

[1] Näslund-Hadley, E., Duenas, X., Alonzo, H., Guimaraes, B., Grossi, A., Bando, R. What Builds Literacy Skills? Global Evidence from a Systematic Review of Education Interventions. 2026. Washington, DC: Inter-American Development Bank.

[2] Kraft, M.A., Interpreting effect sizes of education interventions. *Educational Researcher,* 2020, 49(4): p. 241–253. doi.org/10.3102/0013189X209127

[3] Castles, A., Rastle, K. and Nation, K., Ending the Reading Wars: Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest,* 2018, 19(1): p. 5–51. https://doi.org/10.1177/1529100618772271

[4] *National Reading Panel, Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction*. 2000, National Institute of Child Health and Human Development: Bethesda, MD. https://www.nichd.nih.gov/publications/product/247

[5] Gough, P.B. and Tunmer, W.E., "Decoding, Reading, and Reading Disability." *Remedial and Special Education*, 1986, 7(1): p. 6–10. https://doi.org/10.1177/074193258600700104

[6] Palazón López, J. (ed.), *Intervenciones especializadas para la mejora de la lectura, la escritura, el lenguaje y el habla.* Madrid: Ediciones Pirámide, 2025.

[7] Rivas, A., Dueñas, X., and Vinacur, T., *Políticas de alfabetización inicial en América Latina: un mapa de las acciones en curso.* 2025, Banco Interamericano de Desarrollo: Washington, DC. https://publications.iadb.org/es/politicas-de-alfabetizacion-inicial-en-america-latina-un-mapa-de-las-acciones-en-curso

[8] Akyeampong, K., Andrabi, T., Banerjee, A., Banerji, R., Dynarski, S., Glennerster, R., Grantham-McGregor, S., Muralidharan, K., Piper, B., Ruto, S., Saavedra, J., Schmelkes, S., Yoshikawa, H., Cost-effective Approaches to Improve Global Learning: What Does Recent Evidence Tell Us Are Smart Buys for Improving Learning in Low- and Middle-income Countries? 2023, Global Education Evidence Advisory Panel (GEEAP), World Bank Group: Washington, DC. http://documents.worldbank.org/curated/en/099420106132331608

[9] Scarborough, H.S., "Connecting Early Language and Literacy to Later Reading (Dis)Abilities: Evidence, Theory, and Practice." In S.B. Neuman and D.K. Dickinson (Eds.), *Handbook of Early Literacy Research*, 2001, New York, NY: Guilford Press, p. 97–110.

[10] Goldenberg, C., Tolar, T.D., Reese, L., Francis, D.J., Bazán, A.R., & Mejía-Arauz. R., "How important is teaching phonemic awareness to children learning to read in Spanish?" *American Educational Research Journal*, 2014, 51(3), 604–633. https://doi.org/10.3102/0002831214529082

[11] Borman, T.H., Borman, G.D., Houghton, S., Park, S.J., Zhu, B., Martin, A., and Wilkinson-Flicker, S., "Addressing Literacy Needs of Struggling Spanish-Speaking First Graders: First-Year Results From a National Randomized Controlled Trial of Descubriendo la Lectura." *AERA Open*, 2019, 5(3): p. 1–14.
https://doi.org/10.1177/2332858419870488

[12] Rosenshine, B., Principles of instruction: Research-based strategies that all teachers should know. 2012, *American Educator*, 36(1): p. 12–19. https://www.aft.org/sites/default/files/periodicals/Rosenshine.pdf

[13] Smith, F., Joining the Literacy Club: Further Essays into Education. 1988, Portsmouth, NH: Heinemann.

[14] Piper, B., Zuilkowski, S.S., Dubeck, M., Jepkemei, E., and King, S.J., "Identifying the Essential Ingredients to Literacy and Numeracy Improvement: Teacher Professional Development and Coaching, Student Textbooks, and Structured Teachers' Guides." *World Development*, 2018, 106: p. 324–336.
https://doi.org/10.1016/j.worlddev.2018.01.018

[15] Álvarez Marinelli, H., Berlinski, S., Busso, M., and Martínez Correa, J., "Improving early literacy through teacher professional development: Experimental evidence from Colombia." *Journal of Public Economics Plus*, 2023, 4: 100019. https://doi.org/10.1016/j.pubecp.2023.100019

[16] Bransford, J.D., Brown, A.L. and Cocking, R.R. (Eds.), *How People Learn: Brain, Mind, Experience, and School.* 2000, National Academy Press: Washington, DC. https://nap.nationalacademies.org/catalog/9853/how-people-learn-brain-mind-experience-and-school-expanded-edition

[17] Miciak, J., G. Roberts, W.P. Taylor, M. Solis, Y. Ahmed, S. Vaughn, and J.M. Fletcher. *The Effects of One versus Two Years of Intensive Reading Intervention Implemented with Late Elementary Struggling Readers. Learning Disabilities Research & Practice*, 2018, 33(1): p. 24–36. doi.org/10.1111/ldrp.12159

CHAPTER

6

FALLING BEHIND ≠ LEFT BEHIND. REMEDIAL EDUCATION AND TUTORING: WHAT WORKS, WHY IT WORKS, AND ESSENTIAL INSTRUCTIONAL INGREDIENTS

Emma Näslund-Hadley & Pablo Zoido

The classroom moves at one pace, follows one sequence, and relies on students arriving ready to learn. Yet students learn in different ways and at different rhythms. Some need more time, others a different approach, and yet others both. For those who fall behind because they missed acquiring foundational skills early, lost ground along the way, or simply needed more time than the system allowed, that gap grows quietly, year by year, until it becomes hard to close.

Falling behind is not the same as being left behind.
For these students, tutoring and remedial education offer a second chance through targeted instructional support provided outside or alongside regular classroom teaching. In Latin America and the Caribbean, interest in these approaches has grown sharply, with promising programs suggesting that individualized support can help students recover ground faster than classroom instruction alone [1]. But identifying the specific features that drive learning outcomes and are cost-effective remains an active area of inquiry. Who should deliver the support? How should it be organized? Does it matter whether instruction is delivered one-on-one or in small groups? Is a trained teacher better than an external tutor?
This chapter brings together the global evidence to answer those questions, and the findings are both encouraging and specific.

Tutoring and remedial education have been studied at every stage of schooling, making it possible to identify what works and for whom.
This chapter draws on a systematic synthesis of 70 studies — 12 at the pre-primary level, 47 at the primary level, and 11 at the secondary level — evaluating the impact of tutoring and remedial education programs across a wide

range of contexts and implementation models [2]. Effect sizes are expressed in standard deviations to enable comparisons across studies: an effect of 0.40 standard deviations corresponds to approximately one additional year of learning, while an effect of 0.10 corresponds to about 9 to 10 weeks [3]. All studies include a treatment and comparison group, which allows a causal estimation of the programs' impacts.

Unlike the preceding sections, this chapter combines studies across both mathematics and literacy, yielding a sufficient number of studies at the pre-primary, primary, and secondary levels of education to report strong findings.

NOT ALL EVIDENCE IS CREATED EQUAL, BUT THE STUDIES REVIEWED MEET A CONSISTENT STANDARD OF METHODOLOGICAL RIGOR

Across the 70 studies included in this synthesis, findings vary in both magnitude and direction, reflecting the diversity of existing tutoring and remedial education models. Thirty-seven are classified as high rigor and 33 as moderate rigor. As in the preceding chapters, high-rigor studies are primarily large randomized controlled trials (RCT) with credible randomization procedures, balanced baseline characteristics, and strong treatment fidelity. Meanwhile, moderate-rigor studies include smaller RCTs and quasi-experimental designs that rely on stronger assumptions to causally link observed outcomes to the intervention. No low-rigor studies were included. For a full description of the methodology (see online Methodological Annex).

The evidence on tutoring and remedial education is global, and its lessons travel. One study was conducted across multiple countries, accounting for a total of 72 country evaluations. Of these, 44 were conducted in North America, 16 in Europe and Australia, and 12 in low- and middle-income countries, including 8 from Latin America and the Caribbean (see Annex 6).

- **Context shapes how much tutoring helps, and the gaps between settings are worth understanding.**

In high-income settings, effects average 0.36 standard deviations — equivalent to roughly one academic year of learning. In low- and middle-income countries, average effects are more modest, at 0.10 standard deviations, based on 10 high-rigor studies. Scale matters too: most of these evaluations cover programs at pilot scale, and effects tend to shrink as programs grow: from a pooled 0.40 standard deviations in samples of 400 or fewer students to 0.18 in larger ones. For policymakers in Latin America and the Caribbean, this gap is an important signal: effectiveness may depend heavily on the availability of trained tutors, adequate instructional materials, and the systems needed to identify and reach students who have fallen behind. Understanding what drives this difference is as important as recognizing that it exists.

ARE STUDENTS BEING HELPED TO KEEP UP OR CATCH UP?

Definitions of remedial education and tutoring vary across the literature, and the boundary between the two is often fluid. For this synthesis, the distinction is captured through curriculum alignment: some programs are fully aligned with the grade-level curriculum; others are partially aligned, bridging grade-level content with foundational gaps; and others follow a remedial foundational curriculum, targeting students who need to recover core competencies before engaging with grade-level material. The vast majority of programs were implemented during the school day.

Definitions of remedial education and tutoring vary across the literature, and the boundary between the two is often fluid.

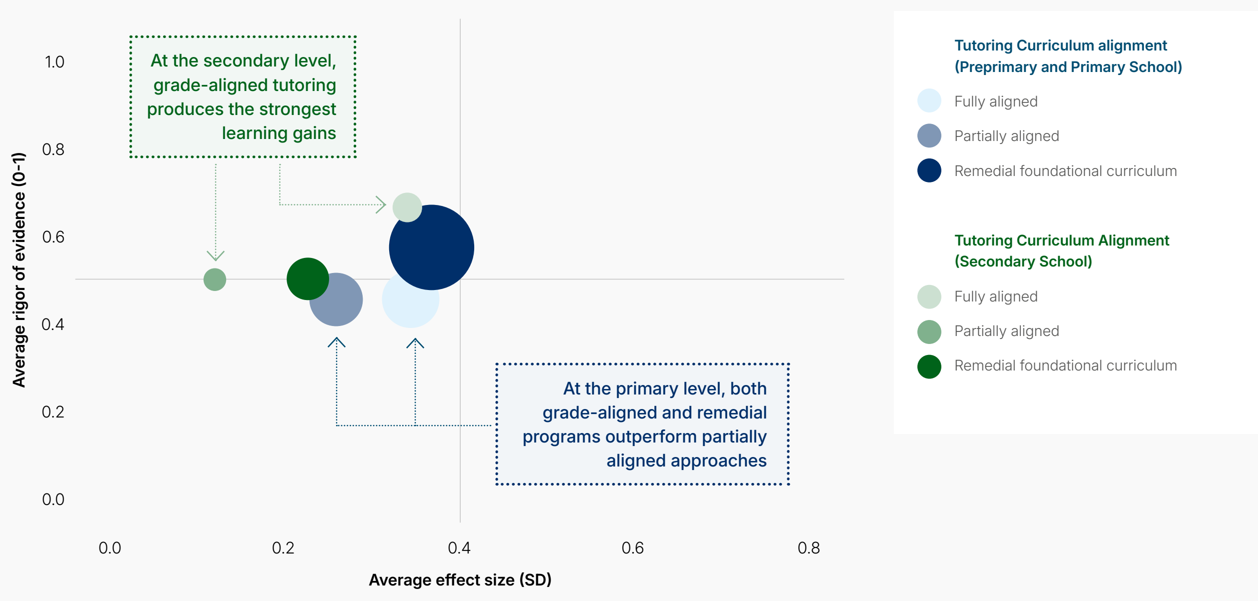
Typically, tutoring is designed to help students to keep up by reviewing homework and/or reinforcing materials already learned in the classroom. Remedial education is focused on rebuilding foundational skills that a student is missing, so as to overcome learning gaps and catch up.

For this synthesis, the distinction is captured through three curriculum alignments. The first are programs fully aligned with the grade-level curriculum; a second approach is partially aligned, bridging grade-level content with foundational gaps; and a third approach follows a remedial foundational curriculum, targeting students who need to recover core competencies before engaging with grade-level material. The vast majority of programs were implemented during the school day.

Targeted support works best when it knows what it is trying to do: help a student to keep up or to catch up.

All three approaches produce meaningful learning gains, but the pattern carries a clear policy message: programs that make a clear curricular choice — either fully aligned with grade-level content or explicitly remedial — consistently outperform those that approach tutoring and remedial education simultaneously (see Figure 6.1). At the pre-primary

FIGURE 6.1: KEEP UP OR CATCH UP: DOES CURRICULUM ALIGNMENT MATTER?
At the pre-primary and primary levels, grade-level and remedial foundational programs outperform partially aligned models; at secondary, grade-aligned tutoring pulls ahead



Note for Figures 6.1–6.5: Effect sizes are reported in standard deviations (SD). Following Kraft (2020), effects below 0.05 SD are considered small, effects between 0.05 and 0.19 SD are moderate, and effects of 0.20 SD or larger are large. Rigor refers to the methodological strength of the causal identification strategy. Higher rigor studies include large randomized controlled trials and stronger quasi experimental designs. No low rigor studies without a credible comparison group or causal identification strategy are included in the review. The size of the circles represents the number of studies.
Source: Näslund-Hadley et al 2026.

and primary levels, programs using a remedial foundational curriculum produce strong average gains of 0.36 standard deviations across 35 high-rigor studies — teaching students at their actual learning level rather than their enrolled grade level. In Colombia, for example, a remedial literacy program provided struggling third graders with tutoring focused on second grade reading skills, improving literacy outcomes by 0.36 standard deviations [4]. Programs fully aligned with the grade-level curriculum produce comparable gains of 0.34 standard deviations across 13 studies. Partially aligned programs average 0.24 standard deviations across 11 studies, suggesting that trying to bridge grade-level and foundational content simultaneously may dilute the effectiveness of both.

■ **As students get older, the case for grade-aligned tutoring grows stronger.**
At the secondary level, a different pattern emerges: programs fully aligned with the grade-level curriculum are associated with stronger gains (0.33 standard deviations) than remedial foundational approaches (0.22 standard deviations) — a difference equivalent to roughly four months of learning. This suggests that as students progress through school, support that keeps pace with grade-level content may become more important than going back to foundational skills. This does not mean that remediation loses its value at the secondary level, a 0.22 standard deviation effect remains large, nor that foundational gaps have disappeared by this stage: as documented in Chapter 2, many students in the region advance to secondary school without mastering minimum competencies. Rather, for these students, support anchored to grade-level content appears to offer more leverage than returning to foundational material alone. However, with only 11 studies at this level, these findings should be interpreted with caution and might point to an important gap in the evidence base.

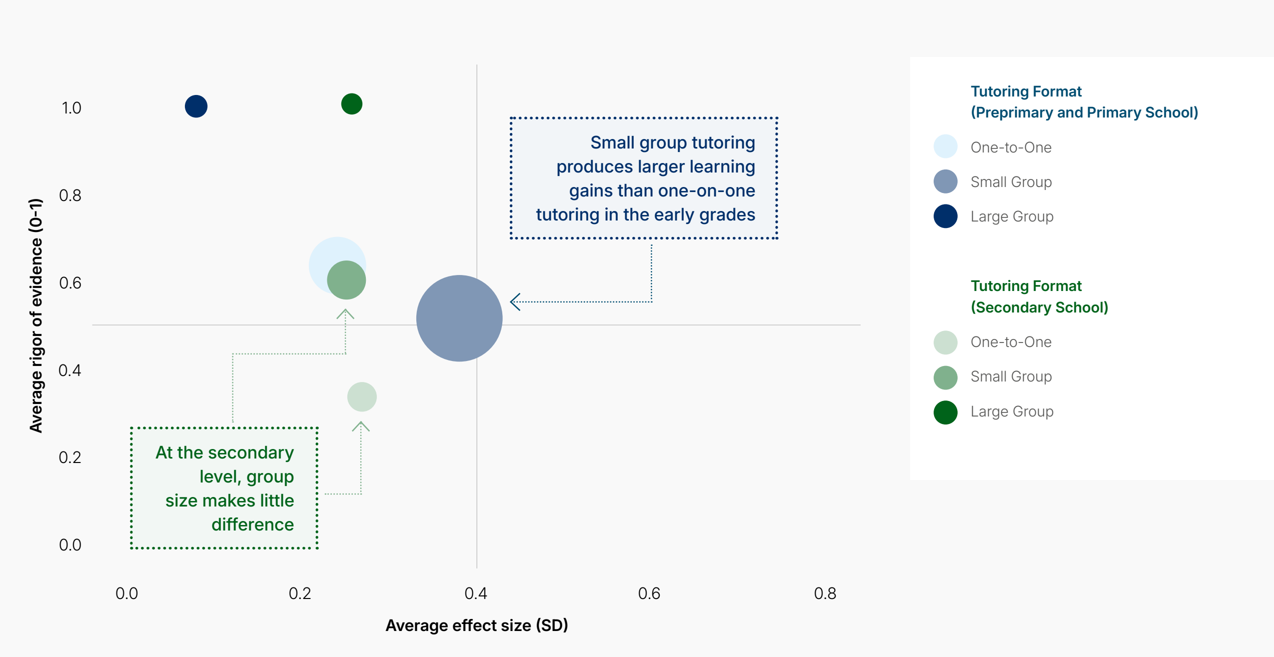
TUTORING ONE-ON-ONE OR IN GROUPS — DOES IT MATTER?

One of the most consequential —and costly— decisions in designing a tutoring program is how many students to assign to each tutor. One-on-one tutoring is intuitively appealing: undivided attention, fully personalized instruction, and immediate feedback. But it is also expensive and difficult to scale. Small group tutoring, by contrast, can reach more students at lower cost — but does it work as well?

■ **Small group tutoring outperforms one-on-one at the primary level — a counterintuitive finding that warrants a careful interpretation.**
Of the 70 studies included in this synthesis, 68 reported information on tutoring format. At the pre-primary and primary levels, small group tutoring, defined as 2 to 6 students per tutor, produces stronger average gains than one-on-one instruction, at 0.37 standard deviations across 37 studies compared to 0.24 standard deviations across 13 studies. This finding is surprising and not fully consistent with other reviews of the tutoring literature, which more commonly find one-on-one instruction to be at least as effective as small group formats. One plausible explanation is that peer interaction — through collaboration, modeling, or even mild competition — adds something to the learning experience that a solo session cannot replicate (see Figure 6.2).

However, the evidence on this mechanism is thin, and it is not clear whether peer interaction in small groups helps or hinders learning. It is also possible that differences in delivery mode, session frequency, or program design across studies contribute to this pattern in ways that are difficult to fully disentangle in a synthesis. Large group tutoring produces negligible gains across 2 high-rigor studies, suggesting that beyond a certain group size — likely somewhere in the range of 5 to 7 students, though the precise threshold will depend on context — the benefits of targeted support begin to diminish. Programs combining one-on-one and small group formats show strong gains of 0.42 standard deviations across 5 studies, suggesting that flexible delivery models may capture the benefits of both approaches.

FIGURE 6.2: TUTORING ONE-ON-ONE OR IN SMALL GROUPS — WHAT WORKS BETTER?
Small group tutoring outperforms one-on-one instruction at the primary level, while group size matters less at secondary level



Note: Group size was classified as one-to-one (1 student), small group (2 to 6 students), and large group (more than 6 students).
Source: Näslund-Hadley et al 2026.

For older students, it barely matters how many are in the room — one-on-one, small group, or large group, the learning gains are nearly the same.
At the secondary level, all three formats produce gains in the range of 0.24 to 0.25 standard deviations — a convergence that may reflect not only the diminishing importance of group size, but also older students' greater capacity to engage productively in group and virtual settings, and the potentially stronger role of peer relations in adolescent learning. For policymakers, the message is encouraging at the secondary level, larger groups can stretch limited resources further without sacrificing learning outcomes.

CAN VIRTUAL TUTORING MATCH FACE-TO-FACE INSTRUCTION?

A major bottleneck for delivering high-quality remedial education and tutoring, particularly in low resource contexts, is the cost and availability of qualified instructors. Structured materials and monitoring systems have made it increasingly possible for non-specialist tutors to deliver effective support. But recruiting, training, and retaining enough tutors, whether paid or volunteers, remains a practical challenge. This has led many education systems to turn to virtual and hybrid models in which students receive academic support partially or fully through digital platforms. The question is whether these models work and whether they are as effective as face-to-face instruction.

In this synthesis, face-to-face interventions refer to remedial education or tutoring delivered through in-person interaction between instructors and students. Virtual interventions refer to support delivered fully online through videoconferencing, phone conversations, educational software, or digital learning platforms, while hybrid interventions combine in-person and virtual forms of instruction.

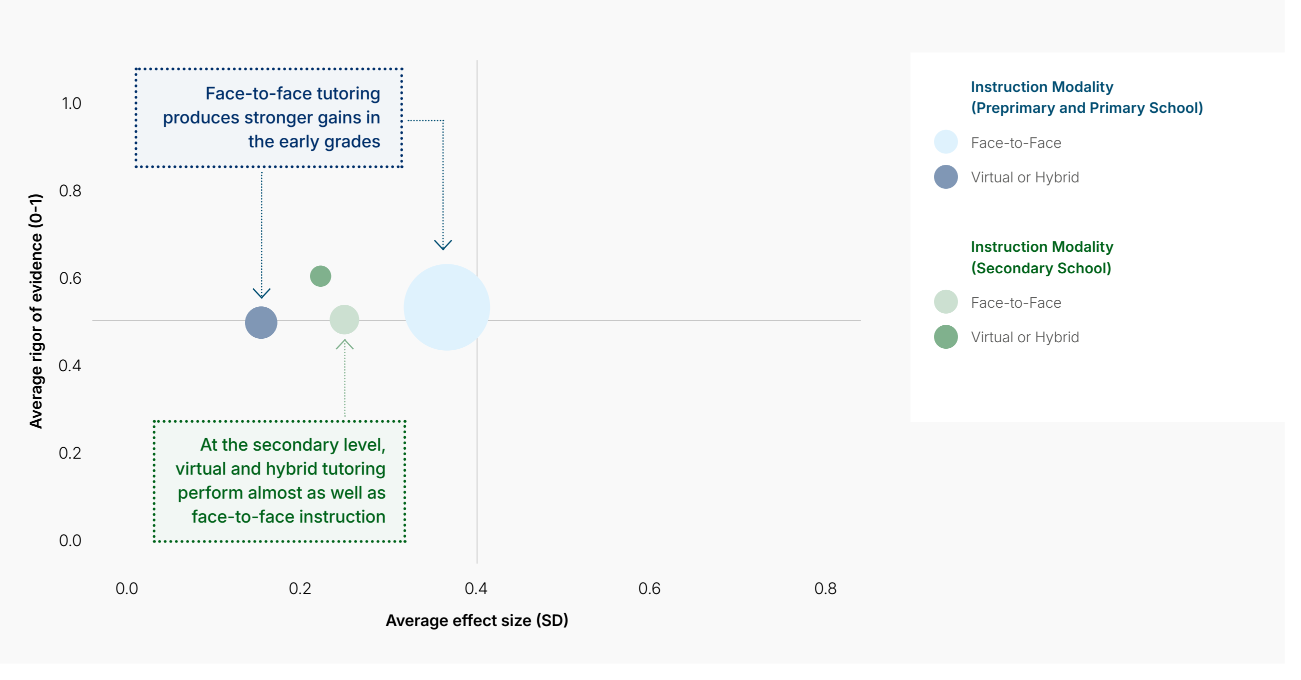
- **The bulk of the evidence favors face-to-face delivery in the early grades.**

Across all 59 studies at the pre-primary and primary levels, in-person tutoring and remedial education generate an average effect of 0.35 standard deviations— a meaningful gain equivalent to roughly 30 weeks of additional learning. Virtual and hybrid approaches, based on 6 studies, produce more modest gains of 0.15 standard deviations at the same level, equivalent to roughly 13 weeks of additional learning. Most studies are of moderate rigor, but the consistency of the findings strengthens the confidence in the overall pattern. The gap is especially pronounced in mathematics, where face-to-face tutoring yields 0.46 standard deviations across 23 studies, compared to just 0.10 for virtual and hybrid formats across 4 studies. In literacy, face-to-face instruction averages 0.26 standard deviations across 23 studies, but with only one study evaluating virtual or hybrid literacy tutoring at the primary level, a meaningful comparison between modalities cannot be made. Expanding the evidence base for virtual literacy tutoring remains an important goal for future research (see Figure 6.3).

- **For secondary students, virtual tutoring is not a compromise — it is a contender.**

At the secondary level, face-to-face instruction averages 0.25 standard deviations, while virtual and hybrid models average 0.22 — a difference too small to be practically meaningful. Older students may be better equipped to engage productively with digital learning environments, and well-designed virtual platforms can support the kind of sustained tutor-student relationship that drives learning gains, even at a distance. In contexts where face-to-face tutoring is difficult to implement — due to geography, cost, or tutor availability —

FIGURE 6.3: IN-PERSON OR VIRTUAL TUTORING: WHAT WORKS BETTER?
Face-to-face tutoring drives stronger gains at the primary level, while the gap narrows considerably at the secondary level



Source: Näslund-Hadley et al 2026.

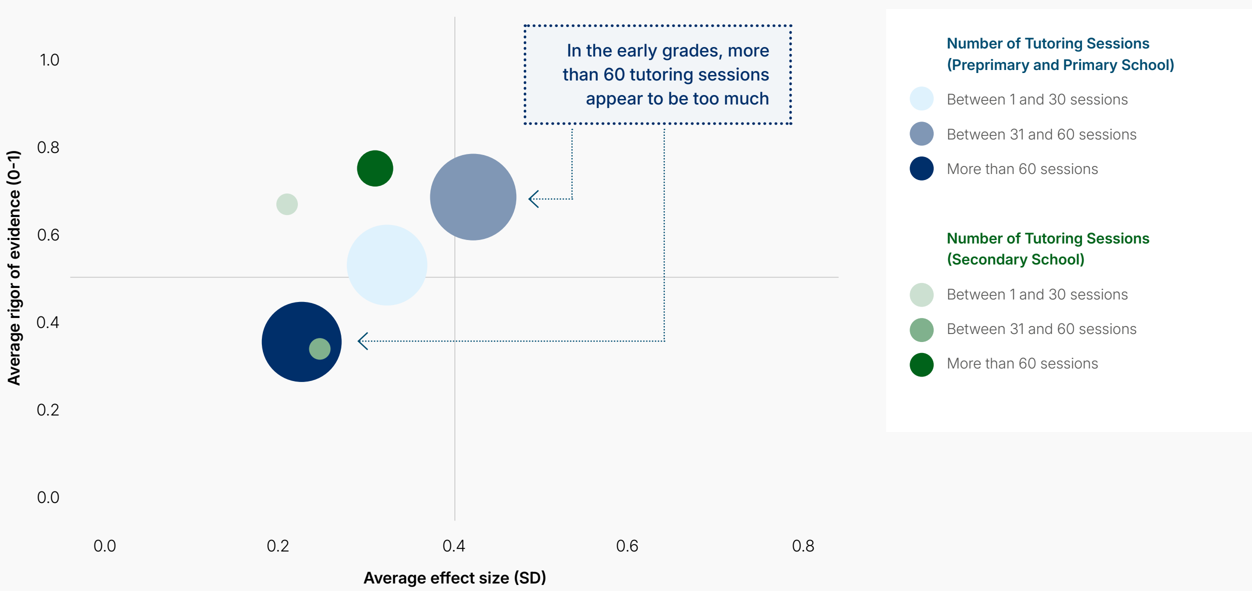
virtual and hybrid models appear to be a genuinely viable alternative. While systematic cost comparisons across studies are not possible given that few report cost data, the lower resource requirements of virtual formats suggest they are likely to be more cost-effective than in-person alternatives at the secondary level, a consideration that matters greatly in lower-income countries like Latin America and the Caribbean. Recent regional evidence is encouraging on this point: a randomized evaluation of phone tutoring across seven Latin American countries (~7,000 students, US$35–109 per student) found average gains of 0.20 standard deviations [5].

DO THE HOURS AND DAYS DEVOTED TO REMEDIAL EDUCATION AND TUTORING HAVE AN IMPACT ON LEARNING GAINS?

Tutoring and remedial programs vary widely in how much support is provided. Most programs have sessions of 30 to 60 minutes, from one day per week to up to five days [6]. One of the most practical questions for any policymaker designing a tutoring or remedial education program is deceptively simple: how much is enough? Should sessions be long or short? How many sessions does a student need before gains become meaningful? Of the 70 studies included in this synthesis, 68 provide information on the number of sessions, and the evidence offers some clear answers — and a few surprises.

■ **When it comes to the number of sessions, the intuition that more is better does not hold up.**
At the primary level, there is a sweet spot. Programs offering between 31 and 60 sessions — in practice, for

FIGURE 6.4: HOW MANY TUTORING SESSIONS ARE ENOUGH?
More sessions are not always better at the primary level, but at secondary, sustained exposure may pay off



Source: Näslund-Hadley et al 2026.

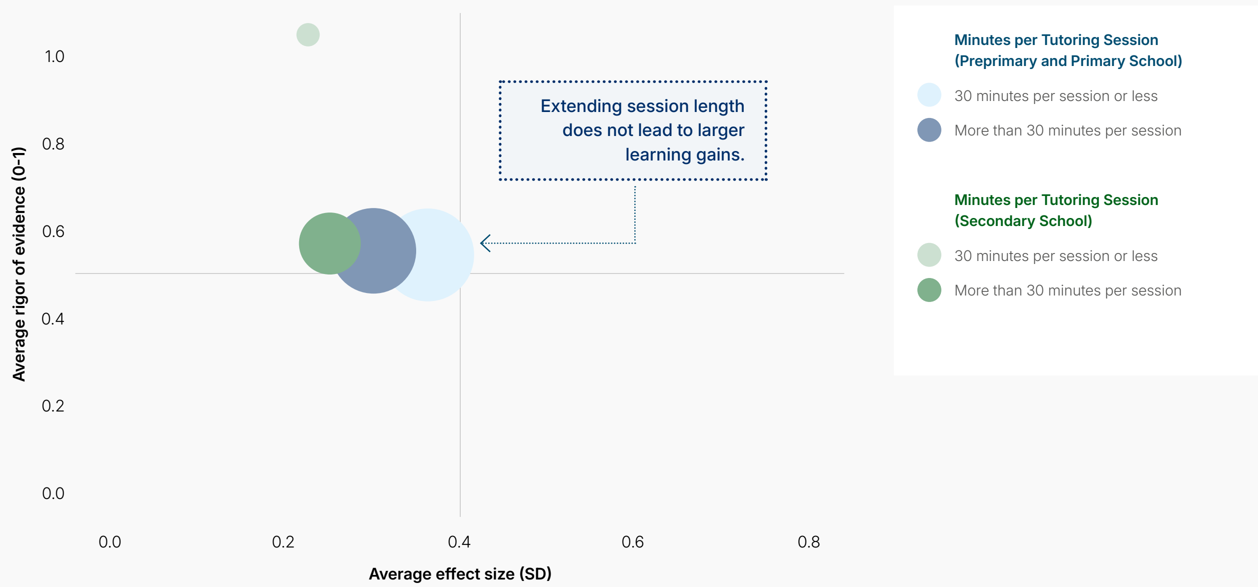
example, two to three sessions per week sustained over one semester to a full school year — produce the strongest gains, averaging 0.41 standard deviations across 22 studies, equivalent to roughly a full academic year of learning (see Figure 6.4). This is consistent with findings in the literature, which find that moderate-dosage programs often outperform longer ones [6]. Below that range, gains are somewhat smaller at 0.33 standard deviations, but still meaningful. Above it, however, returns begin to fade. Programs with more than 60 sessions average 0.23 standard deviations — still an educationally meaningful gain, equivalent to roughly 20 weeks of additional learning, but notably smaller than what shorter programs achieve. This suggests that piling on more sessions does not automatically translate into more learning. A compelling illustration comes from Paraguay, where a phone-based remote tutoring program targeting fourth through sixth-grade students with below-grade-level language skills delivered just 15 sessions over 8 weeks — a compact, low-dose design — and still produced meaningful gains of 0.11 standard deviations at low cost, with 98 percent of treated students completing on average 91 percent of all sessions offered [7]. The program's efficiency is notable. In a context where tutoring is rarely easy to implement, it achieved near-universal participation — 98 percent of treated students took part, completing on average 91 percent of the sessions offered — consistent engagement, and positive learning outcomes within a compact and replicable model.

The message for policymakers is clear: investing in longer sessions is unlikely to pay off. For program designers working within tight time and resource constraints, this is an encouraging finding — effective tutoring does not require long sessions, and shorter models can be just as powerful.

■ **At the secondary level, more tutoring appears to mean more learning.**
At this level, gains grow steadily as the number of sessions increases: from 0.21 standard deviations for programs with 30 sessions or fewer — roughly a semester of weekly or twice-weekly support — to 0.30 for those offering more than 60, the equivalent of two to three sessions per week across a full school year. The pattern suggests that older students benefit from more sustained support. The evidence for the secondary level is still limited, there are just 3 to 4 high-rigor studies per group, so these findings should be read with caution. But the overall message is consistent: for primary students, finding the right dose of educational support matters as much as providing more; for secondary students, staying the course may be worth it.

■ **When it comes to session length, 68 of the 70 studies provide valuable information on this dimension — and the findings are consistent: shorter sessions are just as effective as longer ones.**
At the pre-primary and primary levels, tutoring sessions of 30 minutes or less produce average gains of 0.35 standard deviations across 32 studies —virtually identical to the 0.31 standard deviations generated by longer sessions across 26 studies. At the secondary level, the story is the same: shorter and longer sessions perform almost equally, at 0.24 and 0.26 standard deviations respectively, and while there are fewer studies at this level, the high methodological rigor of the studies allows for a high confidence in their findings. The message for policymakers is clear: investing in longer sessions is unlikely to pay off. For program designers working within tight time and resource constraints, this is an encouraging finding — effective tutoring does not require long sessions, and shorter models can be just as powerful (see Figure 6.5).

FIGURE 6.5: DOES SESSION LENGTH MATTER FOR TUTORING EFFECTIVENESS?
At both the primary and secondary levels, shorter and longer tutoring sessions produce comparable learning gains



Source: Näslund-Hadley et al 2026.

TEACHERS OR OUTSIDE TUTORS - WHO IS BETTER FOR REMEDIAL EDUCATION AND TUTORING?

One of the most consequential decisions in designing educational support programs is who delivers that help, and at what cost. Should it be a teacher from the child's school, such as a special education teacher, who knows the curriculum and the child? Or an outside tutor — a volunteer, a college student, or a contracted coach — who may bring fresh energy and a different kind of relationship?

At the primary level, outside tutors can be just as effective as teachers when they are well prepared and well supported.

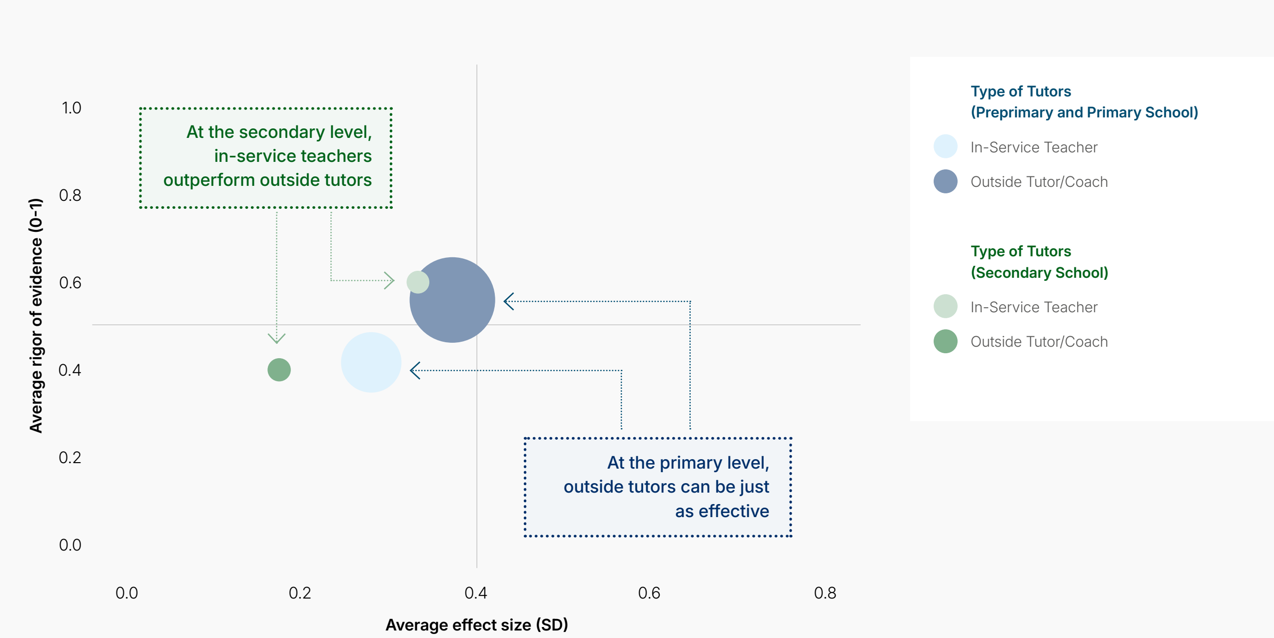
All studies in this synthesis provide information on who delivers the tutoring support. At the pre-primary and primary levels, programs delivered by outside tutors or coaches are associated with stronger gains than those delivered by in-service teachers: 0.36 standard deviations across 41 studies compared to 0.27 standard deviations across 17 studies. This is a difference equivalent to roughly eight weeks of additional learning, and it holds across a large and relatively rigorous evidence base.

Nevertheless, this finding is not universal. Other large reviews of the tutoring literature find teachers and paraprofessionals outperforming outside tutors [8] [9], though these cover a time period going back four decades and may not reflect what is possible today with structured materials, digital platforms, and better-trained non-specialist tutors. A well-documented example is the Future Forward literacy program in the United States, where paraprofessional tutors delivered one-on-one reading support to struggling early-grade students in low-income schools and backed by strong materials, training, and supervision, produced an average effect of 0.23 standard deviations on foundational literacy skills across multiple randomized controlled trials [10]. This suggests that tutors that are not trained teachers can achieve substantial gains when the conditions for success are in place.

At the secondary level, subject knowledge may matter more than the tutor-student relationship.
At this educational level, however, the pattern reverses: in-service teachers are associated with stronger gains (0.32 standard deviations across 5 studies) than outside tutors (0.17 standard deviations across 5 studies). This reversal at the secondary level carries a clear policy implication. As mathematical and literacy content becomes more complex, the subject expertise that in-service teachers bring may become more important than the personalized attention and relational warmth that outside tutors tend to offer. For policymakers, this implies that the right choice of tutor depends heavily on the age and level of the students being tutored — and that a one-size-fits-all approach to tutor recruitment is unlikely to be optimal (see Figure 6.6).

How tutors are trained matters as much as who they are.
What tutors learn in an initial workshop is a starting point, not a finished skill. The real learning happens in the room, and ongoing support gives tutors the chance to reflect on what they observe and to adjust accordingly. Training is, ultimately, a proxy for what matters most: the quality of the tutoring interaction itself. Of the 52 studies that report on tutor training, programs combining in-person workshops with follow-up and refresher sessions

FIGURE 6.6: WHO SHOULD DELIVER TUTORING — TEACHERS OR OUTSIDE TUTORS?
Outside tutors drive stronger somewhat stronger gains at the primary level, but in-service teachers pull ahead at secondary



Source: Näslund-Hadley et al 2026.

produce substantially stronger gains, 0.47 standard deviations, compared to 0.28 standard deviations for programs relying on initial training alone. The message is clear: one-off training is not enough, and investment in continuous support — through monitoring tools, structured feedback, and communities of learning— leads to stronger learning outcomes.

KEY FINDINGS AND WHAT THEY CONTRIBUTE TO POLICYMAKING

Tutoring and remedial education are not a single intervention, but a set of design choices, each of which shapes how much students learn and who benefits most. The findings, summarized in Figure 6.7, do not point to one universal model, but they offer a consistent and actionable picture for policymakers seeking to close learning gaps.

■ **Know what you are trying to do, and design for it.**
Programs that make a clear curricular choice, either fully grade-aligned or explicitly remedial, consistently outperform those that try to do both at once. Both approaches work well at the primary level; at secondary, grade-aligned tutoring pulls ahead.

■ **Small groups stretch resources without sacrificing learning.**
Small-group tutoring yields gains comparable to one-on-one instruction at a fraction of the cost, but once group size exceeds six students, the benefits of targeted support begin to disappear.

■ **Face-to-face instruction remains the best option in early education — but virtual models are a viable alternative at the secondary level.**
Given that virtual formats typically cost less and can reach more students, they represent a genuinely cost-effective option for secondary programs, particularly in contexts where qualified tutors are scarce or geographically dispersed.

■ **More is not always better — calibrate the dose.**
At the primary level, gains peak between 31 and 60 sessions and fade beyond that, while session length makes little difference at any level: invest in the right number of well-structured sessions, not longer ones.

■ **Who delivers the tutoring matters, but it depends on students' age.**
Well-supported outside tutors — paraprofessionals, volunteers, college students — are highly effective at the primary level, while at secondary, subject expertise tips the balance toward in-service teachers.

■ **Training tutors once is not enough.**
Programs combining initial training with follow-up and refresher sessions substantially outperform one-off preparation — one of the most actionable levers available to program designers.

■ **Tutoring is promising in Latin America and the Caribbean, but more regional evidence is urgently needed.**
Effects in low- and middle-income countries are more modest, tend to shrink as programs scale, and regional evidence remains thin. Expanding the evaluation of tutoring programs across the region is as urgent as expanding the programs themselves.

FIGURE 6.7: SO, HOW TO DESIGN TUTORING AND REMEDIAL EDUCATION?
Instructional components are associated with different learning outcomes across educational levels

Instructional approach	Strongest learning outcomes	Key implication for policy and practice
WHERE TO TEACH		
Grade-aligned tutoring	Learning gains at both primary and secondary levels	Keep students on pace with the curriculum — clear curricular focus outperforms mixed approaches
Remedial foundational programs	Primary level learning gains	Teach students at their actual level, not their enrolled grade — foundational catch-up works best in early grades
HOW TO DELIVER		
Small group tutoring	Primary level learning gains	Small groups retain most of the effect at a fraction of the cost — avoid groups larger than six
Face-to-face instruction	Primary level gains, particularly in mathematics	Face-to-face delivery is stronger at primary, particularly for math and struggling learners — virtual is viable at secondary and where cost and reach are priorities
Virtual and hybrid models	Secondary level learning gains	Cost-effective alternative at the secondary level where the gap with face-to-face narrows to insignificance
HOW MUCH		
31–60 tutoring sessions	Primary level learning gains	More sessions are not always better — returns begin to fade beyond 60 at the primary level
Higher session count at secondary	Secondary level learning gains	At the secondary level, more tutoring sessions consistently produce larger gains — invest in program duration
THE TUTOR: WHO AND HOW		
Outside tutors	Primary level learning gains	Outside tutors can be highly effective when well trained and supported
In-service teachers	Secondary level learning gains	Subject expertise matters more at secondary — qualified teachers outperform outside tutors
Ongoing tutor training	Learning gains across levels	One-off training is not enough — follow-up and refresher sessions substantially improve tutor effectiveness

Source: Developed by authors based on the synthesis of 70 tutoring and remedial education impact evaluations.

REFERENCES

[1] Zoido, P., Albornoz, F., Prieto, A., Almeyda, G., Hernández, J.C., Oubiña, V. & Calderón, M. Multiplying Learning: Remote Tutoring to Enhance Schooling. 2023. Washington, DC: Inter-American Development Bank.

[2] Näslund-Hadley, E., Alonzo, H., Grossi, A., & Bando, R. What Boosts Learning Recovery? Global Evidence from a Systematic Review of Tutoring and Remedial Education Interventions. 2026. Washington, DC: Inter-American Development Bank.

[3] Kraft, M.A., Interpreting effect sizes of education interventions. *Educational Researcher,* 2020, 49(4): p. 241–253. doi.org/10.3102/0013189X209127

[4] Berlinski, S., Busso, M., and Giannola, M. *Helping Struggling Students and Benefiting All: Peer Effects in Primary Education.* IDB Working Paper No. IDB-WP-1338. 2022. Washington, DC: Inter-American Development Bank. doi.org/10.18235/0004268

[5] Reference: Albornoz, F., Almeyda Torres, G., Lombardi, M., Oubiña, V., & Zoido, P. (2026). Remote tutoring in Latin America. Journal of Development Economics, 179, 103687. https://doi.org/10.1016/j.jdeveco.2025.103687

[6] Kraft, M.A. and Falken, G.T. A Blueprint for Scaling Tutoring and Mentoring Across Public Schools. *AERA Open,* 2021, 7(1): p. 1–21. doi/10.1177/23328584211042858

[7] Acevedo Rebolledo, N., Almeyda Torres, G., Granada Donato, D., Lombardi, M., Oubiña, V. and Zoido, P. Remote Tutoring in Language: Evidence from Paraguay. IDB Working Paper No. IDB-WP-01798. 2026. Washington, DC: Inter-American Development Bank. doi.org/10.18235/0013975

[8] Nickow, A., Oreopoulos, P. & Quan, V. The Promise of Tutoring for PreK–12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence. American Educational Research Journal, 2024, 61(1): p. 74–107. https://doi.org/10.3102/0002831223120868

[9] Kraft, M.A., Schueler, B.E. & Falken, G. What Impacts Should We Expect from Tutoring at Scale? Exploring Meta-Analytic Generalizability. EdWorkingPaper No. 24-1031. 2024. Providence, RI: Annenberg Institute at Brown University. https://doi.org/10.26300/zygj-m525

[10] Jones, C.J. and Christian, M. The Results of a Randomized Control Trial Evaluation of the SPARK Literacy Program: An Innovative Approach that Pairs One-on-One Tutoring with Family Engagement. Journal of Education for Students Placed at Risk (JESPAR), 2021, 26(3): p. 185–209. doi.org/10.1080/10824669.2020.1809419

Part III

From evidence to Scale

KNOWING WHAT WORKS IS NOT ENOUGH. SYSTEMS HAVE TO DELIVER IT, EVERYWHERE, EVERY DAY.

CHAPTER

7

THE SCIENTIST, THE ENGINEER, AND THE PLUMBER: WHAT LITERACY PROGRAMS TEACH US ABOUT DESIGN, THE REALITY ON THE GROUND, AND PEOPLE WHO MAKE CHANGE POSSIBLE

Ximena Dueñas & Silvana Godoy Mateus

Decades of investment in foundational skills programs have not translated into better learning outcomes. An alarming share of children in Latin America and the Caribbean fails to reach minimum levels of reading proficiency by the end of primary school, and these gaps accumulate from the earliest grades (see Chapter 2).

Programs backed by strong evidence often stall because implementation and scale-up are complex. Chapters 4, 5 and 6 illustrate this point clearly. In the systematic reviews of mathematics, literacy and tutoring, only a third of the hundreds of studies analyzed evaluated interventions implemented at large scale. And on average, those interventions at scale had effects on learning close to half those achieved by small programs (those reaching 400 students or fewer). In other words, as they scale up, programs tend to lose a good part of their impact. The economist John A. List has described this phenomenon as the "voltage effect": the tendency of interventions to lose power, in part because their implementation has received little attention [1].

This chapter focuses on the details of the implementation process: the gap between a validated design and what is ultimately delivered in schools. Even when the intervention remains the same, it is no longer the same pilot, the same schools, or the same people. Once outside the controlled environment in which evidence was generated, the program is truly put to the test.

Theory and evidence offer limited guidance on the implementation details that shape results, such as how processes are organized, change is communicated, and how the people doing the work actually use it. In the words of Nobel Laureate in Economics Esther Duflo, "as those designs are implemented in the real world, a new responsibility arises, which is to attend to numerous details on which our models and theories offer little guidance" [2].

Our analysis examines implementation through the lens of the barriers and the facilitators of this process, drawing on implementation science to understand how evidence-based interventions can be adopted, delivered with fidelity, implemented with quality, and sustained in real settings [3] [4].

To apply this focus to the Latin American context, we conducted a qualitative inquiry based on seventeen interviews with coordinators from fourteen organizations in six countries: Argentina, Brazil, Chile, Colombia, Panama, and Venezuela. The cases were chosen through purposive sampling: not at random to represent a population, but by deliberately seeking the widest possible diversity of experiences [5]. The sample brings together programs that differ from one another, with different scales and degrees of maturity, varied sources of funding, resources and technical capacity, and under different institutional configurations (governments, foundations, private and community initiatives). This heterogeneity is intentional: to identify whether the same pattern recurs despite dissimilar contexts. The selection criterion is not to have a representative sample, but that each case has the capacity to provide in-depth insight into the implementation process.

The evidence on implementation processes is more developed in the field of literacy than in other areas, which makes it especially fertile ground for this inquiry. The lessons that emerge, however, apply more broadly to the implementation of foundational skills programs. Rather than estimating prevalence, the findings indicate where to look, what to verify, and what to anticipate when implementing literacy programs. A description of the analytical framework and the full methodology is in online Methodological Annex.

IMPLEMENTATION HAS A SCIENCE OF ITS OWN

■ **Implementation science asks how to ensure that what has worked can be adopted, used, and sustained in routine practice.** This leads it to examine three issues that impact assessments often do not capture.

First, it analyzes the conditions that enable or hinder implementation: the barriers and facilitators that explain whether a program is adopted in practice [6]. This approach is the development framework of this study. Second, it assesses a set of outcomes specific to implementation: whether the program is acceptable to those who have to use it, whether it is actually adopted, whether fidelity is possible, if it is feasible and sustainable over time, and at what cost [7]. These implementation outcomes can affect whether a foundational skills program can be scaled successfully. And third, it focuses on processes and mechanisms: how is the program executed? What is adapted along the way? Why does the same design perform differently from place to place? In short, the aim is to examine the path from evidence to practice: where a proven pilot program is sustained at scale, and why it might lose traction.

■ **Evidence on how to implement effective programs in schools remains scarce.**
Duflo argues that good public policy requires three kinds of symbolic actors. Through rigorous evidence, **the scientist** establishes what works. **The engineer** takes those findings and designs the program. And **the plumber** installs that program in the real world, with its broken pipes, its unexpected pressures, joints that don't fit, and adjusts everything it until it works [2].

■ **The region has invested heavily in discovering what works and designing programs around that evidence. What it has often overlooked is the plumbing: the work of making those programs function in schools.** Implementation science helps make visible the human, institutional, and contextual processes that sustain or erode results when a program is implemented in practice [8]. Throughout this chapter, we analyze implementation experiences and identify the barriers and facilitators, observing where they appear and what patterns they follow. This makes it possible to take the first step toward improving how learning programs are implemented.

■ **Effective foundational skills programs share five enabling conditions.**
By an early literacy program we mean an intervention oriented to the teaching of reading and writing, grounded in the science of reading, which promotes explicit, systematic and progressive instruction [9], and which intentionally articulates five pillars in its design: (i) political and technical commitment, (ii) materials, (iii) teacher training, (iv) situated coaching — continuous pedagogical support to teachers in their own classroom, adapted to their teaching context (observation, feedback and modeling) — and (v) assessment for learning [10 to 13].

- **Pillar 1. Political and technical commitment.**
  The key lies in coordinating multiple actors (from the school to the Ministries of Education) with different levels of agency and power. Clear objectives aligned with the science of reading are also required. Strong political and technical commitment translates into multisectoral coalitions that protect the budget, ensure accountability, and prevent the program from collapsing or being discontinued when the central government changes.

- **Pillar 2. Materials.**
  The materials that reach the classroom must be grounded in the science of reading and easy to use, to reduce the teacher's planning burden and guide daily practice, such as: structured lesson guides and an explicit teaching sequence for the teacher; levelled or decodable readings, and workbooks for systematic student practice. But success does not depend on design alone, but on equitable and timely distribution: good materials that do not arrive on time, or reach only some schools, cease to be a learning tool.

- **Pillar 3. Teacher training.**
  This is a change management strategy: it seeks to shift beliefs and classroom practices that are deeply rooted in everyday school culture. To ensure an effective transition towards explicit and systematic instruction, the design requires intensive initial training — in person and before the start of the school year — complemented by periodic refreshers during the year and peer support: regular spaces where teachers applying the program resolve difficulties and share what works in their classrooms (communities of practice).

- **Pillar 4. Coaching in the classroom.**
  This is the pillar responsible for protecting implementation quality and fidelity; in other words, it is the plumbing work. It centers on continuous pedagogical support to teachers rather than merely administrative support, allowing early detection of deviations and iterative adjustments based on modeling and classroom observation.

- **Pillar 5. Assessments for continuous learning.**
  This pillar consists of appraisals that generate actionable information for teachers. Their function is to identify students falling behind and to enable data-based decision-making for continuous improvement.

For each pillar, the study organizes the findings into five domains, following the path a program travels from paper to the classroom.
The first domain is the design of the program, or the engineer's work in the terms used by the economist Esther Duflo. The four that follow concern the realities in which the plumber operates: the external context (what lies outside the program and beyond the control of the design), the school context (the conditions of the school where it is applied), individuals (the teachers and coaches who filter the intervention through their experiences and values), and the process (how the program is executed over time). The heat map in online Methodological Annex synthesizes this reading.

The analysis reveals three important points. First, the problem is rarely the design itself. The challenge begins when programs reach schools and classrooms, where teachers and school conditions live realities no design can fully anticipate. Second, these difficulties are not always unpredictable: the same obstacles appear repeatedly in the same pillars, in different countries and in different programs. Third, knowing these patterns in advance allows implementers and governments to anticipate barriers instead of improvising when it is already too late.

WHERE ARE THE BARRIERS?

In the programs we analyzed, the barriers lie not in the design, but in implementation realities and in the people who bring the program to life.
"This business of the design being so good and then in practice not managing to translate into reality... that it doesn't become real, I think that's a great frustration", said a from the sample of programs. This is no minor finding. The fourteen cases studied represent programs with designs based on the science of reading and at different stages of maturity. If even these programs face systematic barriers in implementation, the problem seems not to be the design itself, but the conditions the design cannot control.

The challenge begins before implementation itself: many teachers enter the classroom without adequate preparation to teach reading.
None of the implementing units that reported information referred to initial university training as a facilitator: "There are teachers from the best universities (...) who were never properly instructed in teaching reading and writing", says an expert from an education foundation. Universities and teacher training colleges tend to privilege pedagogical theories and general approaches to education, but do not consistently prepare teachers to teach reading using evidence-based approaches, as documented in detail in Chapter 5. Each program has to remedy that gap on its own, at scale and indefinitely.

Added to this is teacher turnover of systemic origin—transfers, temporary appointments, and movement between territories and schools for personal and administrative reasons. On the administrative side, within the school itself, principals often move the most experienced and already-trained teachers to higher grades in order to improve standardized test scores, breaking the continuity and specialization of the first-through-third-grade cycle. Each departure from the school or shift between cycles dissolves the knowledge the program managed to embed, and the accumulated technical investment leaves with the person.

Even when a program begins with political support, that backing may not survive the next electoral cycle.
Most of the programs analyzed plan joint work with principals, ministries, secretariats, and local authorities. Yet that political architecture often works only until the next change of government. "We survived the change of administration... there was a lot of resistance, because deep down there was this thing of that's the previous lot's plan and I want my

own", says a person from the sample. What one government builds, the next can dismantle. No program design can solve this without institutional mechanisms that protect continuity beyond electoral cycles.

■ **When a program survives a change of government, the next obstacle is more prosaic: books and supplies may not arrive on time.**

The implementation units in the sample report materials of good technical and pedagogical quality. The problem lies in how they are produced, distributed and received in the classroom: "The common mistake is to think the whole program is in the books… [and then] the materials never arrive on time", stresses a person from the sample.

The most important barrier in the materials pillar is the logistics chain: the remoteness of many rural schools and, in some cases, situations of armed conflict and insecurity make distribution a dangerous and costly operation: at times, distribution costs exceed the value of the book. Add to that barrier a recurrent operational difficulty: a lack of contractual anticipation. Public procurement timelines often delay the delivery of materials until several months after the school year has begun: "The whole process starts late. It's not that it starts well and [the materials] arrive late", laments a person from the sample. The solution is as much contractual as logistical, and multi-year purchasing and greater local autonomy to anticipate production would help overcome this barrier.

■ **When materials do arrive, teachers need support to use them. Yet that rarely arrives with the frequency and depth promised in the design.**

Unlike the other four pillars, coaching already shows weaknesses at the planning stage. In the first iterations, compliance is verified through answers to questions such as: did the material arrive? Did the teacher attend? But the key is also the pedagogical support that gives teachers useful feedback to improve their classroom practice. Theoretical frameworks say one thing; the operational structure does another: "Monitoring, reporting, follow-up and coaching systems… have to be very well thought through from the design stage and I think those are the things we sometimes don't prioritize enough", says a person from the sample.

This weakness is compounded by the region's implementation realities. Geography and distance systematically interrupt the frequency of in-person coaching in remote schools. Digital connectivity makes it possible to sustain part of the follow-up remotely. However, when these solutions are not enough, programs are forced to reduce the intensity of coaching envisaged in the design and to delegate greater responsibility to teachers to implement the program autonomously.

■ **Even where a culture of assessment is in place, results arrive late, breaking the last link in the learning cycle.**

"It was our Achilles heel for many years; teachers would say, when are you going to tell us how we did? Besides, if you assessed me at the end of the year, I no longer have those children next year", says a person from the sample. Teachers are receptive when assessment is not linked to appraisal of their performance. The challenge lies in the feedback: results arrive late, are tallied administratively, or do not arrive in a shape the teacher can use to adjust their teaching.

The ten subthemes classified as barriers share one feature: none is in the design domain. They are unevenly distributed: six in the context conditions, two in the people who have to adopt the change, one in the school, and one in execution. Literacy programs face structural, organizational, and relational conditions that affect how they are put into operation. That is why the chapter now turns to what policy and management can influence: the plumber's work of anticipating, adapting to, and managing those barriers.

WHY IS THE HUMAN BARRIER THE HARDEST TO OVERCOME?

The most persistent obstacles are neither logistical nor political, but relational and cultural. Teacher resistance is not an expression of indifference: teachers care about their students' learning. But they are uncomfortable moving away from familiar practice, from the way of doing things that gave them confidence. There is also learned distrust in settings that have repeatedly seen pedagogical interventions arrive, disrupt practice, and disappear within months. Reversing this distrust requires consistency, respect for teachers' professional trajectories, and evidence of tangible results among their own students.

Trust does not start with the teacher, but with the school principal.
"It is not by starting with teachers that the operation works, but by starting with the actors who make the decisions", stresses an expert from another foundation. Effective programs work first with principals, supervisors and local authorities to build an institutional commitment before entering the classroom. "A program... cannot get going if there is no consensus about the urgency of addressing the problem... that technical consensus with schools can take us up to six weeks", says an official from a foundation. This time devoted to coordination is not a delay; it's an investment.

Once in the classroom, change is not imposed but is built upon what teachers already know.
"Training has to empathize with the teacher; it has to recognize the teaching work they have been doing... use all the resources that the teacher has and add something more to their baggage but not wipe out everything they have done", emphasizes a person from the sample. Building openness to change requires a precise stance: honoring the knowledge teachers already have, not dismissing it, and accompanying them from there. To address the initial resistance rooted in their professional identity, the most effective programs have learned that changes in practice are sustained mainly through on-site coaching. Acceptance usually emerges only after the program has accumulated enough experience in teacher training and in the science of reading to shift the focus toward practice and modeling. This is a work challenge that only mature implementation teams can do well. The most powerful facilitator is the moment teachers see learning happen: "The moment when the teachers realize that the children are learning, it's as if a barrier breaks", says a person from the sample. That moment of visible progress can turn teacher resistance into conviction, something training alone rarely achieves.

The coach —a trainer responsible for building the teacher's capacity in their own classroom— must be a peer, not an inspector.
In the words of a researcher from a think tank, "it's a relationship of trust, of great respect, very professional, and of important complementarity". The coach must be perceived as a technical peer who supports teachers' work, reducing the anxiety of being observed and building a long-term relationship of trust. This is a process of building trust through technical credibility.

Teachers' resistance to programs that affect their work is not overcome in a training session, but when teachers see their students learn.
Timely feedback of results allows the teacher to regroup children or adjust their instruction: "They saw in the possibility of assessment something very powerful for knowing how their children are doing... they build trust through assessment and results", says a person from the sample. Faced with the challenge of getting children to learn, situated modeling works directly on the real difficulty of executing what the program proposes. In its most demanding form, the coach takes the risk of teaching the class in front of the teacher, with their own students and under their real conditions, demonstrating that the strategy works in that context. It follows the logic of gradual

release of responsibility: “I do it, we do it together, you do it”, explains a person from the sample. By the end of the cycle, the teacher leads, and the coach observes.

Quality materials are not only pedagogical inputs; they are a lever of trust.
Quality materials are not only pedagogical inputs; they are a motivational lever. Teachers accept the program more readily when they see that the material is useful for students and comes with structured planning that reduces their own lesson preparation burden. “We have a wonderful book… you get less annoyed with someone who gave you a book that makes your students beg you to read them literature”, says a person from the sample.

Trust requires time, presence and sustained relationships. Implementation science has documented this: personal readiness for change depends on having the capability, the opportunity and the motivation to modify behavior [14] [15]; organizational readiness is even more complex, since it requires individuals to feel committed and confident in their collective capacity to transform their practices [8].

WHAT DO EFFECTIVE PROGRAMS DO WHEN THE CONTEXT IS AGAINST THEM?

Resilient programs do not improvise; they build repertoires of responses.
No program unfolds exactly as designed once it reaches schools. The literature confirms that operational adjustments of this kind are precisely what make the difference. When Pratham’s well-known remedial education program based on grouping children by actual learning level rather than grade, initially failed at scale, the problem lay not in the pedagogical design, but in the fact that nothing in the teacher’s daily life had really changed with that program. Nor was the solution pedagogical: when a dedicated hour and a coaching supervisor were added, results reversed [2].

When books and supplies do not arrive, effective programs build alternative distribution networks and keep teaching.
They turn to local carriers, parcel services, networks from other programs, regional hubs and even the personal vehicles of supervisors or principals to get books to hard-to-reach areas. If the difficulty persists, they resort to pedagogical compensation strategies through digital and multimodal files to avoid interrupting instruction. In emergencies or extreme contexts, they apply the “Escuela Solidaria” (Solidarity School) model: children copy content into their notebooks so that a single set of books can serve several cohorts or classrooms. Depending on the accessibility of the terrain, they organize differentiated coaching routes (full, express or autonomous) and, in remote areas or those affected by conflict, they use radio capsules and sustain learning with offline resources. A person from the sample explains one of these alternatives: “(We send) audio files that, through songs and sounds, practice each of the lessons… it’s a pen drive with the 60 audio files and the printed manuals, so they can use it”.

When teachers leave, resilient programs institutionalize the curriculum instead of depending on individuals.
They promote learning plans that formalize it in the classroom and agreements with principals to form stable teaching teams from first through third grade.

In every pillar there is a degree of freedom about how to operate, depending on the context and its particularities. However, no adaptation replaces action on the underlying structural conditions, nor should it extend to the technical and pedagogical core that gives the model its meaning.

HOW CAN THE GAP BETWEEN DESIGN AND IMPLEMENTATION BE CLOSED?

The question is not whether a program will lose power as it moves from design to reality, but where, how much, and why.

Under the plumber's gaze proposed by Duflo, what shapes final impact, beyond the soundness of the design, is attention to implementation details: the points where the machine jams as it starts up. The ten barriers identified are not random failures; they are the predictable places where power escapes when a proven program leaves the controlled conditions in which its evidence was generated. Closing that gap requires coordinated action on two fronts: public policy, which can modify structural conditions, and implementation practice, which learns to operate within and around them. Figure 7.1 organizes the ten barriers by pillar and indicates, for each one, who can resolve it and with which levers.

What distinguishes resilient programs is their ability to adapt continuously.

The cases that sustain the pillar despite difficult implementation conditions deliberately build a repertoire of responses, but this has a limit. The study identifies a non-negotiable core that no implementation should sacrifice: the technical sequence of explicit, systematic and progressive instruction; tangible materials of good technical quality; and situated coaching delivered by coaches with high relational and technical capacity. The practical rule is to protect the core and allow flexibility at the operational margins, not the other way around. Since two of those three elements depend directly on individuals, this points to the need for a people-centered design [16]. Through observation and active listening, this approach emphasizes understanding people well; it recognizes that what happens with teachers cannot be assumed, but must be observed, understood, and acted upon.

Studying implementation adds two valuable elements to the traditional impact assessment.

First, it brings understanding and depth to the processes and reveals details that the numbers do not capture. And second, it helps to identify which variables are worth observing and measuring in future impact assessments and their effect on learning outcomes (the quality and timeliness of the arrival of materials, or the intensity and nature of coaching). These elements are seldom measured, despite being entirely quantifiable.

Three takeaway messages emerge.

1

Implementation is too complex to be treated solely as a budgetary, technical, or administrative task. Its structural, social, and cultural dimensions must move to the center of the policy debate.

2

People, not only evidence or inputs, must be at the center of program design.

3

The region needs more **plumbers**: professionals willing to look closely at how implementation unfolds, identify bottlenecks, and make the practical adjustments that no pilot, however rigorous, can anticipate.

FIGURE 7.1: THE TEN BARRIERS TO IMPLEMENTATION AND WHO CAN RESOLVE THEM
Organized by the five pillars of an effective literacy program

BARRIER	WHAT GOVERNMENTS CAN DO	WHAT IMPLEMENTERS CAN DO
PILLAR 1. POLITICAL AND TECHNICAL WILL		
Sustainability of political will: The program starts with political support. That support does not survive the elections.	Institutional mechanisms that protect continuity beyond electoral cycles.	Roadmaps and handbooks for ministers that institutionalize the program as state policy.
PILLAR 2. MATERIALS		
Contractual rigidity: the public contracting and financing framework is slow and complex.	Reform the contracting and financing framework: multi-year purchasing, greater contractual autonomy for provinces, departments and municipalities, and clear procurement routes for supplies.	Advocate to simplify processes and build contracting timelines into design and planning assumptions.
Contractual anticipation: even under the current framework, failing to start the operational chain on time means materials arrive months after the school year has begun.	Start contracting processes in advance and align the purchasing calendar with the school year, so that materials are in place on the first day of class.	Anticipate the operational chain: plan for a first year of readiness; print the first unit and fall back on PDFs or WhatsApp scripts while the tender proceeds; in extreme contexts, "Escuela Solidaria" to share a single set of books.
Logistics chain and geography: rural areas and zones that are hard to reach or affected by conflict make distribution more expensive.	Fund last-mile distribution and strengthen supply chains to rural and hard-to-reach areas.	Alternative distribution networks: local carriers, parcel services, regional hubs and even the vehicles of supervisors or principals.
Ideological resistance to the approach of the material: the teacher uses the book without sharing its foundations.	Promote curricular alignment and dissemination with the materials distributed, to avoid contradictions of approach.	Teach how to use the material. Build conviction with evidence (use of data) to show that there is a link between material and learning.
PILLAR 3. TEACHER TRAINING		
College training does not prepare teachers to teach reading with evidence-based approaches.	Reform university and teacher training college curricula.	The program supplies that missing training, at scale and on a permanent basis.
Teacher turnover across contexts.	Monetary incentive programs to minimize turnover.	Rapid induction protocols for new teachers; materials and self-training routes to sustain continuity.
Teacher turnover within the school: trained teachers are moved to higher grades.	Internal school organization policies that protect the retention of trained teachers within the literacy cycle.	Institutionalize the curriculum; agreements with principals to form stable teams from 1st to 3rd grade.
Initial receptivity to training.	Recognize teachers' trajectories in continuing training policies. Advance the construction of national curricula or guidelines that incorporate the science of reading.	Pedagogical leadership by principals. Build trust through technical credibility so that the program's coach is seen as a peer rather than as a supervisor. Invest in training coaches with a pedagogical rather than merely administrative profile.
PILLAR 4. SITUATED COACHING		
Geography and distance interrupt the frequency of in-person coaching.	Differentiated coaching routes: intensive in-person, express, or autonomous, depending on school accessibility	Define implementation routes adapted to the terrain. Hire people from the local area.
PILLAR 5: EVALUATION FOR LEARNING		
Late and non-actionable feedback: results arrive when the teacher no longer has those students, or in administrative formats that do not guide teaching	Align the assessment calendar with the school year and require feedback within the same cycle; keep formative assessment separate from teacher performance ratings	Short teacher-administered instruments with immediate results; one-page classroom reports indicating the suggested action (regrouping, reinforcing a phoneme); use instructional coaching to read the data alongside the teacher

Source: Authors' own elaboration based on seventeen interviews with coordinators from fourteen implementing organizations in six countries.

REFERENCES

[1] List, J.A., The Voltage Effect: How to Make Good Ideas Great and Great Ideas Scale. 2022, New York, NY: Currency.

[2] Duflo, E., The Economist as Plumber. American Economic Review, 2017, 107(5): p. 1–26. https://doi.org/10.1257/aer.p20171153

[3] Patsopoulos, N.A., A Pragmatic View on Pragmatic Trials. Dialogues in Clinical Neuroscience, 2011, 13(2): p. 217–224. https://doi.org/10.31887/DCNS.2011.13.2/npatsopoulos

[4] Ryan, A., Prieto-Rodríguez, E., Miller, A., and Gore, J., What Can Implementation Science Tell Us About Scaling Interventions in School Settings? A Scoping Review. Educational Research Review, 2024, 44: 100620. https://doi.org/10.1016/j.edurev.2024.100620

[5] Patton, M.Q., Qualitative Research & Evaluation Methods. 4th ed. 2015, Thousand Oaks, CA: SAGE.

[6] Damschroder, L.J., Reardon, C.M., Widerquist, M.A.O., and Lowery, J., The Updated Consolidated Framework for Implementation Research Based on User Feedback. *Implementation Science*, 2022, 17: 75. https://doi.org/10.1186/s13012-022-01245-0

[7] Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., and Hensley, M., Outcomes for Implementation Research: Conceptual Distinctions, Measurement Challenges, and a Research Agenda. Administration and Policy in Mental Health, 2011, 38(2): p. 65–76. https://doi.org/10.1007/s10488-010-0319-7

[8] Moir, T., Why Is Implementation Science Important for Intervention Design and Evaluation Within Educational Settings? Frontiers in Education, 2018, 3: 61. https://doi.org/10.3389/feduc.2018.00061

[9] The Reading League, *Science of Reading: Defining Guide*. 2022, The Reading League: Syracuse, NY. https://www.thereadingleague.org/what-is-the-science-of-reading/

[10] Argentinos por la Educación, Hoja de Ruta para la alfabetización: Guía para una política transformadora en Argentina. 2025, Argentinos por la Educación: Buenos Aires.

[11] Crawford, M.F., Barron Rodríguez, M.R., Ding, E.Y.Z., Gutiérrez Bernal, M., Saavedra, J., and Arias Díaz, O.S., Attaining the Learning Target: A Policy Package to Promote Literacy for All Children. 2021, World Bank: Washington, DC. https://documents.worldbank.org/curated/en/565171627618840464

[12] Global Education Evidence Advisory Panel (GEEAP), Cost-Effective Approaches to Improve Global Learning: What Does Recent Evidence Tell Us Are "Smart Buys" for Improving Learning in Low- and Middle-Income Countries? 2023, World Bank Group: Washington, DC.

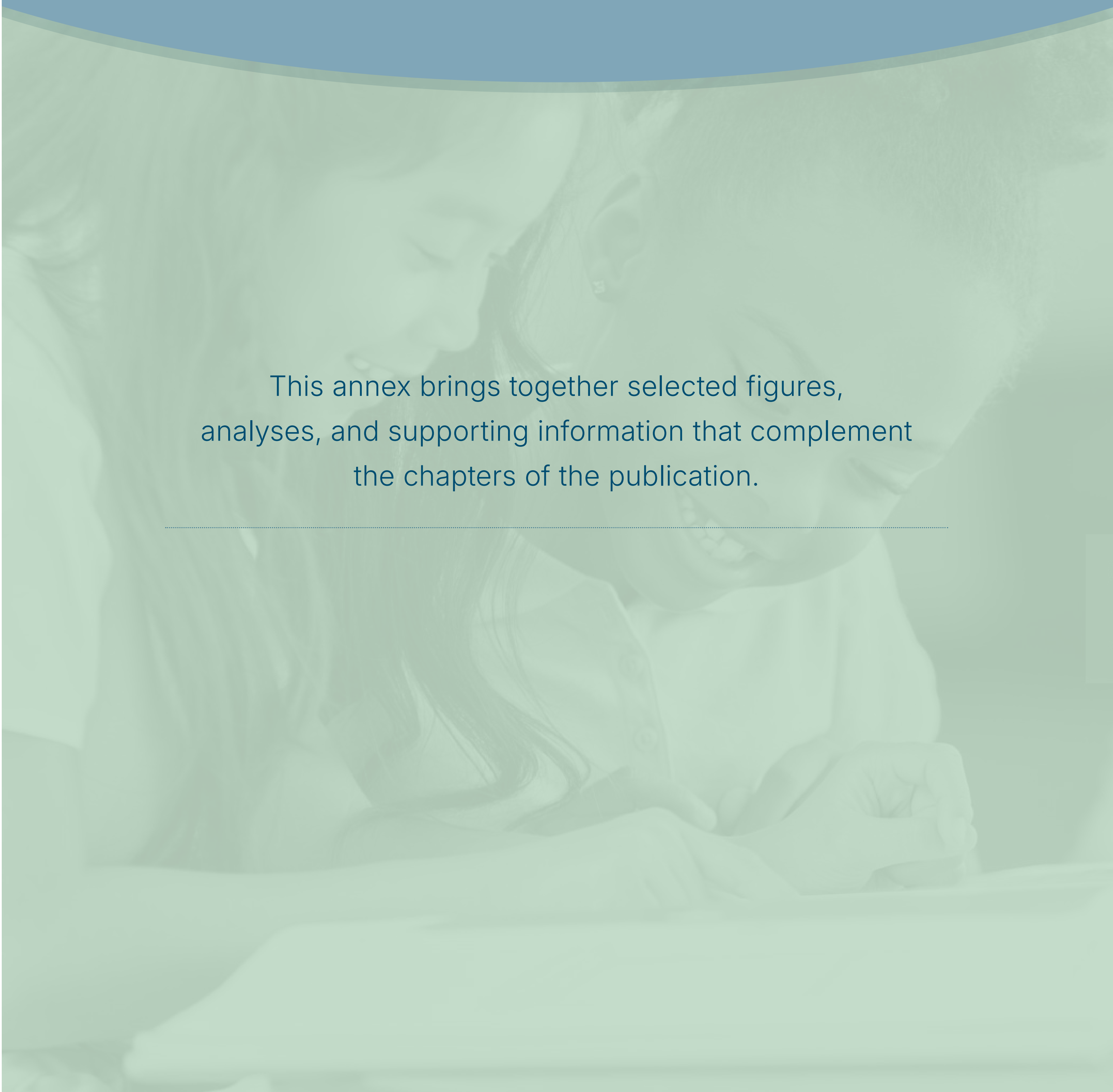
[13] Alvarez Marinelli, H., Boggild-Jones, I., Crawford, M., Dubeck, M.M., Jhingran, D., Lack, C.J., Mohohlwane, N., Oviedo Buitrago, M.E., Piper, B., Saavedra, J., and Taha, H., *Effective Reading Instruction in Low- and Middle-Income Countries: What the Evidence Shows*. 2025, Global Education Evidence Advisory Panel (GEEAP) and World Bank Group: London, Washington, DC, and New York.

[14] Michie, S., Fixsen, D., Grimshaw, J.M., and Eccles, M., Specifying and Reporting Complex Behaviour Change Interventions: The Need for a Scientific Method. Implementation Science, 2009, 4: 40. https://doi.org/10.1186/1748-5908-4-40

[15] Fallon, L.M., Cathcart, S.C., DeFouw, E.R., O'Keeffe, B.V., and Sugai, G., Promoting Teachers' Implementation of Culturally and Contextually Relevant Class-Wide Behavior Plans. Psychology in the Schools, 2018, 55(3): p. 278–294. https://doi.org/10.1002/pits.22107

[16] Norman, D.A., The Design of Everyday Things. Revised and Expanded Edition. 2013, New York, NY: Basic Books.

Annexes



This annex brings together selected figures, analyses, and supporting information that complement the chapters of the publication.

ANNEXES

BOX A2

Assessment Methodologies and Country Coverage

This annex provides methodological detail on the three assessments used in Chapter 2: the Regional Comparative and Explanatory Study (ERCE), the Programme for International Student Assessment (PISA), and the Programme for the International Assessment of Adult Competencies (PIAAC). Each uses a different design, proficiency scale, and target population. The chapter consistently focuses on the share of learners performing below the minimum proficiency threshold defined by each assessment.

ERCE — Regional Comparative and Explanatory Study

ERCE is UNESCO's regional assessment of learning outcomes in primary education, administered by the Latin American Laboratory for Assessment of the Quality of Education (LLECE). The 2019 cycle assessed reading and mathematics among third- and sixth-grade students in sixteen countries. Results are reported using four performance levels (I to IV) reflecting mastery of skills expected at each stage of primary education. Students performing at Level I are considered to fall short of minimum proficiency: in reading, they cannot yet reliably locate explicit information in a short text or draw simple inferences from it.

Participating countries in ERCE 2019: Argentina, Brazil, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, and Uruguay. Chile and Bolivia are members of the LLECE but did not participate in ERCE 2019, as the socioeconomic and political situation in late 2019 prevented the application of the assessment instruments; Chile therefore does not appear in the primary-level analyses in this chapter, although it is included in the PISA and PIAAC analyses. Country-level results are presented in Box A2 in the Annex. ERCE 2019 is comparable to TERCE 2013, which uses an equivalent methodology and a largely overlapping set of participating countries, enabling trend analysis across the two cycles.

PISA — Programme for International Student Assessment

PISA is the OECD's triennial assessment of the knowledge and skills of 15-year-old students in reading, mathematics, and science. Rather than focusing on curriculum coverage, PISA measures how well students can use their skills in unfamiliar contexts. Results are reported using proficiency levels ranging from below Level 1 to Level 6; students performing below Level 2 are considered to fall short of the baseline level of proficiency.

PISA assesses students who are enrolled in school and have completed at least six years of formal education. Adolescents who have dropped out or are more than two years behind their expected grade are excluded. In 2022, 89 percent of 15-year-olds were eligible to participate in OECD countries, compared to only 71 percent in Latin America and the Caribbean, meaning PISA results in the region reflect a population that is already more educationally advantaged than the full cohort of that age.

Participating countries in PISA 2022 from Latin America and the Caribbean: Argentina, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, El Salvador, Guatemala, Jamaica, Mexico, Panama, Paraguay, Peru, and Uruguay. Long-term trend analysis uses PISA 2006 as the baseline, or 2009 for countries that did not participate in that round. Four Latin American countries — Chile, Colombia, Costa Rica, and Mexico — are OECD members; throughout the chapter they are included in both the OECD average and the regional average, which means the reported gaps between the two groups are, if anything, conservative.

PIAAC — Programme for the International Assessment of Adult Competencies

PIAAC is the OECD's large-scale survey of the literacy, numeracy, and problem-solving skills of adults aged 16 to 65. Proficiency is classified into six levels, ranging from below Level 1 to Level 5. In numeracy, Level 3 is considered the benchmark for

continue...

BOX A2

Assessment Methodologies and Country Coverage / *...continuation*

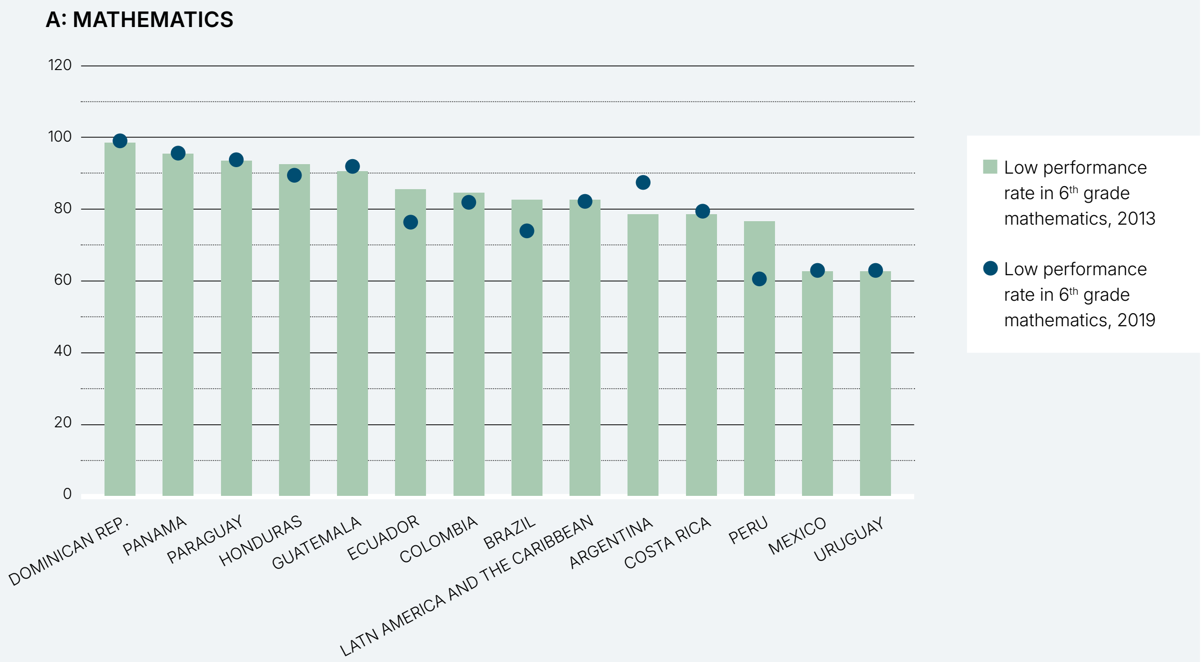
minimum proficiency: it requires adults to interpret less explicit information, work with more complex representations, apply multiple steps, and select appropriate problem-solving strategies. Adults performing below Level 3 are considered not to have reached minimum proficiency. The same benchmark is applied in literacy.

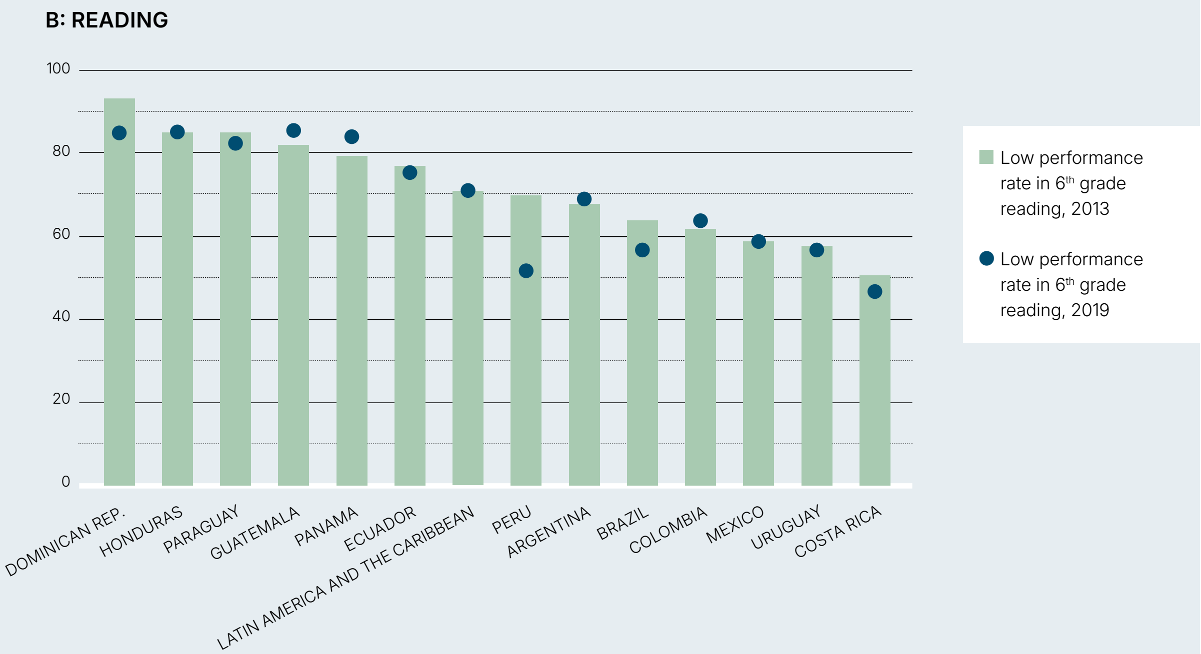
Four Latin American countries participated in the first cycle of PIAAC: Chile, Ecuador, Mexico, and Peru. Chile is the only country with data from two comparable cycles (2018 and 2023), enabling a limited comparison over time. The second cycle expanded country coverage, but comparable data for the other three countries are not yet available.

Comparability Across Assessments

The three assessments differ in design, target population, and proficiency scales and cannot be directly compared to one another. Chapter 2 does not attempt to merge results across assessments. Instead, it uses each assessment to examine foundational skill levels at the relevant stage of the life cycle — primary school (ERCE), secondary school (PISA), and adulthood (PIAAC) — focusing consistently on the share of learners below each assessment's defined minimum proficiency threshold.

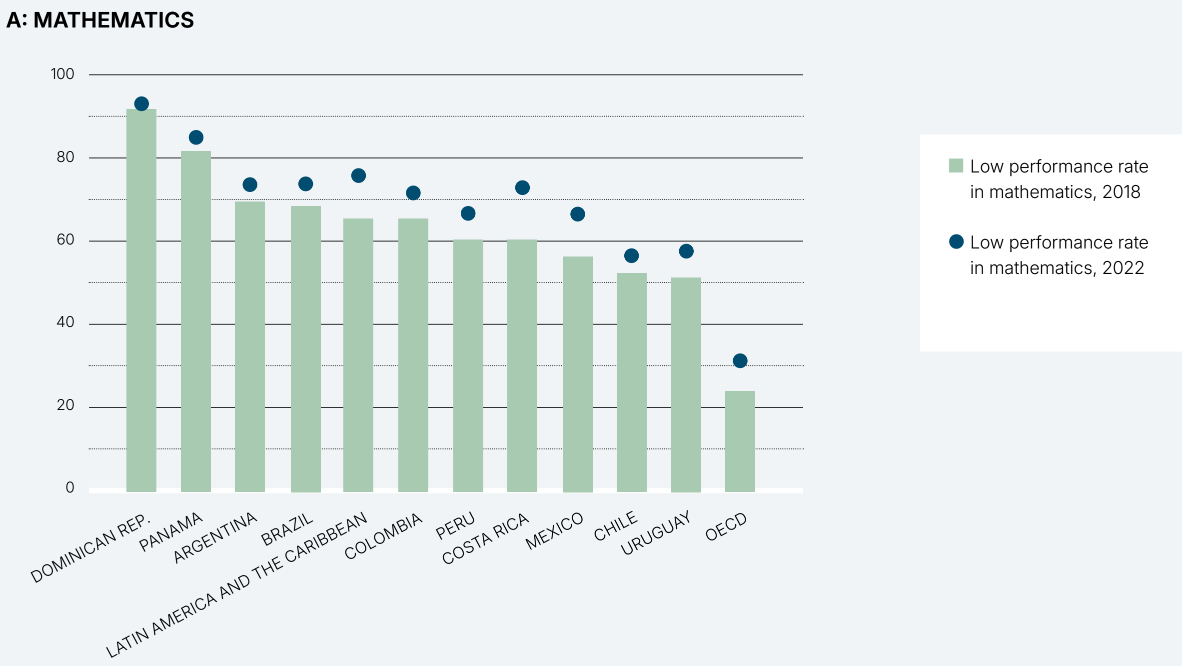
FIGURE A2.1: HAS LEARNING IMPROVED IN THE EARLY GRADES? CHANGE
between 2013 and 2019 in percentage of students not reaching minimum proficiency in primary school, ERCE 2019 and TERCE 2013

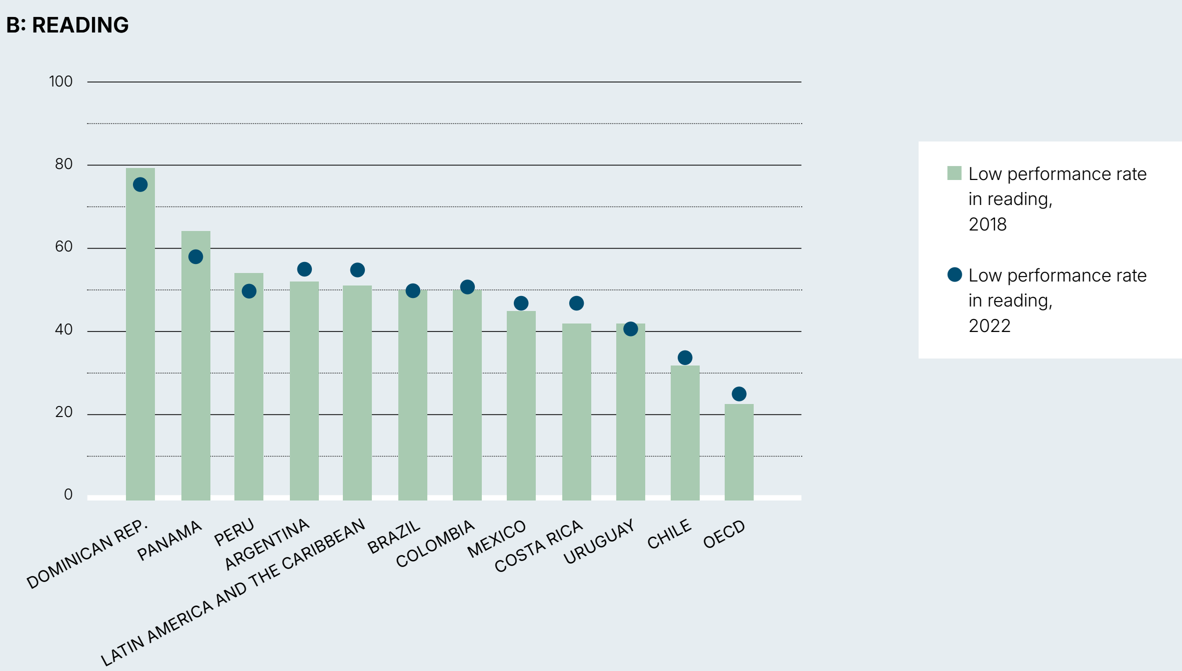




Source: CIMA (2025) based on ERCE 2019 and TERCE 2013 (LLECE-UNESCO).

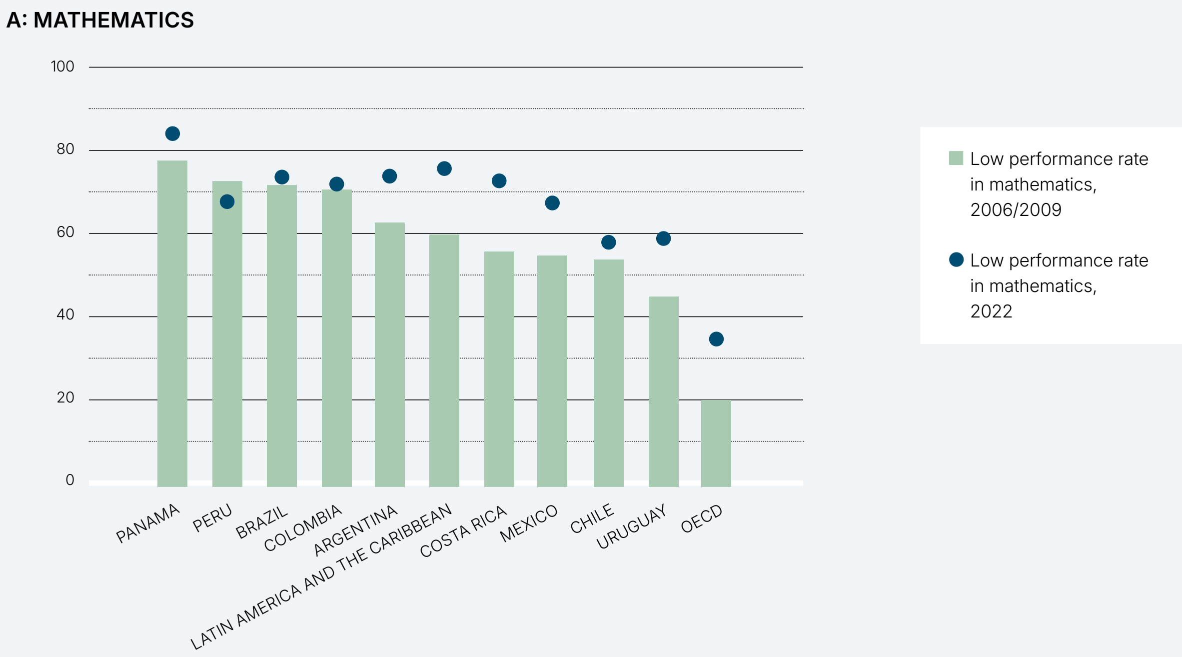
FIGURE A2.2: HAS LEARNING IMPROVED RECENTLY?
Change between 2018 and 2022 in low-performance rate secondary school, PISA 2018 and 2022

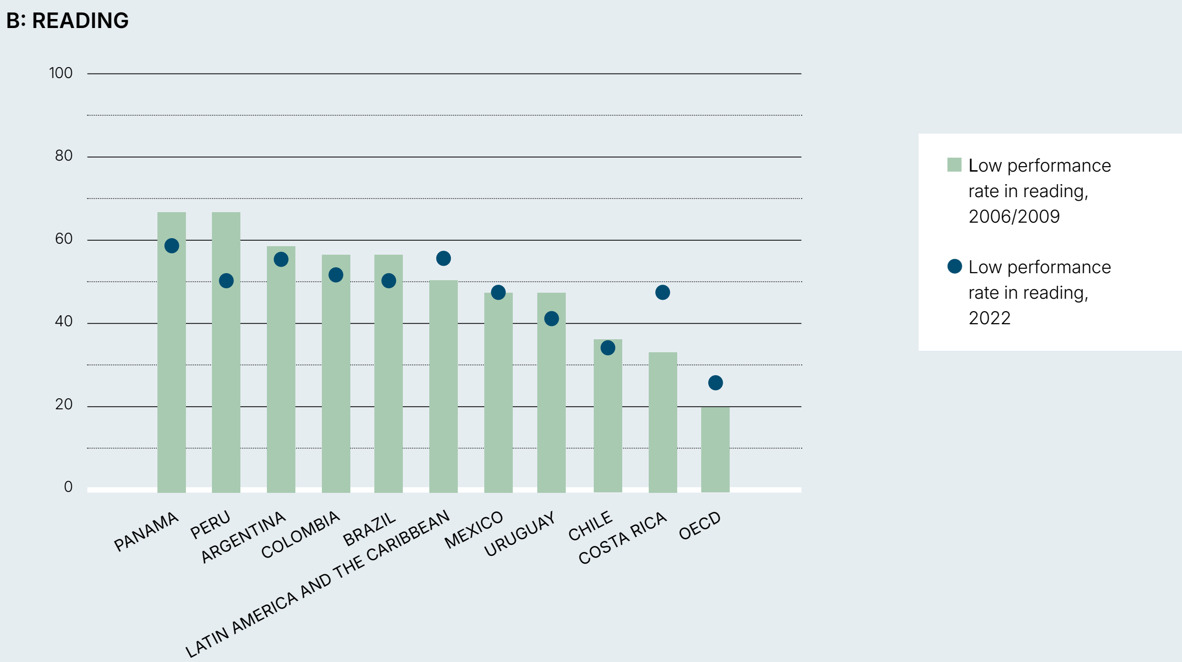




Source: CIMA (2025) based on PISA 2022 and PISA 2018 (OECD).

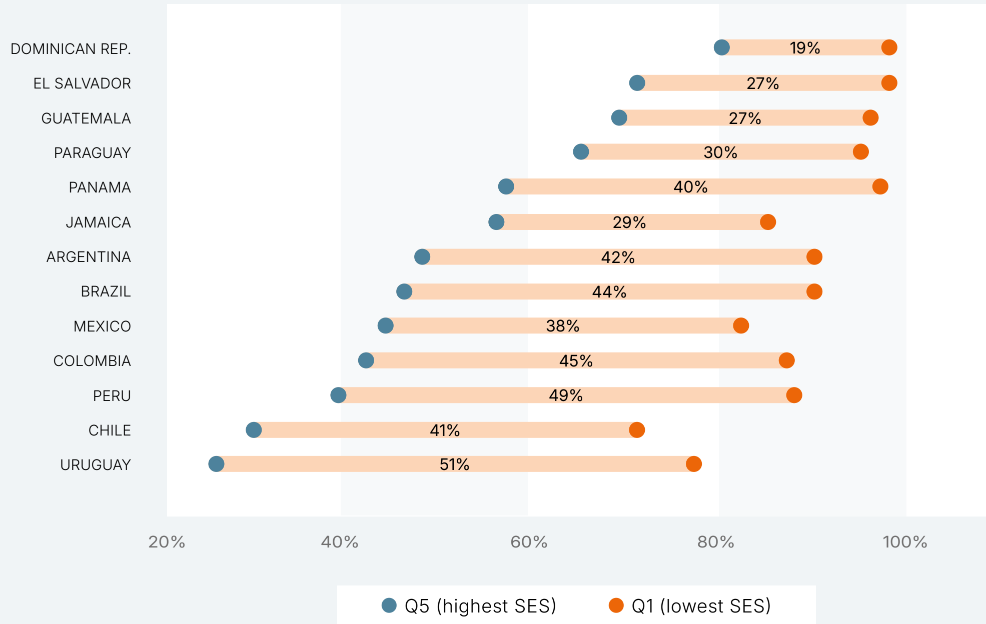
FIGURE A2.3: HAS LEARNING IMPROVED OVER THE LONG RUN?
Change between 2006/2009 and 2022 in percentage of students not reaching minimum proficiency in secondary school, PISA 2006, 2009 and 2022





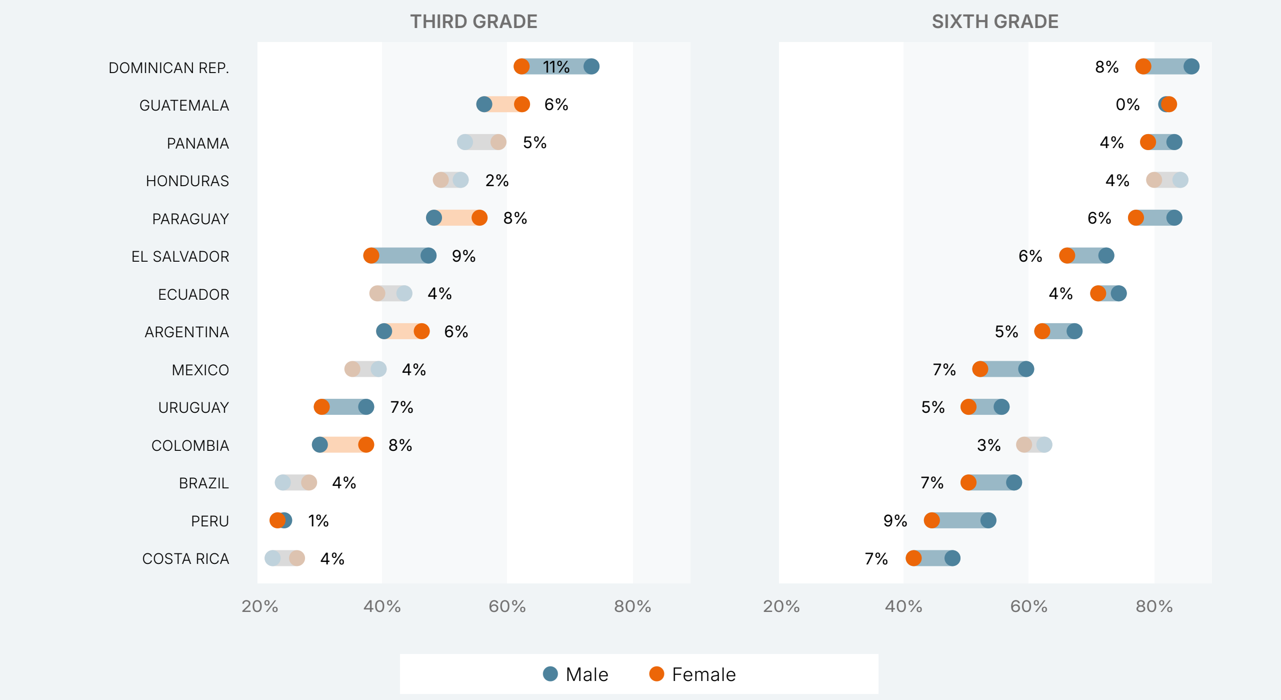
Source: CIMA (2025) based on PISA 2006, 2009 and 2022 (OECD).

FIGURE A2.4: DOES FAMILY INCOME DETERMINE WHO LEARNS?
Gap by socioeconomic status in percentage of students not reaching minimum proficiency in mathematics in secondary school, PISA 2022



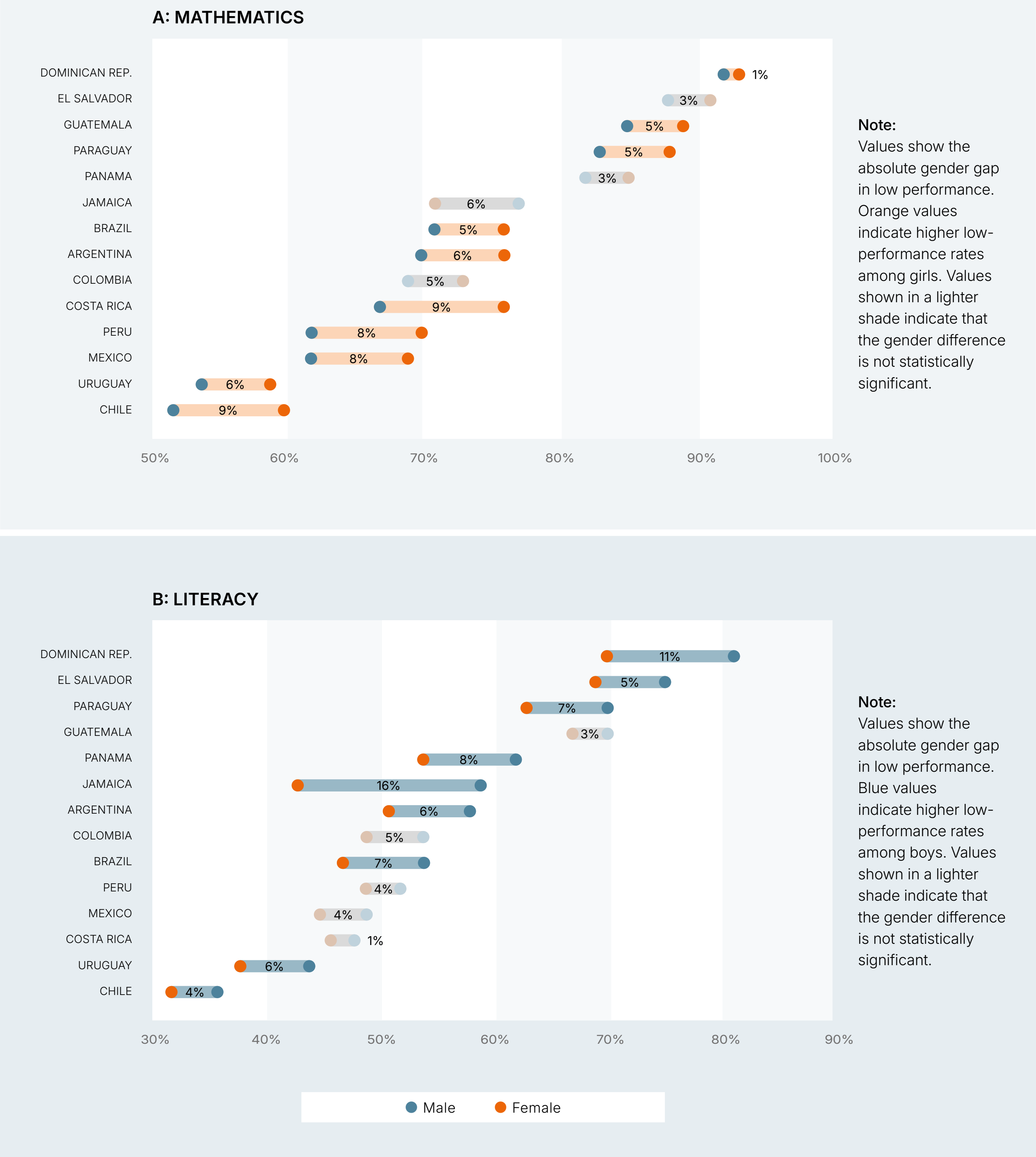
Note: Values show the gap in low performance by socioeconomic status, calculated as the low-performance rate among quintile 1 (poorest) students minus the rate among quintile 5 (richest) students. Orange values indicate higher low-performance rates among the poorest students.
Source: CIMA (2025) based on PISA 2022 (OECD).

FIGURE A2.5: DO GIRLS PULL AHEAD IN READING EARLY?
Gap by gender in percentage of students not reaching minimum proficiency in reading in primary school, ERCE 2019



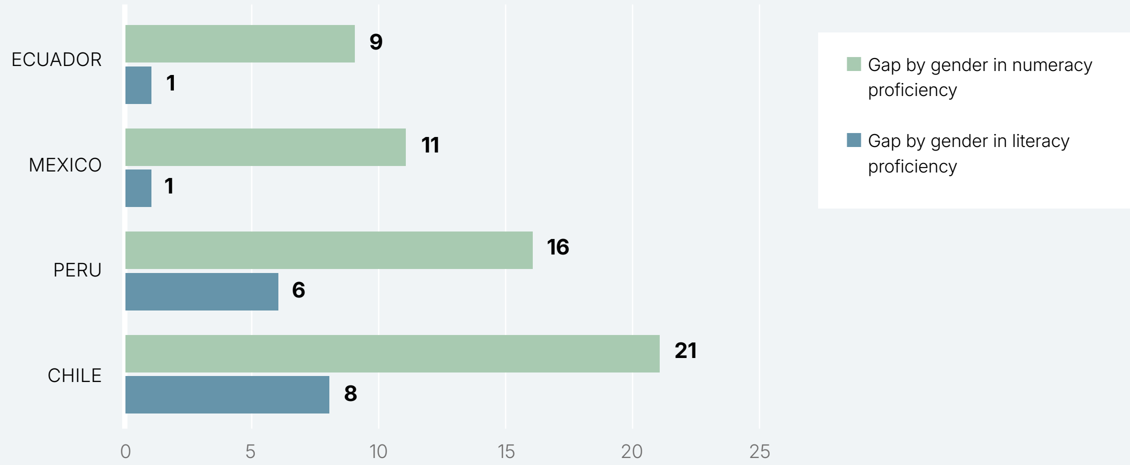
Note: Values show the absolute gender gap in low performance. Blue values indicate higher low-performance rates among boys; orange values indicate higher low-performance rates among girls. Values shown in a lighter shade indicate that the gender difference is not statistically significant.
Source: CIMA (2025) based on ERCE 2019 (LLECE-UNESCO).

FIGURE A2.6: DO BOYS AND GIRLS LEARN DIFFERENTLY?
Gap by gender in percentage of students not reaching minimum proficiency at secondary school, PISA 2022



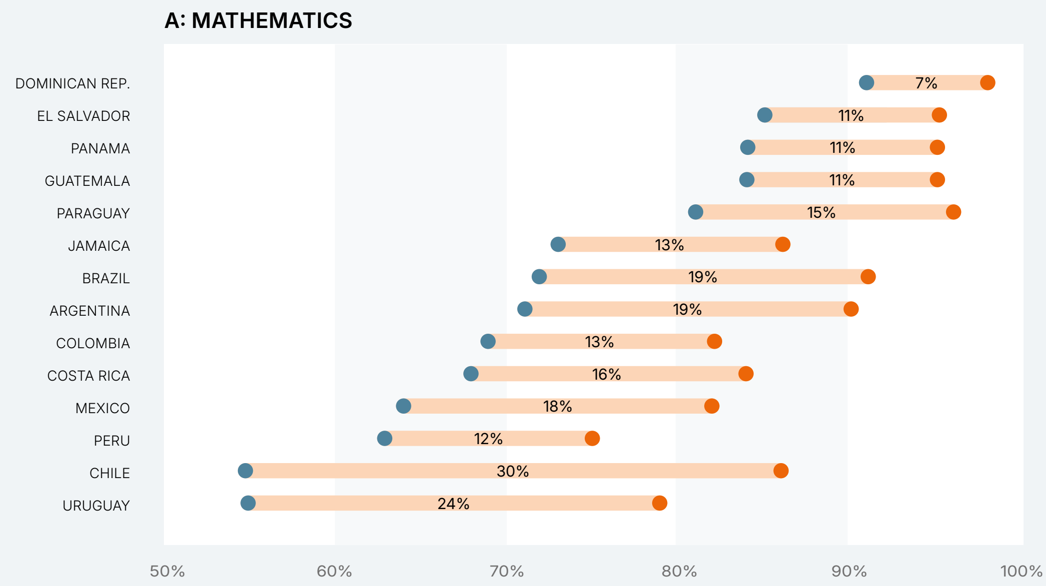
Source: CIMA (2025) based on PISA 2022 (OECD).

FIGURE A2.7: DO GENDER DIFFERENCES FOLLOW US INTO ADULT LIFE?
Proficiency gaps by gender in Survey of Adults Skills, 2018

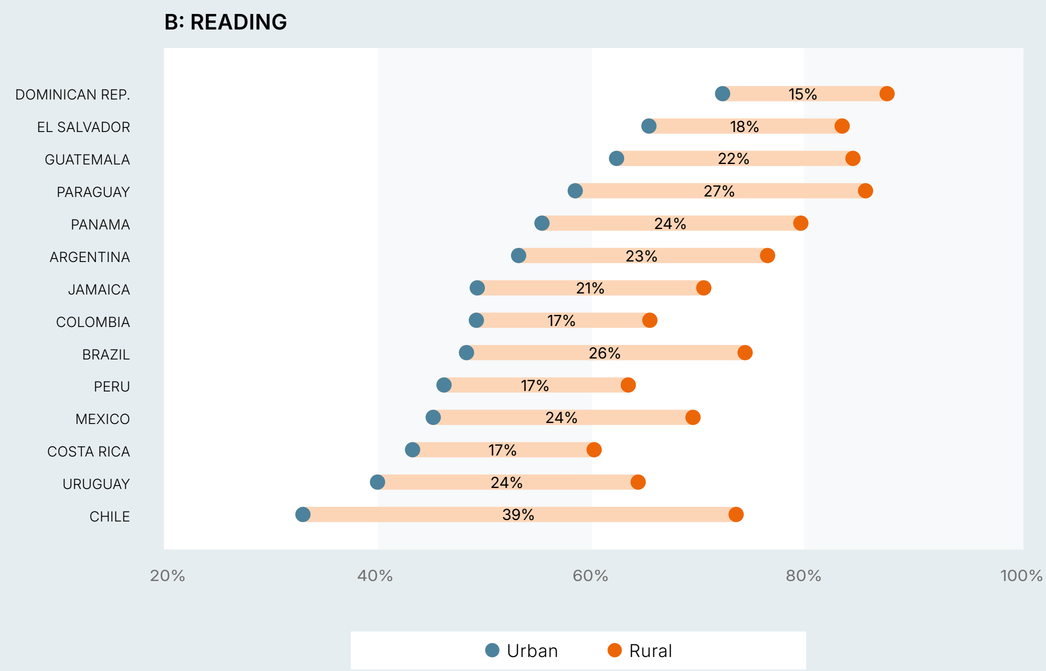


Source: Survey of Adults Skills, PIAAC first cycle (OECD).

FIGURE A2.8: DOES WHERE STUDENTS LIVE SHAPE WHAT THEY LEARN?
Gap by geographic area in percentage of students low performers in secondary school, PISA 2022.



Note: Values show the gap in low performance by geographic area, calculated as the low-performance rate among rural students minus the rate among urban students. Orange values indicate higher low-performance rates among rural students.

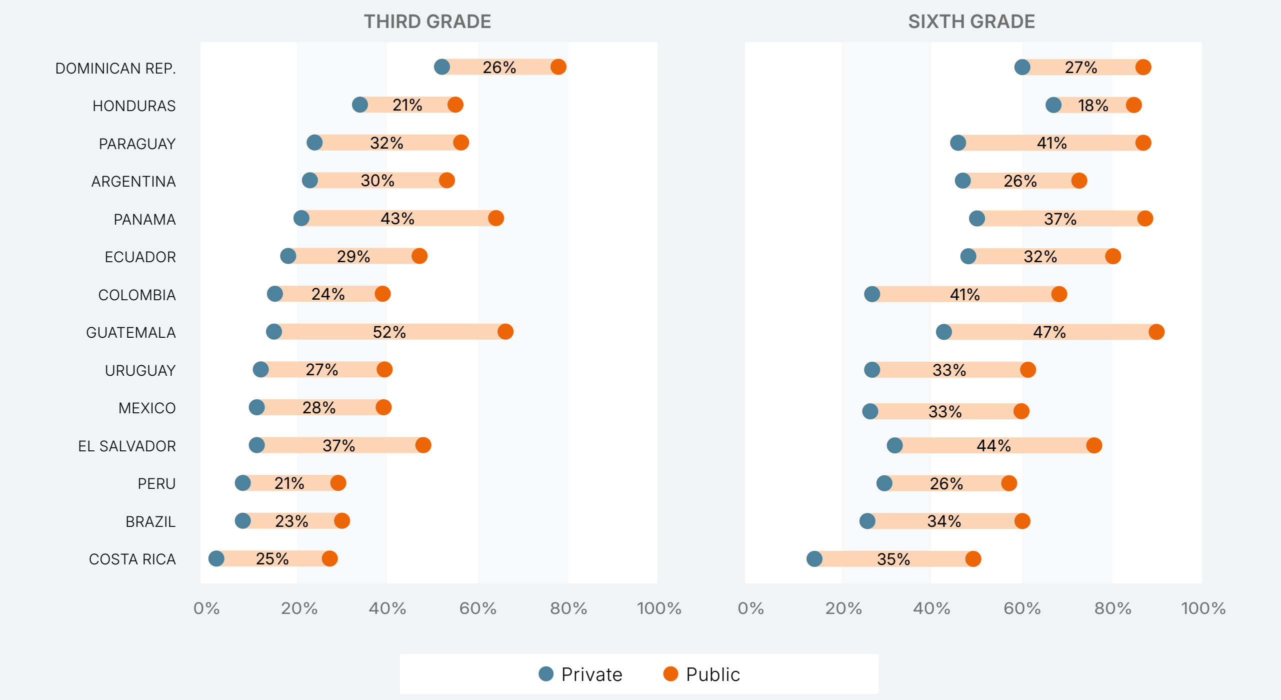


Note: Values show the gap in low performance by geographic area, calculated as the low-performance rate among rural students minus the rate among urban students. Orange values indicate higher low-performance rates among rural students.

Source: CIMA (2025) based on PISA 2022 (OECD).

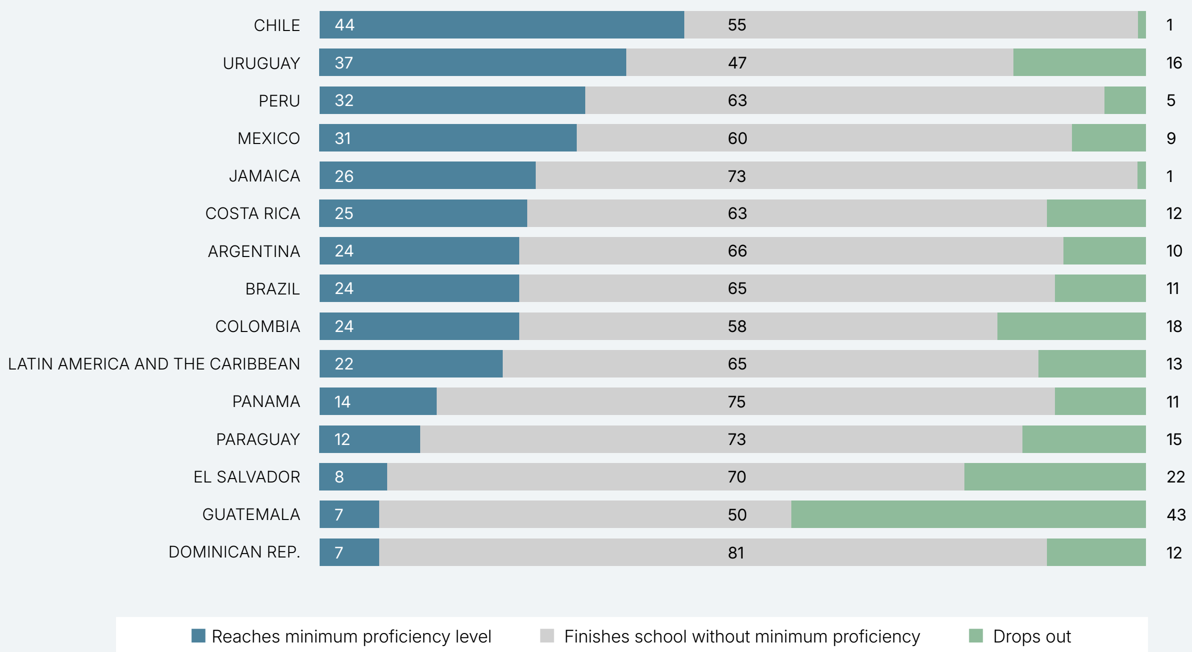
FIGURE A2.9: DOES SCHOOL TYPE MAKE A DIFFERENCE?

Gap by type of institution in percentage of students low performers in reading in primary school, ERCE 2019



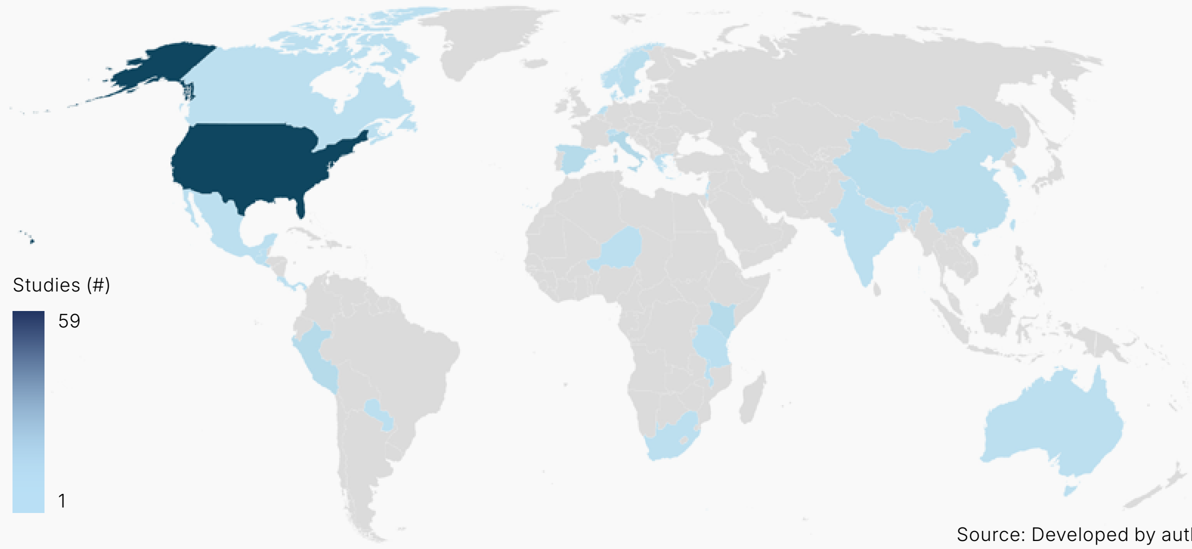
Note: Values show the gap in low performance by school type, calculated as the low-performance rate among public school students minus the rate among private school students. Orange values indicate higher low-performance rates among public school students.
Source: CIMA (2025) based on ERCE 2019 (LLECE-UNESCO).

FIGURE A2.10: THE FUNNEL EFFECT: HOW LEAKY IS THE REGION'S EDUCATION PIPELINE?
Percentage of 15-year-olds who reach minimum proficiency in reading, finish school without it, or drop out, PISA 2022



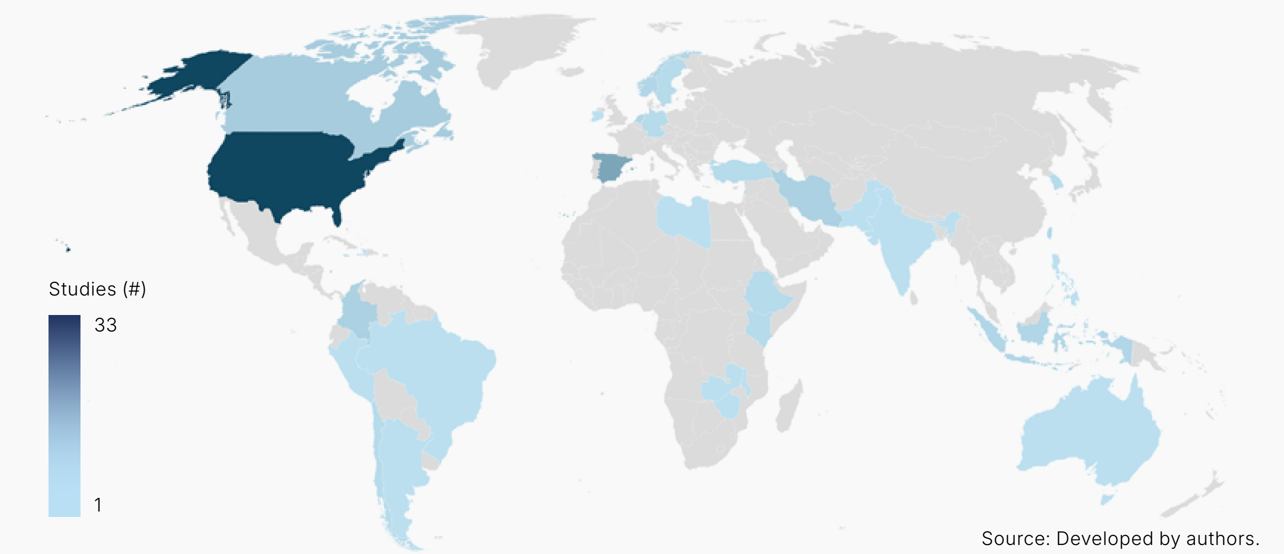
Source: CIMA (2025) based on PISA 2022 (OECD).

FIGURE A4.1: WHERE HAVE RIGOROUS IMPACT EVALUATIONS ON MATHEMATICS BEEN CONDUCTED?
Evidence remains concentrated in North America, though one in five studies comes from low -and middle-income countries



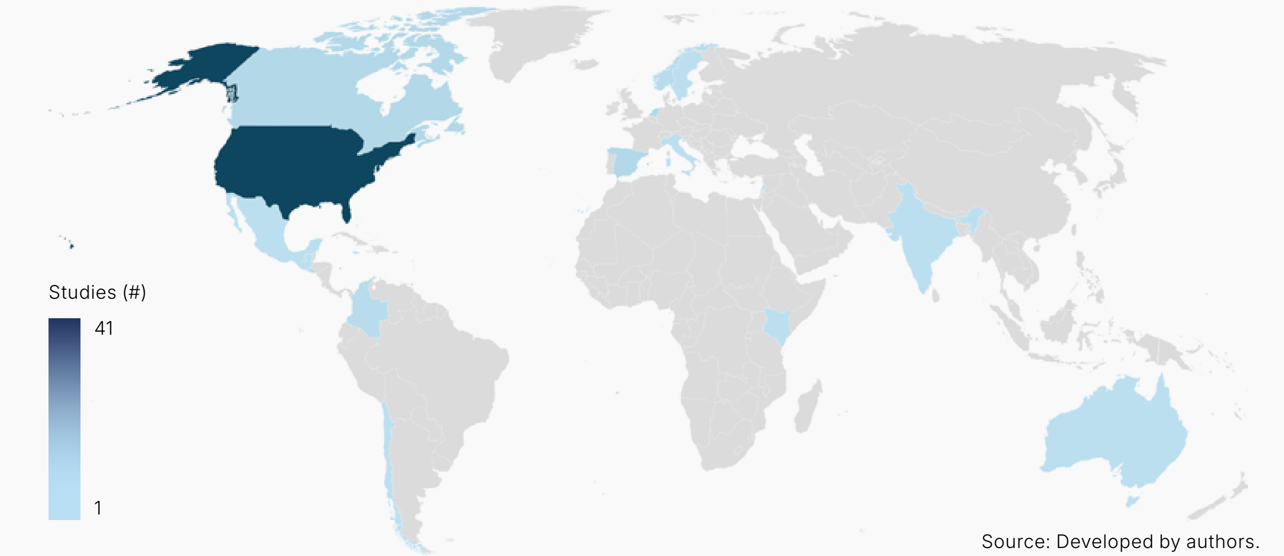
Source: Developed by author.

FIGURE A5.1: WHERE HAVE RIGOROUS IMPACT EVALUATIONS ON LITERACY BEEN CONDUCTED?
Evidence remains concentrated in North America, though close to a third comes from low -and middle-income countries.



Source: Developed by authors.

FIGURE A6.1: WHERE HAVE RIGOROUS IMPACT EVALUATIONS ON REMEDIAL EDUCATION AND TUTORING BEEN CONDUCTED?
Evidence is concentrated in North America and Europe, with only one in seven studies coming from low- and middle-income settings



Source: Developed by authors.