IMAGO
Global Grassroots

How Recife responded to the challenge of learning deficits in the post-COVID era

An adaptive evaluation of a complex intervention

Caique Bellato, Vinicius Bueno, Carlos Palacios and Michael Walton

May 9th, 2025

How Recife responded to the challenge of learning deficits in the post-COVID era

An adaptive evaluation of a complex intervention

Caique Bellato, Vinicius Bueno, Carlos Palacios and Michael Walton[1]

Abstract

School closures during COVID-19 caused learning losses worldwide, often compounding existing deficits. Brazil, with one of the longest closures, is a particularly striking case. This paper examines how the city of Recife responded after reopening schools in 2022 and 2023. With education largely decentralized and managed at the municipal level in Brazil, Recife's education department rolled out various system-wide measures—adjusting pedagogy, offering opportunities to group students by learning levels, and sharing learning outcome data with schools to inform teachers. Given the multifaceted nature of the problem, heterogeneity in implementation, and the evolving nature of the intervention, a quasi-experimental design was neither feasible nor appropriate. Instead, a simple application of an adaptive evaluation approach—involving systems diagnostics and process tracing—was employed for a structured assessment of policy implementation, school-level variation, and shifts in learning outcomes. While the core guiding principles remained stable, 2023 saw a shift toward increased school-level discretion, reduced testing burdens, and more timely feedback. Process tracing of schools showed that guidance was generally transmitted reasonably effectively, but uptake varied, revealing limitations in translating data into classroom practice. Although statistical causal attribution was not possible, average learning outcomes in Portuguese and Math recovered to or surpassed pre-pandemic levels, with substantial variation across schools. Qualitative analysis of high performing "positive deviant" schools highlighted the importance of data use, internal collaboration, and flexible regrouping. These findings support reform strategies that combine central coordination with local autonomy and show how an adaptive evaluation can illuminate the workings of complex education systems, where traditional causal inference is not possible.

[1] Carlos Palacios and Caique Bellato are with CAEd, Vinicius Bueno was with the Harvard Kennedy School during the research, and Michael Walton is with Imago Global Grassroots and the Harvard Kennedy School. Thanks to Fred Amâncio, Ana Selva, and Alexsandra Felix from the Recife Department of Education for making the research in Recife possible and for the valuable lessons learned throughout the project; to Lina Kátia Mesquita de Oliveira and Manuel Palacios for their institutional support for the research within CAEd; and from Imago, to Jossie Fahsbender for her support on design and workshopping, to Siddhant Gokhale and Ronaldo Rodrigues Alves Braga for the careful work on process tracing.

Table of Contents

1. Introduction

Throughout the world, children's learning suffered significant setbacks in the wake of school closures during the Covid-19 pandemic. These setbacks were particularly problematic when they interacted with longer-term learning problems within school systems, often magnifying pre-existing inequalities between children from different socio-economic, cultural and geographic contexts. Brazil is a striking case. It entered the democratic era in the late 1980s with huge learning disadvantages outside elite groups. There have been major government efforts in the subsequent thirty years that have led to substantial gains in learning outcomes of subsequent cohorts of children. However, progress had significantly stalled in the decade from 2010, across reading, maths and science, at levels substantially below OECD benchmarks. There have been persistent stark inequalities across socio-economic groups in the population— disproportionately affecting children of households of black and mixed race, and in poorer parts of Brazil, notably the North and Northeastern states. While Brazil's learning outcomes are not markedly different from other large Latin American countries, such as Argentina, Colombia, Mexico, and Peru, this is small comfort, given the broader assessment of a long-term learning crisis in the region.

When Covid-19 arrived, Brazil, along with much of the world, pursued a strategy of school closure to reduce the spread of the disease. Schools were fully closed for almost 270 days, and partially closed for a further 290 days in the period from February 2020 to early 2022. While there was a major effort to sustain schooling through online methods, this was particularly hard for younger children, and for those from poorer households. This exacerbated learning deficits, both on average and across groups. Such deficits can be perpetuated across future years of education, given the cumulative nature of learning processes. There was a major risk that the school closures would lead to permanent disadvantage for large parts of the population.

This is the context for major efforts throughout Brazil's schooling system to restore learning levels, and to do so in ways that would also tackle longer-term learning challenges. As Brazil's educational system is highly decentralized to the state and municipal levels, it is of value to explore public action and outcomes at these levels. The municipality of the city of Recife provides an interesting case. It is the capital of the (relatively poor) Northeastern state of Pernambuco, with learning levels slightly below the Brazilian average. It comprises a wide and

complex network of schools. According to the 2022 School Census, Recife's municipal network had 328 schools with 2,587 teachers and 67,910 enrolled students (INEP, 2022). In the context of the return to school opening, Recife's education department developed a creative range of interventions within this school system, with pedagogies designed to adapt to learning challenges—termed "Learning Recomposition" (*recomposicão)* at the center of these. Recife's education department invited a team from CAEd and Imago Global Grassroots to assess this process.[2] This paper explores the implementation of key elements of this effort for the 2022 and 2023 school years.

The interventions were implemented throughout Recife's network of public schools, and involved a set of complementary activities, including ongoing learning assessments, feedback to schools on the finding of the assessment, adapted pedagogies, and for most of the system, regrouping of students according to their learning levels, for specific sessions of the week. There was also some adaptation of the pattern of interventions over time, especially between the 2022 and 2023 school years. The socio-economic context and initial position of schools varied significantly across the system. Because of this intervention structure, an experimental approach was neither feasible nor desirable. Rather the research involved a mixed-method observational approach, that we call "adaptive evaluation", that is, an evaluation strategy designed for complex settings, involving system diagnosis, theory-based assessment of processes of change in an implementation sequence, and, potentially, feedback on iterative design. The core empirical assessment of changes was based on process tracing of hypothesized theories of change, built from descriptions, interviews, observations and quantitative data (including on learning outcomes). In-depth qualitative research was undertaken on a small-subset of schools, to support within-case analysis of implementation at the individual school level, that was then extended to comparative case analysis, and quantitative assessment of system-wide trends in outcomes. This supported an in-depth analysis of implementation. While this analysis is embedded in the description of learning outcomes, direct causal attribution was not feasible either from logical or statistical tests, given the structure of the data. However, the research does support a rich analysis of a substantial policy effort, in the context of what emerged as a significant recovery of learning outcomes for major parts of the system and a majority of schools.

[2] CAEd is the Centro de Políticas Públicas e Avaliação da Educação of the Universidade Federal de Juiz de Fora, specialized in assessing student learning outcomes, Imago is a non-profit organization that works on scaling, systems change and adaptive evaluation.

The main findings were as follows. In terms of design, the Recife Education Department developed a pattern of interventions involving innovations in pedagogy, regrouping and assessment, that was well-adapted to the complex challenge faced in terms of both the Covid-19 shock and long-term learning deficits. While the principles remained largely constant between the two school years of 2022 and 2023, there was an important shift in 2023 to allowing more flexibility at school level in applying the department's guidance, lowering the burden of testing, while making efforts to provide test results in a timely fashion to schools. The in-depth process tracing in 2023 found that transmission of the guidance to schools from the Education Department was largely effective, with respect to executing and sharing test results from the regular "formative assessments" (undertaken by another part of CAEd), and advice on regrouping and pedagogies. While the schools in the subsample mostly followed subsequent steps in the implementation of the interventions, there was significant variation, for example in the extent to which assessment results were discussed and acted upon by principals and teachers. This is indicative of some fragility in the implementation support down to school levels within the system. On average, learning outcomes in Portuguese and Math had fully recovered to pre-Covid levels and beyond by 2023, again with significant variation across schools. While the research could not attribute the general learning gains, in statistical terms, to the Learning Recomposition there is a presumption that this was an important factor. There are revealing insights from qualitative work schools who experienced most gains in learning outcomes of children below grade level ("positive deviants") and those with least gains ("negative deviant.") While there are clear interactions with the local socio-economic context, especially for the negative deviant school, the positive cases indicate that the guidance on Learning Recomposition was combined with more extensive school-based activities, involving principals and teachers, including more intensive review of results, adapting responses and flexible re-grouping of students within the same classroom, school-based reviews, and teacher interactions, that were then taken into the classroom.

We draw two types of broader lessons from the research. In terms of design, the findings support a view aligned with the "third generation of education reforms" (Filho, 2022) that interventions should allow for some degree of flexibility at school level, with support and priorities aligned to the varying needs and capabilities at this level. Iterative exploration and ongoing feedback and learning can be incorporated into the system, with information on student learning levels provided to principals and teachers in a practical and timely fashion.

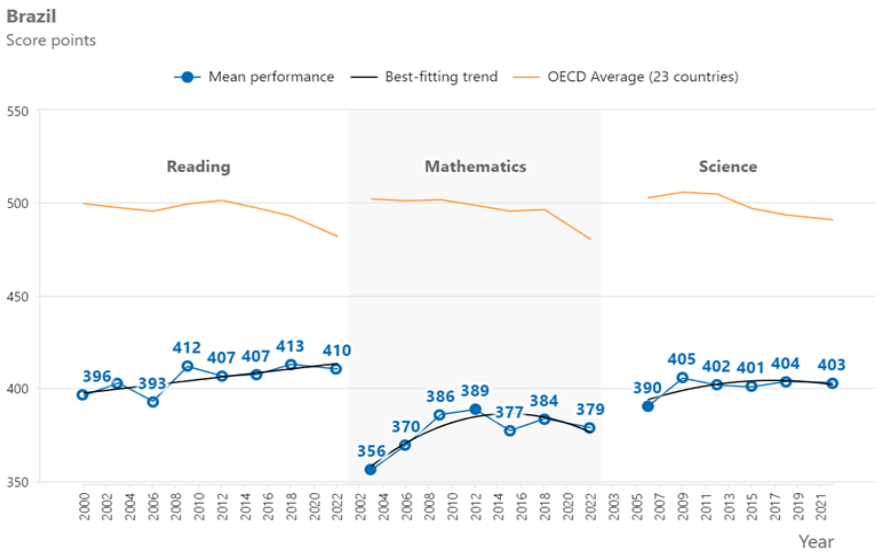
In terms of research methodology, the adaptive evaluation allowed for an assessment of a complex overall system. The building and testing of theories of change from the interventions through this system supporting a structured assessment of the implementation process. While specific engagement with an iterative design process was not feasible in this case, this could be incorporated in future work through the adaptive process.

2. Recife's education challenge in 2021-23

2.1. The educational situation

Even before the COVID-19 pandemic, Brazil faced significant challenges in education, especially related to its quality. Learning levels are substantially below the Organization for Economic Cooperation and Development (OECD) average in the internationally comparable Programme for International Student Assessment (PISA), which focuses on relatively sophisticated learning achievements of **15-year-old** students. Moreover, after some progress until 2008-10, measures of learning outcomes have essentially stagnated. (Figure 1).

Figure 1 – Brazil's has low and stagnating performance in PISA measures between 2000 and 2022



Source: OECD (Organization for Economic Co-operation and Development)

Brazil's poor performance is shared by several middle-income neighbors, including Colombia, Argentina but is significantly worse than Chile, (Table 1).

Table 1 – Brazil's learning outcome in the PISA test was in the lower half of comparable Latin American countries in 2022

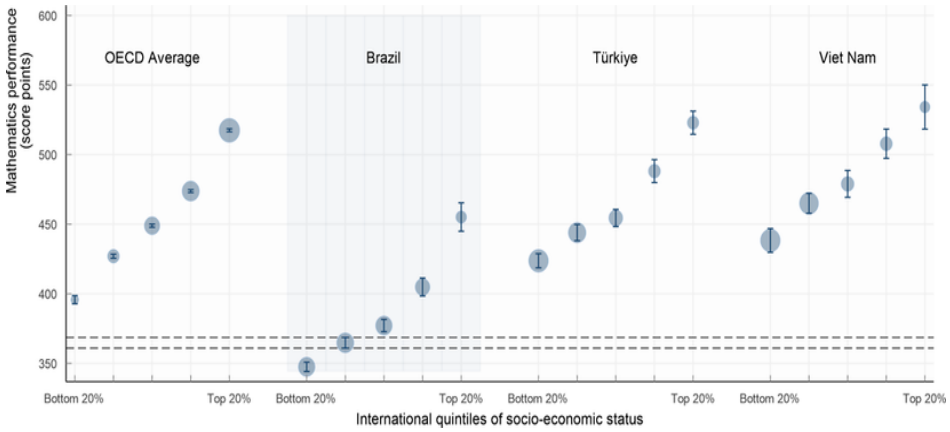
(PISA scores as average of math, science and reading; and ranking out of 81 geographic units)

Country	Score	Ranking
Chile	434.7	44
Mexico	406.7	52
Uruguay	404.3	55
Peru	402.3	58
Colombia	401.7	59
Brazil	**397.3**	**60**
Argentina	395.0	62
Panama	379.0	67
Paraguay	359.7	74
Dominican Republic	350.0	80

Source: OECD

Brazil's education quality problem also displays large inequalities, with a strong correlation with the socio-economic status of households, with a particularly large gap for the top quintile—see Figure 2.

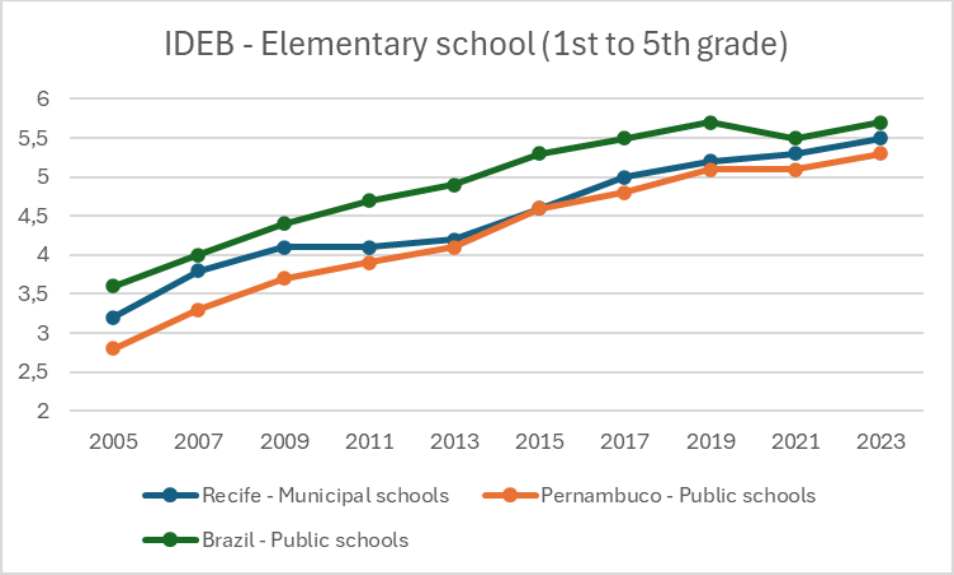
Figure 2 – Performance in maths tests is strongly correlated with socio-economic status, in Brazil and other countries.



Source: OECD

Within Brazil, children in Recife had learning levels slightly better than the average of its state, Pernambuco, that was also just below the Brazilian average (Figure 3). This is for the Brazilian Basic Education Development Index (IDEB) that measures the quality of education in the 5th grade of elementary school. The IDEB is calculated based on a combination of test scores and approval ratings, and evaluates more basic skills than PISA, indicating somewhat more gains at this level than at higher levels of learning.

Figure 3 – Recife's IDEB was comparable to Pernambuco but below the Brazil average between 2005 and 2023

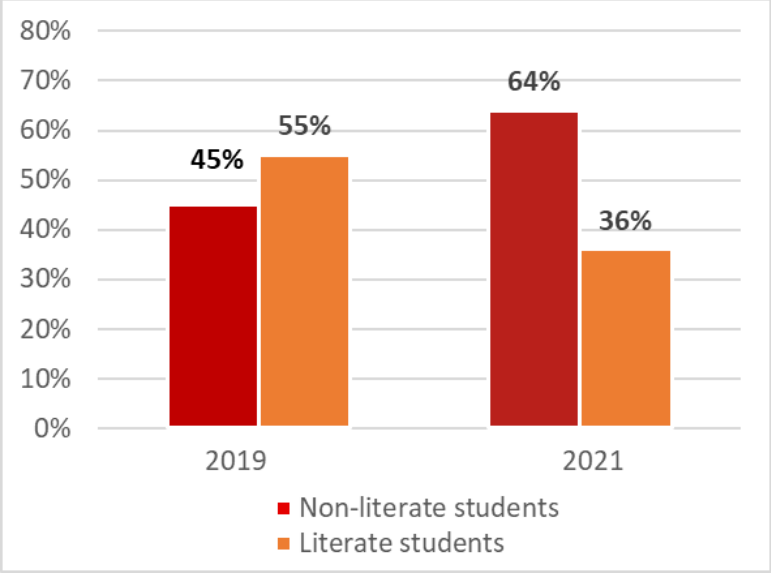


Source: Qedu and INEP,2024

The COVID-19 pandemic brought significant further challenges. The negative effects, observed around the world, were especially large in Brazil, given that the country was one of the nations with more days with schools closed, and there are significant gaps in access to the internet in the country. According to UNESCO's database of the duration of school closures, schools in Brazil were fully closed for 267 days and partially closed for 288 days in the period between February 16, 2020 and April 30th 2022[3]. Brazil is in the top 10% of countries with more days closed. Students in elementary school were especially affected by the remote teaching. Younger children had less autonomy to study at home, suffered more from lack of close support from teachers, and were more dependent on the support of their family members. This was exacerbated for children in poorer households with less educated parents.

At a national level, the size of the COVID-19 impact in the initial years is illustrated by a study conducted by the Anísio Teixeira National Institute of Educational Studies and Research (INEP) based on data from the Brazilian National Basic Education Assessment System (Saeb. This found a large increase in measured illiteracy of 2nd year students in public schools between 2019 and 2021 (Figure 4).

Figure 4 – The COVID-19 closures were associated with a large increase in illiterate students at the end of the 2nd year in Brazilian public schools

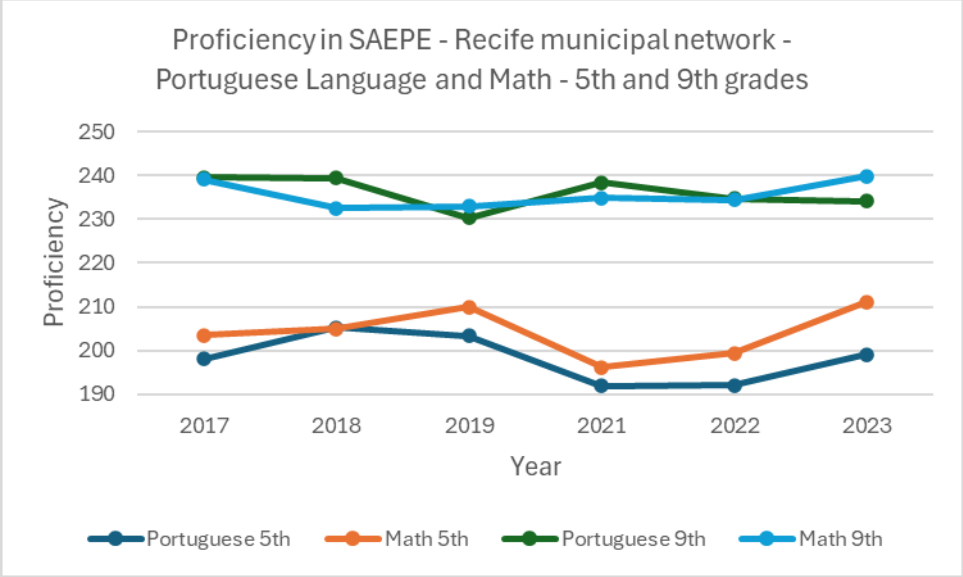


Source: INEP, 2023.

[3] UNESCO. (2022). Database of duration of school closures. Retrieved from https://covid19.uis.unesco.org/wp-content/uploads/sites/11/2022/09/SDG-duration-of-school-closures-by-country.xlsx

Results for the Recife municipal network from the state assessment system of Pernambuco (SAEPE), finds a larger impact in elementary school. Figure 5 below shows how there was a significant drop in proficiency for students in the 5th grade between 2019 and 2021, while in the 9th grade there was a small improvement. On the other hand, the learning recovery in elementary was more significant between 2021 and 2023, a pattern that will be explored later in this paper

Figure 5 – There were large declines in Portuguese and math proficiency for 5th grade students between 2019 and 2021, followed by some recovery. Learning at 9th grade experienced smaller impact but overall stagnated



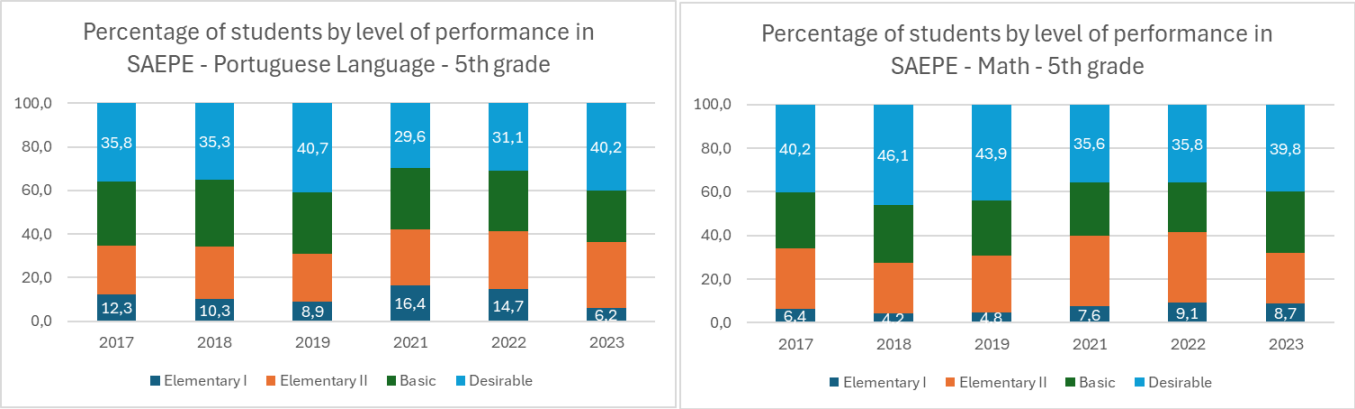
Source: CAEd/UFJF, 2024.

The learning recovery that can be observed between 2021 and 2023 in the 5th grade occurred in Recife at a similar pace as neighboring municipalities, Jaboatão dos Guararapes and Olinda (Annex A Figure A1). Between 2021 and 2022, Recife's performance stagnated in Portuguese and improved slowly in Math, but it recovered quickly between 2022 and 2023. In contrast, Jaboatão dos Guararapes and Olinda evolved more consistently in the period between 2021 and 2023. In the 9th grade, Recife and Olinda reached 2023 proficiency levels superior to the pre-pandemic levels in 2019, although with different trajectories (Annex A Figure A2 . In contrast, Jaboatão dos Guararapes still in 2023 had a lower performance in Portuguese than in 2019, while in Math it was at a similar level.

In addition to the negative impact on average learning levels in Recife, the pandemic also increased inequalities between students in the school system, including within the same

school and in the same classroom. This adds to the range of teaching challenges owing to the diversity of learning levels within the same class. For instance, Figure 6 shows that between 2019 and 2021 there was a considerable decrease in the percentage of students in the "Desirable" learning category - the highest performance level in SAEPE - and at the same time a considerable increase in the percentage of students in the Elementary I - the lowest performance level in SAEPE. There was a significant recovery by 2023, especially in Portuguese, but with continued substantial diversity in learning outcomes.

Figure 8 – Percentage of students of Recife by performance level in SAEPE – 5th grade



Source: CAEd/UFJF, 2023.

3. The Government of Recife's response

To respond to the learning challenges, the Recife Education Department designed and implemented from 2022 a broad and ambitious Learning Recomposition Plan. This brought together different actions related to the use of technologies and expansion of workload, curricular prioritization and regrouping of students by learning levels in schools, and was linked to an internal system of external formative assessments, called SAERE. The assessments were also conducted by CAEd and were applied at different times of the school year in order to provide quick feedback to schools on the students' learning process.

The Learning Recomposition Plan began to be implemented in 2022. The term "recomposition" was chosen instead of "recovery" based on the understanding that most students had not even been introduced to the pedagogic content, making it impossible to recover something that was never offered. Additionally, the plan was designed to support all

students, and not only those who had suffered major deficits during Covid.[4] The strategy was designed to reach students from the 1st to the 9th year of Elementary School, with specific interventions for each year. Some of the actions already existed in the Recife network. However, according to the reports of the interviewed professionals, the approach brought greater alignment and pedagogical coherence across the range of activities.

The research followed the implementation of some of these actions in 2022 and 2023, including adjustments and changes between the two school years, with an in-depth focus on implementation in 2023 (see below).

A first step for the Recomposition Plan involved a curricular prioritization on content and skills for 2022. Based on that, the Department chose strategies to ensure learning of those prioritized skills, including use of instructional materials, training of education professionals (principals, pedagogical coordinators, and teachers), assessment to follow students' learning over time, and the use of technological tools.

The "formative assessment system" of Recife (SAERE) was a key element for monitor student learning throughout the year, including the impacts of the learning recovery actions. It comprised periodic assessments for all students in the 1st to 9th grades. For 1st to 5th grade, the assessment applied to Portuguese and Mathematics. From the 5th to 9th grades, it also included History, Geography and Natural Sciences. SAERE aims to produce indicators of learning levels and students' needs over time. Based on that, principals and teachers can develop pedagogical interventions aligned with observed learning outcomes. This was also intended to strengthen the culture of evaluation and analysis of school data evaluations in the municipal network. Beyond that, SAERE data are used to inform the (re)design of public policies by the central municipal team.

As part of the recovery strategy, the Department also decided to introduce regrouping of students for selected sessions, in a way that is similar to Pratham's "Teaching at the Right Level" experience in India[5]. This is designed to group students based on the different learning levels with pedagogies aligned to each level, for at least part of a school week. This had been

[4] Regarding the term "Learning Recomposition," it is important to note that there is currently no consensus in Brazil on the adoption of this concept, nor is there significant academic production reflecting on the intended distinction between the terms recovery and recomposition. For the purposes of this article, we will use "learning recovery" to refer to actions aimed at regaining lost learning. The term recomposition is reserved for specific references to the Recife Learning Recomposition Plan.

[5] For more on the experience of Teaching at the right level in India, see Banerjee et al (2016).

adapted in other parts of Brazil, including the State of São Paulo, and Recife drew on this experience to develop its practices. In 2022, the strategy of regrouping students by learning levels consisted of reorganization of classes, at least once a week, with students from different grades based on their learning levels, targeting students in classes from 1st to 9th grades of elementary school. From 1st to 5th grades, it was mandatory, and the main criterion for regrouping was the literacy level of the students. From 6th to 9th it was not mandatory and also used additional criteria, such as performance in math.

Another strategy the Department used was digital classes to expand the workload. The pilot project to expand school hours by offering digital classes was implemented from October to December 2022 with the offering of two new subjects: Life Project and Innovation. The pilot was designed to test new models and formats and to design the curriculum for these new subjects. Initially, two models were adopted: one for full-time schools and another for part-time schools. Seven schools participated in the pilot project, reaching approximately 20% of the network's 6th to 9th grades students, varying the number of classes per school unit.

To focus the scope of the research, initially three interventions related to the learning recovery strategy were selected: (i) formative assessment, (ii) regrouping of students by learning levels and (iii) expansion of the workload through digital classes. Therefore, throughout 2022, we monitored these three initiatives with an adaptive evaluation methodology.

During the research, it was observed that the Pilot Project to expand the workload through digital classes had a weak relationship with the other Learning Recomposition strategies. In 2023, the Pilot Project was discontinued. Hence, the research focused on assessing the two other actions: formative assessment and regrouping of students by learning levels.

Between 2022 and 2023, other significant changes to the Learning Recomposition Plan were also made, with a combination of a change in leadership in the pedagogical team of the Education Department and learnings from the first year of implementation. In particular, the strategy of regrouping students became optional for schools, although all would still have to have a learning recovery plan. In addition, the number of SAERE formative assessments was reduced so that schools were not overwhelmed and had more time to use the data. The research challenge was how to assess what was an integrated learning recovery approach, given the complementarities across activities, especially between assessments, and pedagogical interventions.

4. Research design

In partnership with the Recife Education Department, CAEd and Imago Global Grassroots designed an "adaptive evaluation" approach to monitor and assess the policy of Learning Recomposition. This was undertaken over the two years of 2022 and 2023.

An adaptive evaluation was appropriate in this context due to the complexity of both the problem and the intervention. The **problem** was complex because learning levels depend on multiple factors, some related to the school, others to students' families, and to the broader context of the city, state and country, including, of course, the impact of the Covid pandemic. Beyond that, the **intervention** was also complex in design and implementation, and was specifically **not** designed as one intervention in a controlled setting. As noted, the policy approach evolved over time, going through significant changes especially between the years of 2022 and 2023. This meant that an experimental or quasi-experimental approach was neither feasible nor desirable. We instead applied an evaluation approach that allows for flexibility, exploration of variation across schools, and changes over time.

Adaptive evaluation is an approach designed to interpret complex processes of change and public action. Many development challenges are intrinsically complex, involving multiple interactions within a system with many moving parts and inter-relationships. Educational dynamics, and policies focused on improving learning are examples of such inherently complex processes. An adaptive evaluation approach seeks to support policymakers and practitioners to interpret change and to test and refine policies (Gokhale and Walton, 2022). Traditional evaluation methods, such as Randomized Control Trials (RCTs) and quasi-experimental methods, that can be very effective in assessing specific intervention effects, fall short in addressing the dynamic change processes integral to promote learning, innovation and scaling within a system.

The adaptive evaluation framework can be designed to provide feedback and continuous improvement, making it particularly suitable for dynamic and complex environments. Unlike conventional evaluations that often occur post-implementation, adaptive evaluation is ideally integrated into the policy implementation process, supporting design improvements over time. The approach typically involves dialogue and participation of the various stakeholders,

to inform the interpretation of system functioning, theory-building and assessment of implementation, or theory testing.

The flexibility of adaptive evaluation makes it applicable at any stage of an intervention's lifecycle. From initial system interpretation and change process analysis to rapid prototyping, testing, and adaptation, this approach is designed to accommodate the diverse and shifting contexts of development interventions. It can incorporate RCTs and similar techniques for analyzing specific processes but situates them within a broader, more holistic assessment and interpretation framework. This adaptability is crucial for achieving impact at scale, allowing interventions to be tailored and refined based on real-world feedback and evolving needs.

The adaptive evaluation is grounded in three primary strategies:

- **Systems Diagnosis:** This involves analysis of the structure, nature and behavior of the relevant system, including inter-relationships between actors, types of leverage points within the system (e.g. rules, resources, power relations, information flows, and mindsets), and the nature of system dynamics, including blockages and pathways for change.

- **Theory-Based Assessment of Change Processes:** This involves practical theory-building around potential pathways for change, including delineation of the sequence of processes, that can be tested through logical or, in some conditions, statistical tests. Assessment is typically through mixed methods, including both within-case and comparative case analysis.

- **Iterative Designs:** The approach both supports iteration in system interpretation and theory-building during the evaluation process, and can be used for structured experimentation with intervention partners over design choices. (Specific iterative design requires close collaboration with implementing partners and was not used in the Recife case.)

4.1. Application of Adaptive Evaluation in Recife

The adaptive evaluation in Recife was applied in a customized way, considering that the policies faced some changes during the application of the methodology, and its design is inherently flexible to the different contexts.

Empirical strategy

The research was divided into two parts. In the first, we undertook interviews with managers from the Recife Education Department, school principals and teachers for a general understanding of the Learning Recomposition Plan actions - first in August 2022 and later in January 2023. The purpose was to evaluate two distinct moments in the implementation of the Plan, to verify whether expectations were achieved and what obstacles had been faced during the period. In February 2023, a workshop was organized, with managers from the Recife Department, to consolidate this first part of the research.

In the second part of the research, interviews were conducted again with Department managers to identify changes and adjustments that policies underwent from one year to the next, considering the challenges faced and the achievements. Table 2 presents the professionals interviewed in the years 2022 and 2023.

Table 2 – Professionals of the Municipal Education Department Interviewed

Interviews with Professionals from the Municipal Education System
Secretary of Education
Executive secretary for projects, technology, and innovation
Executive manager of pedagogical management
Executive manager of pedagogical support
Manager of full-time schools and 6th to 9th grade
Literacy Program Manager
Early years and early childhood education manager
Inclusive education manager
Regional managers
Heads of pedagogical divisions of regional bodies

Coordinator of the nucleus of Assessment
Coordinator of the educator training school
Coordinator of the pedagogical coordination nucleus
Management teams and teachers from 3 schools

Source: CAEd/UFJF, 2023.

After identifying these changes, a sample of schools was selected so that the implementation of the policies could be monitored by field researchers, focusing on the 5th and 9th grades. Nine schools were selected, sampled to capture typical schools across different school types. Two of the nine schools in the sample have students from the 1st to the 9th grades; three are full-time schools and only serve students in the final years - 6th to 9th grade; and the other four serve only students in the initial years – 1st to 5th grades. All selected schools have socioeconomic indexes similar to Recife's average and performed in the SAEPE 2022 assessment close to the Recife average. On the other hand, in relation to 2021, the performance of schools varied, with some stagnating, while others experienced growth or decline. The purpose was to provide a sample that was broadly representative of the municipality but had some heterogeneity in terms of the recent learning dynamics. From a geographic point of view, the composition of the sample covers the four administrative regions of the municipality and represents the diversity of urban realities in Recife. Table 3 shows the list of schools and their characteristics (with school names removed).

Table 3 – Sample of nine schools selected for field research in 2023.

Full-time schools	Enrolled students	School grades	Socioeconomic level index[6]	SAEPE overall performance between 2021-2022 – 5th and/or 9th year
School 1	331	6th to 9th	III	Stagnation
School 2	321	6th to 9th	III	Improvement
School 3	407	6th to 9th	IV	Drop
Part-time schools	Enrolled students	Grades	Socioeconomic level index	SAEPE performance between 2021-2022 – 5th and 9th years
School 4	646	1st to 9th	IV	Improvement (5th year) and Stagnation (9th year)
School 5	404	1st to 5th	IV	Improvement
School 6	668	1st to 5th	IV	Drop
School 7	321	1st to 5th	III	Improvement
School 8	556	1st to 9th	III	Stagnation (5th year) and Drop (9th year)
School 9	371	1st to 5th	III	Improvement

Source: INEP, 2023; CAEd/UFJF, 2022.

As a methodological approach, we treated each school as an individual case, supporting a mixture of within-case, and comparative case analysis in the research. This was then placed within the overall quantitative context of patterns of outcomes across all schools in the Recife municipality. Empirical work involved field visits and interviews with teachers and school principals to understand the details and specificities of the implementation of the actions

[6] The socioeconomic level index for basic education is a Brazilian indicator calculated based on contextual questionnaires from the national assessment system (Saeb). These questionnaires are completed by students and include questions related to their families' educational backgrounds, ownership of household goods, and home infrastructure. The index is calculated at the school level as its minimum unit but can be aggregated to the entire Brazilian public education system. It ranges from Level I, indicating higher levels of vulnerability, to Level VIII, reflecting a more advantaged socioeconomic context (Inep, 2023a).

under analysis in each school context. Four cycles of visits were carried out in 2023, in each case implementing semi-structured interviews with school managers and teachers, as well as observation scripts for actions related to the Learning Recomposition Plan.

Table 4 – Research instruments applied in schools

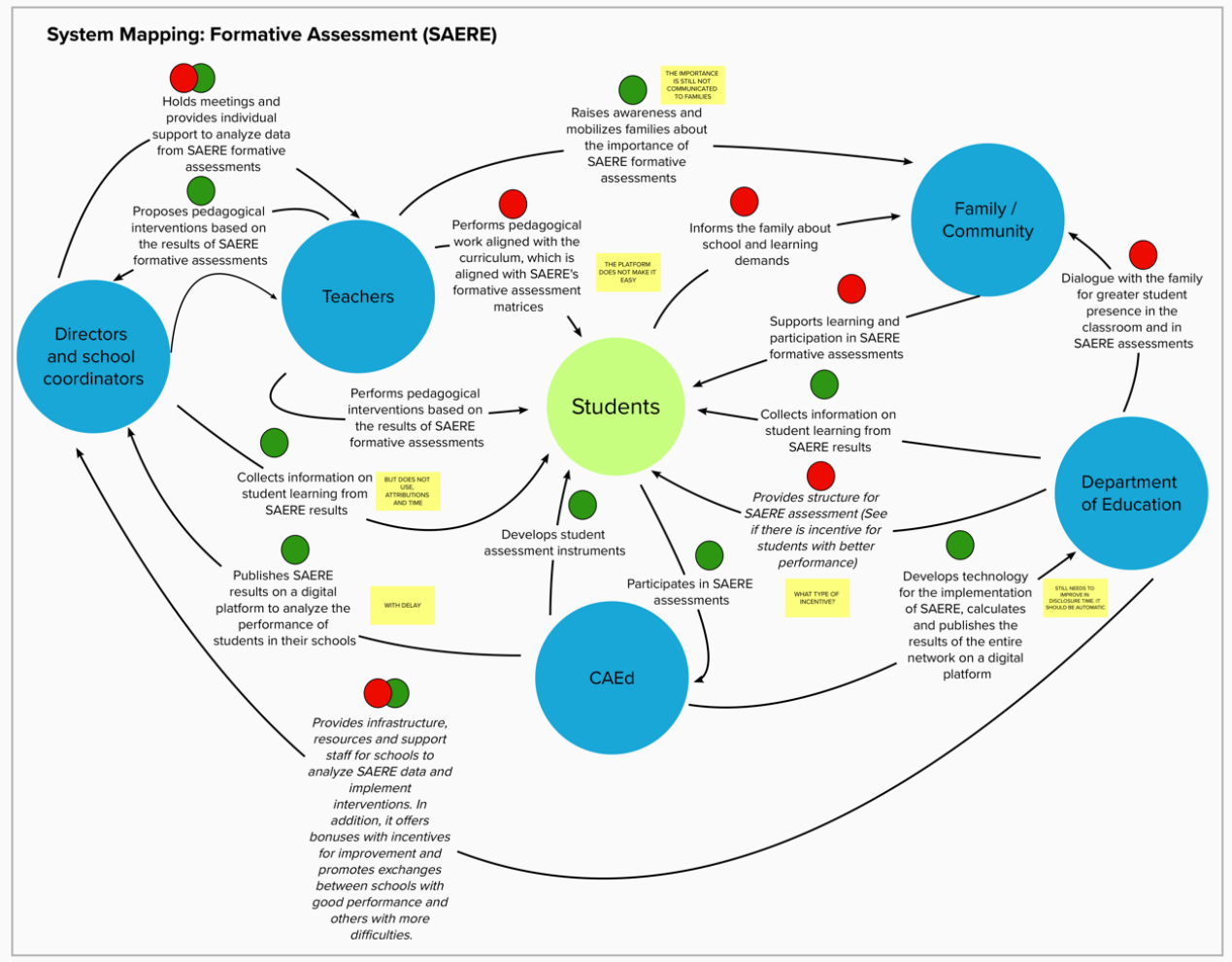
Application of research instruments in schools, 2023	
1st cycle of visits	School Directors' Questionnaire
2nd cycle of visits	Teachers and Management Team Questionnaire
3rd cycle of visits	Observation itinerary - Actions related to SAERE
4th cycle of visits	Observation itinerary - Recomposition actions

Source: CAEd/UFJF, 2023.

5. Findings

As part of the adaptive evaluation process, we developed both an initial system diagnosis and theory of change through participatory engagement and field visits. These form elements of the "findings" of the evaluation work. Figure 9 presents a visual summary of one part of the system mapping work that the evaluation team developed in a workshop with education professionals who worked at different levels in the Recife's Education Department in the beginning of 2023. This focused on the context of the formative assessment (SAERE), that was very similar for the work on the overall Learning Recomposition engagement. An initial version of the system mapping had been done by the research team before the workshop, based especially on the interviews with the different stakeholders, and during the workshop, it was refined

Figure 9 – Mapping of the system for the formative assessment System mapping in Recife – Formative Assessment (SAERE) – 2022 cycle



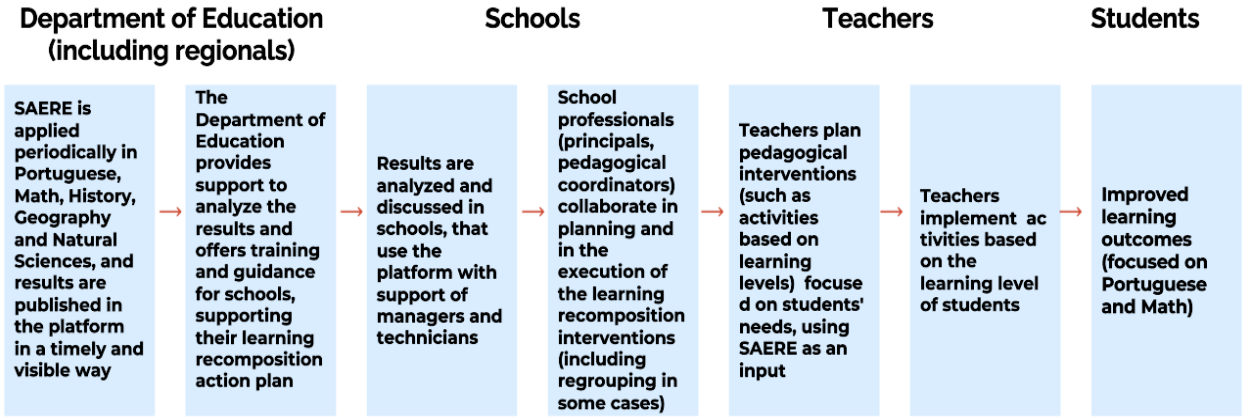
Source: authors, based on workshopping with the Recife Education Department

After mapping the actors and relationships within the system, the group developed an assessment of key "blockages" and promising pathways from the perspective of the intervention approach. The green dots represent processes that are working well in most schools and the red dots represent processes or relationships between stakeholders that can be improved. A combination of red and green indicates processes that work well in some schools but encounter challenges in others. For instance, one of the blockages identified was related to the support offered by the Education Department to school principals and pedagogical coordinators in terms of infrastructure, resources and teams for schools to analyze SAERE data and implement interventions. While challenges remained, there was important progress between 2022 and 2023, for instance with a reduction in the time that the

SAERE data were released to schools after students participated in the test, making them available in a timelier way to be used by principals and teachers.

The system mapping contributed to better understanding the roles played by key stakeholders, the relationships between them and the complexity of the implementation process. It also informed an initial theory of change for Learning Recomposition, and the role of SAERE within this. Workshop participants identified processes that were either working effectively or required improvement and detailed the Department of Education's strategy for tackling the blockages in the system. Participants used this to both delineate a theory of change, that included the role of SAERE, the Department's efforts, and processes for review and pedagogic action in the school, in line with the intentions and guidance of Learning Recomposition. This is summarized in Figure 10 below:

Figure 10 – Theory of change for the Learning Recomposition Plan in Recife



Source: authors, based on workshopping with the Recife Education Department

Based on the theory of change, an essential part of the application of the adaptive evaluation was doing a detailed process tracing. This involved using the empirical observations and interviews from each of the nine schools, to "test" the hypotheses underlying each step in the process in the theory of change. For each school, the steps of the theory of change were classified in relation to two dimensions. The first dimension involves an assessment of whether the step is working or not (i.e. the hypothesis to be tested is that it is working). If a specific step involves a structured experiment (as in A/B testing for instance), this assessment could be undertaken with a statistical test. However, this is commonly not the case, as in the Recife work, in which case the analysis is "within case" at the level of each step. In process tracing this is conceptually a logical test, as to whether observed patterns are consistent with the

hypothesis or could also be caused by alternative hypotheses.[7] A second dimension concerns the quality of evidence, for example whether it comes from reliable quantitative and/or qualitative sources, and how extensively the assessment is triangulated across different sources. As a practical matter, assessments at the level of an individual step in the TOC were classified in relation to four categories:

- **Confident – Working.** The assessment is clearly positive i.e., causal step is working and there are not plausible alternative hypotheses for the observed data *plus* sources are reliable and unbiased

- **Confident – Somewhat Working.** The result is mixed (i.e., for some it works, others it does not) and sources are reliable and unbiased.

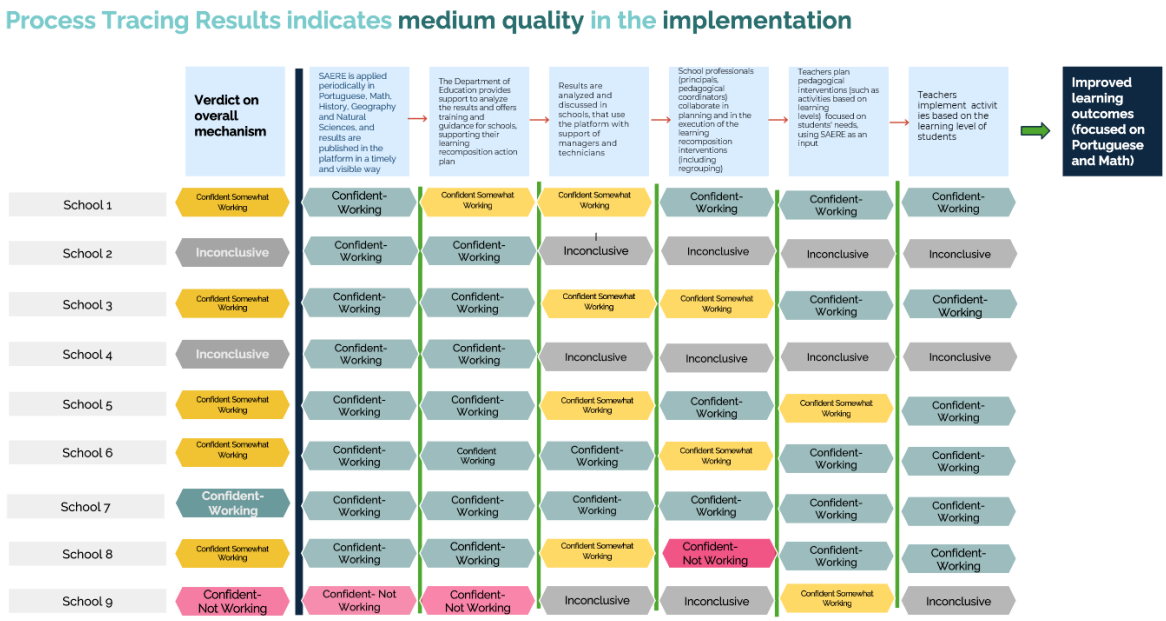
- **Inconclusive.** The result is unclear, or sources may not be or are not reliable and unbiased.

- **Confident – Not Working.** The result is clearly negative (i.e., causal step is not working) and sources are reliable and unbiased

The core empirical analysis was undertaken in depth for each of the nine schools, utilizing all the information gathered from the school visits in 2023. Selected material for three of the schools are in Annex B (and further details, including for the other schools, are available on request from the authors). Figure 11 below summarizes the results from all shows the consolidated process tracing for the nine schools.

[7] Process tracing practice in the literature typically classifies this logical assessment in relation to whether observational data supports or rejects the hypothesis within the process step in the TOC, and, in the case of support, whether it also rejects alternative hypotheses. See Beach (2020) and Raimundo (2020) for extended discussions.

Figure 11 – Process tracing on the implementation of Learning Recomposition for the nine focus schools



Source: authors, based on empirical work from field observations.

Steps 1 and 2 directly related to the Education Department were mostly classified as "Confident - Working"; Steps 3 and 4 related to the school level are more heterogeneous, with a predominance of "Confident - Somewhat Working" or "Inconclusive" observations; at the teacher level (Steps 5 and 6), a combination of "Confident - Working", "Confident - Somewhat working" and "Inconclusive". From the analysis of the process tracing combined with the fieldwork, some key takeaways are described below:

- **The approach of the Learning Recomposition was quite successful on the "supply" of initiatives from the department**. Most schools said that the Education Department applied the formative assessment SAERE periodically and provided support to analyze the results.

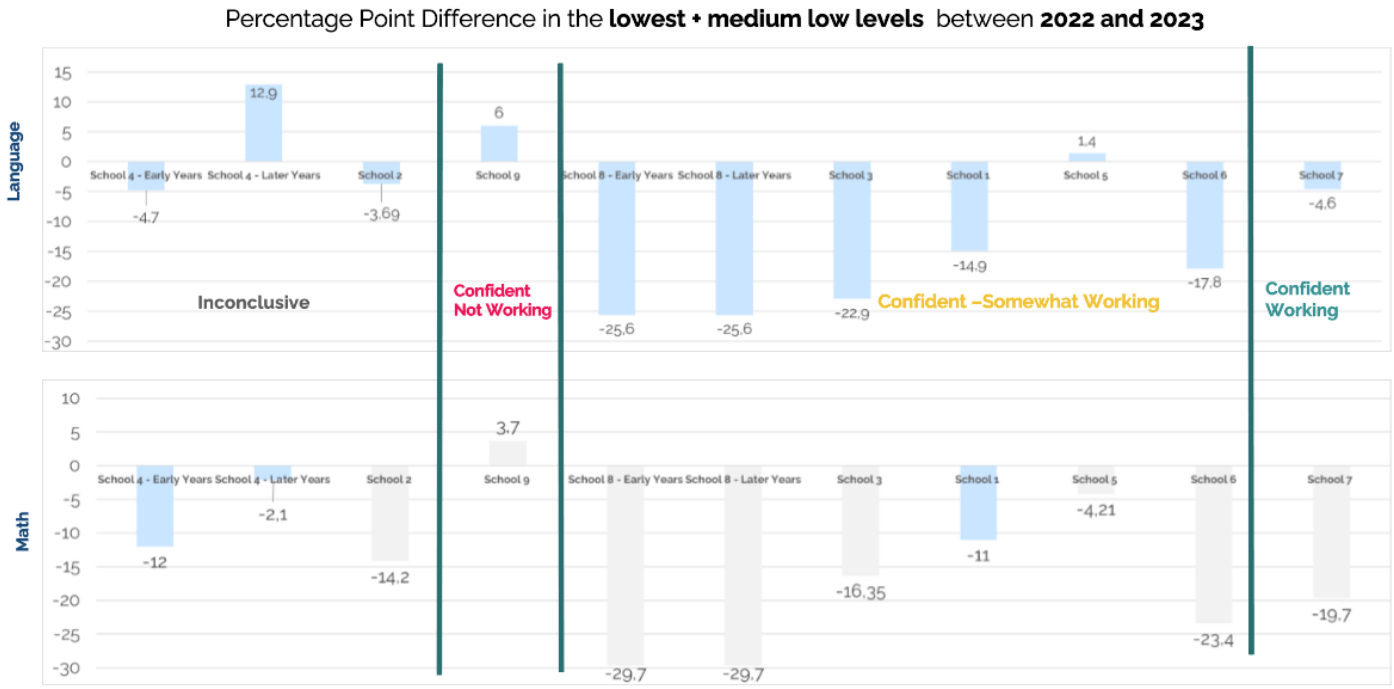
- **Steps in the implementation process that are critical to systemic coherence and consistent improvement of practices in schools are relatively fragile.** While results of SAERE are discussed in schools, their use to plan pedagogical interventions was limited, and there was no evidence of tailored support in this design. While there are cases in which the school has some support from the Education Department and individual teachers take the initiative of planning interventions according to students'

needs, there seems to be limited collaboration that promotes coherence and consistency in terms of pedagogical practices at the school level. A possible factor that could explain that is the lack of time and space for teachers to collaborate with each other and/or with pedagogical coordinators and principals, which leads to teachers planning their classes mostly individually. Another factor that could explain that is the timing in which results of the formative assessment was released – for instance, while school principals, pedagogical coordinators and teachers usually meet in the end of each bimester for the school council, they did not always had the results available to use these collective moments to discuss them.

The process tracing was then combined with the analysis of students' outcomes in the Pernambuco state assessment, SAEPE. The outcomes that are the focus of the analysis was the percentage of students in Elementary I and II - the lowest two categories of learning levels in SAEPE. Students at one of these two levels have not demonstrated adequate development of skills to the educational stage they are at. This was the focus because the learning recovery strategy aims to support students with low performance to improve. Considering that, in the figures below a more negative result is a positive development, since this represents a reduction in the percentage of students at the lowest levels of learning.

Figure 12 below shows the difference in the proportion of students in each school classified in the Elementary I and II levels in SAEPE, and the overall classification of schools according to the process tracing in the four categories - confident working, confident somehow working, confident not working or inconclusive.

Figure 12 – Percentage point different in the proportion of students in the two lowest levels of learning according to SAEPE between 2022 and 2023



Source: authors, based on empirical work from field observations, and SAEPE for learning outcomes.

From Figure 12, some key takeaways from the combination of the process tracing with student outcomes are described below:

- **There is a slight relationship between process tracing and some outcomes, but it at best explains part of the variation.**
    - The only school classified as confidently not working (school 9) stood out negatively for increasing the percentage in the lowest learning levels between 2022 and 2023.
    - With the exception of School 5, there is a set of schools that achieved substantial progress in reducing learning deficits, in which the implementation of Learning Recomposition was "somewhat working."
- **Most schools that stood out positively are full-time**
    - In these cases, students have access to longer instructional time, and teachers generally work in the same school all day and have more opportunities to

collaborate with each other, while in part-time schools the time dedicated for collaboration is significantly smaller.

- However, schools 6 and 8 were part-time and also stood out for their success in reducing the percentage of students in the lowest levels of learning between 2022 and 2023

In addition to assessing overall patterns in implementation and outcomes, we explored in greater depth positive and negative outliers (often referred to "positive or negative deviants" in evaluation terminology). We summarize here findings from the cases of schools 6 and 8 – that stood out positively - and school 9 – that was a negative deviant as being the only of the nine schools in which the proportion of students in the lowest two levels increased between 2022 and 2023. See Annex B for more information on these three cases.

School 6 – a positive deviant

School 6 in 2023 had 736 students, all of them between 1st and 5th grades. According to the national indicator of socioeconomic status, the school is classified as level 3, on a scale from 1 to 8[8]. At this level, the students are up to half a standard deviation below the national average of the indicator of socioeconomic status.

School 6 held collective events, with all students, aimed at engagement and preparation for SAERE - including rewarding students who performed best in the assessment and encouraging actions for others to achieve the same result in the next assessment. We did not identify this type of activity in any other school in the sample.

The same school also conducted regular meetings to analyze SAERE data with the participation of teachers. The school has made notable progress in learning recovery, with significant improvements in student participation and performance in SAERE evaluations. The school's structured approach and commitment to addressing learning gaps highlight its proactive stance in improving educational outcomes.

More specifically, the interventions that stood out in this school were:

[8] The average socioeconomic index of public schools in Brazil is 4.

- **Teaching assistants**: Implementation of the program, which provides an assistant to help with literacy, particularly for students in the 4th and 5th grades.

- **Regular Monitoring and Assessment**: Frequent assessments and monitoring of student progress, including analyzing SAERE results and adjusting teaching strategies accordingly. School professionals frequently used tools such as spreadsheets, Google Forms, and infographics to compare past and present SAERE data and plan further actions.

- **Teacher Collaboration and Training:** regular meetings for teachers to discuss and analyze assessment results and develop action plans.

- **Incentives and Motivational Activities**: rewarding students for their achievements in SAERE evaluations to motivate them. The school organized collective events dedicated to engaging students for SAERE and offered prizes for students who stood out positively in the test.

- **In-Class and regrouping learning Recovery:** each teacher was responsible for the learning recovery of students in their own classroom and regrouping students based on their learning levels for targeted instruction within the classroom using differentiated materials provided by the Education Department twice a week.

<u>School 8 – a positive deviant</u>

School 8 in 2023 had 287 students between 1st and 5th grades, and 238 students from 6th to 9th grades. According to the national indicator of socioeconomic status, the school is classified as level 3, on a scale from 1 to 8. The school serves socially vulnerable students, some of them requiring support from the Child Protection Council and psychological support.

School 8 made a change in 2023 to its learning recovery plan, that ensured greater engagement of its professionals. Instead of maintaining the 2022 model, the school adopted, in 2023, the strategy of regrouping students within each class into different stations and/or work groups. To this end, students underwent a diagnostic assessment to identify learning levels and teachers monitored students' progress and adapted the activities according to the class's performance.

Furthermore, teachers consolidated information related to their school Learning Recomposition plans and shared it with the pedagogical coordination every two months. In 2023, School 8, whose students began the year with performance below the Recife average, had a considerable increase in proficiency in SAERE and SAEPE assessments in Portuguese and Mathematics among students in the 5th and 9th years of elementary school.

The school's learning recomposition plan, initially resisted by teachers, improved student outcomes through regular reassessment and regrouping according to student's needs. Formative Assessment (SAERE) results guided pedagogical interventions, though teachers struggled with data interpretation. Despite the relative success, according to school professionals interviewed during the research, more meetings and psychosocial support were needed to address educational and emotional challenges.

More specifically, the interventions that stood out in this school were:

- **Regular Evaluation, monitoring and focused interventions based on learning levels**: School professionals utilized SAERE results to identify learning gaps and plan targeted interventions based on the analysis of results collectively in bimonthly council meetings and used the Recife Platform for accessing and interpreting assessment data. Additionally, based on assessment results, the school held remedial classes, reallocating time within the school schedule for focused learning recovery sessions, regrouped students based on learning levels and developed pedagogical strategies tailored to students' needs.

- **Family and Community Engagement:** The school maintained close relationships with the community and sought to involve families in the learning process, despite challenges in achieving consistent participation.

- **Environment of support for Teachers and Students**: The school team addressed the social and emotional needs of students by seeking additional support, such as psychologists, to help manage family-related challenges. Additionally, it encouraged a collaborative approach among teachers to align their efforts and share best practices.

- **Infrastructure Improvements:** the school continued with ongoing school reforms, despite the associated disruptions, to create a better learning environment. It provided climate-controlled classrooms and clean, organized facilities to support a conducive learning atmosphere.

- **Change in regrouping strategies.** While in 2022 the school regrouped students from different school classes, in 2023 grouping of students with similar learning needs happened inside the same classroom. Regrouping happened twice a week, lasted 2 hours, and was a strategy used for students between 3rd and 5th grades. The change in strategies helped to reduce resistance of teachers and increase engagement.

School 9 – a negative deviant

School 9 in 2023 had 382 students between 1st and 5th grades, and 22 pre-school students. According to the national indicator of socioeconomic status, the school is classified as level 3 (on a scale from 1 to 8).

The school relies significantly on government programs and has an important role in addressing socio-economic challenges faced by students. It tries to provide a supportive environment to help them cope with external stressors, such as family issues and community violence. This challenging context might explain partially why they had lower performance than the other schools.

Some interventions that stood out in this school were:

- **Teaching assistants:** a national literacy program assists 1st and 2nd-grade teachers with teaching assistants who help students either within the classroom or by taking them out for individualized support.

- **Use of External and Internal Assessments:** The school utilizes results from both the SAERE assessments and internal evaluations to guide instructional strategies and interventions.

- **Teacher Training and Collaboration:** Regular meetings and planning sessions are held with teachers to discuss assessment results and devise strategies to address identified learning gaps.

- **Specific periods for Learning Recomposition:** Learning Recomposition activities are conducted twice a week for students in 3rd to 5th grades. Additional support is provided by the librarian and the pedagogical coordinator.

- **Focus on reading and writing:** The primary focus of the recomposition activities is on literacy, particularly reading and writing skills, as difficulties in understanding written content were linked to poor performance in mathematics.

- **Parental Involvement:** efforts are made to inform and involve parents in the Learning Recomposition process, although the school faces challenges in achieving consistent parental engagement.

Analysis of positive and negative deviants

By analyzing the interventions in the school, they seemed to be similar to the other schools, but with some noticeable differences.

In our sample of schools, we realized that internal meetings to analyze data do not always happen systematically. Another important point of attention refers to the teachers' participation in these meetings, which has been more restricted to school directors and pedagogical coordinators. As we know, the participation of teachers in internal school meetings to analyze data favors the pedagogical interpretation of results and the sharing of practices, which can lead to the implementation of successful interventions (Boudett et al., 2020).

Case studies have pointed to this as an important element of Education Departments that have built a routine of actions based on evidence and achieved good results (Palacios & Bonamino, 2023, 38-39). This means a change in the social place occupied by large-scale assessment in the Brazilian educational system: from an instrument of external control of schoolwork, the assessment becomes a fundamental tool in everyday school life. (Burgos, 2020).

In the schools that stood out positively, they combined actions that enhanced the use of interventions directed to the whole school system – such as having consistent routines and processes to analyze the results of the formative assessment – and specific innovations that were designed locally – such as prizes for students who achieved high scores in the test.

A few specificities that indicate that the learning recovery policy did not work as effectively in the school is the perception of the school principal that results of assessment were not so useful considering the school's reality, and that the school did not regroup students based on learning needs.

In schools that were positive deviants, policies offered by the Education Department were adapted and reinforced according to their contexts with specific interventions created at the local level. In the school that was a negative deviant, some of the policies seemed to be treated mostly as a burden, and not adequate to the local reality.

While relevant, these differences in implementation do not seem as significant to explain all the differences in performance compared to other schools. So, contextual factors beyond the pedagogical interventions seems to explain part of the results observed as well.

Therefore, the variation in results seems to be explained by a combination of school level differences unrelated to the implementation and the way the schools viewed the policy – either as a support that helped to improve management and pedagogical practices or as a burden that schools thought obligated to comply with but did not see as helpful to improve its results.

6. Interpretation

The analysis in the previous section indicates there was considerable heterogeneity in the implementation of the learning recovery policies. In addition to that, there were significant differences between the implementation in 2023 compared to 2022, with improvements made based on lessons learned from what worked well and aspects to improve (see Table 5).

Table 5 – Challenges in the implementation of the Learning Recomposition Plan faced in 2022 and adjustments made in 2023

Challenges faced in 2022	Adjustments made in 2023
Excessive assessments applied throughout the year	Reduction in assessments, giving more time to schools to use the results and leaving more instructional time
Logistical problems with internal and external applicators that were responsible for executing digital and printed assessments undermined credibility of results	Standardization with only printed tests and external applicators increased trustworthiness of results
Long time in releasing assessment results	Faster publication and dissemination of results
Problems in communication for returning results and little support for schools to use the data	Collective presentations of results to schools and additional support to use data in everyday school life
Fixed model of regrouping between classes, mandatory in the initial years and optional in the final years (6th to 9th grades)	Schools are free to create their learning recovery plan, with guidance and support from the Department and flexibility
Adherence in elementary school (1st to 5th grade), but resistance and difficulties in the final years (6th to 9th grades)	Beyond having more flexibility, schools are encouraged to share their plans at events with each other to learn collectively, contributing to minimize resistance
Logistical difficulties with times and spaces available to carry out regroupings between classes	Reports of greater satisfaction with implementation flexibility and exchange of experiences

Source: authors based on interviews with government officials.

In 2022, the main challenges identified by the research in implementing the Learning Recomposition plans referred to the difficulties in adopting a fixed model of regrouping between classes proposed by the Education Department. As seen in the previous sections, the action was mandatory between 1st and 5th grades and optional between the 6th and 9th grades. In the school sample, logistical difficulties with schedules and spaces available to carry out regroupings between classes were mentioned as an obstacle to the adoption of the proposed actions. In general, the schools understood the importance of actions focused on learning gaps but pointed out difficulties in adopting a fixed model of student regrouping and, in some cases, complained about the lack of support from the Department.

Despite the criticism of the fixed model, it was interesting to note that, in 2023, the plans were not that different from what had been proposed in 2022. In fact, many schools seemed to follow the same model, while others made small adjustments to their reality. In the end, the plans were divided between a model of regrouping between classes and another of regrouping within classes. This second would be an alternative for schools that face obstacles in regrouping students from different classes and school years, whether due to teachers' difficulties or space limitations.

This demonstrates the importance of educational reform processes being conducted in a systematic way without leaving aside discretion, allowing for flexibility according to diverse realities while aiming for coherence (Fullan, 2010). On the other hand, it also demonstrates that, as authors such as Hargreaves (2009) argue, government interventions and policies are hardly capable by themselves of sustaining changes in education, but rather involve people working together as a team around common purposes and clear objectives that have a real meaning for them. The analytical perspective aimed at reconciling large-scale educational reforms with an approach capable of guiding the process of change based on the involvement of professionals who work at the “street level” gains strength, seeking new forms of balance between top-down and bottom-up approaches (Burgos and Bellato, 2019).

In 2022, the reforms were promoted in a way more associated with the "2nd generation of educational reforms in Brazil" (that have achieved limited progress in learning outcomes) according to the classification of Filho (2022), while in 2023 they made some progress toward the 3rd generation of reforms, that have achieved more significant progress in learning outcomes), as detailed in the table below:

Table 6 – Characteristics of the 2nd and 3rd generation of reforms

Characteristics of less effective reforms (associated with the 2nd generation of reforms)	Characteristics of more effective reforms (associated with the 3rd generation of reforms)
Focus on standards, targets and incentives as key levers of reform	Standards, targets and incentives are articulated with investments in inputs that strengthen the system's capabilities
Highly centralized management	Shared systemic view with coordinated decentralization
More uniform policies and with short time span	Systemic reforms based in intensive projects at the school level and with lasting implementation

Source: Filho (2022). Adapted and translated from tje Portuguese.

Based on the experience of Recife and the adaptive evaluation process, some of the key insights learned that are more broadly applicable to different education systems. Here, we divided these into two categories: reflections related to the implementation of public policies on on education; and reflections on the use of adaptive evaluation to improve these policies.

6.1. Reflections related to the design and implementation of education policies

Giving flexibility and support for schools contributes to making implementation adapted to diverse contexts and enhances legitimacy. In 2022, there was a perception of some schools that the regrouping of students by learning levels was implemented in a top-down way without considering the diverse needs and contexts of the municipality. In contrast, in 2023 the Education Department gave more flexibility for schools to choose whether to use the regrouping methodology or not. It was well received by schools, as the new format gave them more freedom to choose whether it was appropriate or not to their context, and if so, adapt to their own needs. As a result, resistance to the project declined, and at the same time a considerable portion of the schools chose to continue using the methodology, according to interviews with the Department. However, flexibility alone is not enough to promote significant improvements to education policies and practices. The support offered by the Education

Department, for example with training for teachers and other education professionals involved in the process (such as principals and pedagogical coordinators), and regular visits by professionals of the regionals, focused on helping the implementation over time. Additionally, incentives and communication about the benefits of the projects are essential so that schools understand why they can contribute to the work of schools in improving students' learning and are associated with the accountability mechanisms that exist. Without support and incentives, flexibility alone might lead to projects not implemented on the ground. Therefore, the effective implementation of large-scale education policies requires a combination of support and pressure from the central team and flexibility for schools to adapt the projects according to their reality.

Gradual implementation through iterative design makes it possible for the school system to learn in the process and adjust their habits at a rate they can absorb. In 2022, there was a perception of some schools that the projects (in particular the regrouping of students) was implemented too fast, without giving enough time for schools to get prepared and feel more comfortable to do it. Over time, more schools understood the importance of the project and were able to implement it well, but the initial resistance of schools reduced the projects' legitimacy. A lesson that can be extracted from this experience is that, especially in the cases of complex projects that require significant changes of practice and habits, a gradual implementation, that increases both the schools capacity to implement the project and the legitimacy of the project through iterative adaptation is more likely to generate improvement in results and sustainability over time as this approach allows for schools to change their habits in a more manageable way. Additionally, involving more the different stakeholders that will be responsible for the design and the implementation of the policy also helps in increasing legitimacy and effectiveness of the policies as they feel part of the process, understand better why some decisions were made, and how to implement policies effectively.

Prioritization is crucial to concentrate resources on what can generate more impact. For example, in 2022 many schools complained about the excess of assessments, as they did not have enough time to carefully analyze data and use it before the next assessment. Based on the feedback received, the Recife Education Department adjusted it, reducing the number of assessments in 2023. This is one illustration of a broader challenge faced by Recife and other school systems, in which schools can feel overwhelmed when there are many projects proposed by the Education Department at the same time. It is a case of what Andrews,

Pritchett e Woolcock (2017) call premature load bearing – when change demands are introduced before they can be managed by the organization. Because of that, it is crucial to prioritize better and focus the resources such as time and budget in a few actions that have more potential of impact, rather than dispersing them in various actions.

Coherence and alignment are challenges that can be mitigated by professionals that have the role of bridging different teams around a common purpose and routines that promote frequent collaboration. A common challenge faced by education systems around the world is the lack of alignment between different elements of the education policy - for example, curriculum, instructional materials, teacher training and assessment – and coherence between what is defined as policy and what happens on the ground because of a lack of a shared understanding and commitment across all levels of the educational system—teachers, administrators, policymakers, and communities—toward achieving common goals (Fullan & Quinn, 2016). Part of this problem is related to the way education departments are organized, in different areas that might not collaborate with each other. In the case of Recife, there was one area in the Department of Education, whose leader (Manager of Pedagogical Support) was responsible for promoting coherence in the pedagogical policy, by fostering collaboration between different stakeholders with the common purpose of improving students' learning. This model of having people responsible for bridging different teams is a possible way to enhance coherence in education policy by promoting frequent collaboration. Coherence is also crucial at the school level: the observations from the researchers who did field visits indicate that in schools in which the projects were conducted in a fragmented way, without the coordination of the principal or pedagogical coordination, the implementation tended to be less effective.

For data (quantitative and qualitative) to be used effectively to guide decisions it needs to be timely and reliable. Timely data is important both for quantitative and qualitative data.. An example of quantitative data is the results of the formative assessment conducted by CAEd. In 2022 it took more time for the data to be processed and schools to have the summary of their students' performance in the platform, which made it more difficult to use them to understand students' needs and based on that guide learning recovery actions. In contrast, in 2023 the time of data processing was shortened, which contributed to increase the use of data in a timely way to guide interventions to improve students' learning. Regarding qualitative data, an example is related to its collection and analysis as part of the adaptive evaluation

methodology. One of the moments of data collection and consolidation was in January 2023, that is a month in which education systems in Brazil plan the year that is starting. After that, there was a workshop together with different people from the Education Department in February. Concentrating the process of data collection, analysis and use of the data in the beginning of the year was important so that the insights learned during the adaptive evaluation could be used to improve the policy design and implementation.

6.2. Reflections on the use of an adaptive evaluation approach to improve public policies

Close collaboration between the teams responsible for the adaptive evaluation and the government is crucial. For the researchers who are conducting the adaptive evaluation, collaboration of civil servants is essential to understand the context, the reasoning behind the projects, the process of implementation, the different stakeholders involved, know how to obtain data and better interpret the data analysis. For government officials, collaborating with researchers can be valuable to bring an external perspective and new ideas that can be helpful to improve the policies, as well as to have people with time dedicated to organize data collection and analysis in such a way that in general is not possible without external support because of the many responsibilities and time constraints associated with it that civil servants in general have. In the case of Recife, during some points of the adaptive evaluation in which researchers and government officials worked together more closely, both parties benefited more than in situations in which the connection was more distant. Therefore, finding ways to enhance the proximity between the teams of researchers and of the government, as it happened for example during the fieldwork, is crucial for the adaptive evaluation to be more effective in generating valuable insights and using them to improve public policies.

Adaptive evaluation can facilitate and accelerate the process of collective learning in the organization. The whole process of conducting the adaptive evaluation contributed to facilitating learning in the Recife Education Department and enhancing the team's capacity over time. For example, in the workshops, different stakeholders from different parts of the department worked together, going beyond their day-to-day activities, to reflect on the projects and how to improve them over time from a broader perspective. In the interviews that were part of the process, civil servants had the opportunity to reflect more deeply on the projects. Additionally, the consolidation of results of the adaptive evaluation in different points

of the implementation of the projects contributed to organizing what different people learned in such a way that strengthened collective learning. In summary, adaptive evaluation accelerates learning and the capacity of the team to continuously improve their projects over time.

System diagnosis contributes to understanding the complexity of implementation and building policies that are more sustainable. Mapping the different stakeholders and their interactions through the system diagnosis helped to understand the complexity of implementation and ensure that the different needs and interests are taken into consideration. Additionally, system diagnosis contributes to designing and implementing public policies in a way that is more sustainable over time. For example, in Brazil, it is common for public policies to be discontinued when governments change. However, if public policies are designed and implemented in close collaboration between different stakeholders, such as teachers, school principals, civil servants who work in different parts of the Education Department and civil society organizations, it is possible to build coalitions that help in the continuity of public policies.

Process tracing can help to focus monitoring on the most crucial aspects of implementation. Process tracing, by organizing the steps in the process of implementation of each project, as well as the hypothesis associated with them, made it possible to identify what data was already available and what was missing to better understand the causal links between actions and the expected results. During the process of adaptive evaluation, as more data was collected and different parts of the projects unfolded, some hypotheses were confirmed, while other dimensions required more data, which clarified what crucial data should be the focus of next monitoring visits. This process of defining hypotheses clearly and testing them gradually contributed to organizing the process of learning during implementation in different phases of maturity of the project. For example, in the case of Recife, some of the hypotheses related to schools' knowledge of the assessment mechanisms were confirmed in the first weeks of monitoring by field visits by researchers. However, regarding hypotheses related to more complex steps of the implementation process, such as using data to inform pedagogical interventions, there were mixed results in the beginning, which required further research to understand what challenges schools were facing in using data and see how it evolved over time. Additionally, the observation of

implementation over time indicated parts of the process that required more attention and support from the Education Department, either by the central or the regional teams.

7. Conclusion

This paper about the case of Recife illustrates how an adaptive evaluation approach can be used in dealing with complex development challenges such as improving students' learning. The methodology contributes to facilitating collective learning, which is especially important when dealing with complex challenges in which there are multiple unknowns regarding the problem, its causes, and potential solutions and their impact. Beyond that, it is a flexible methodology that can adapt to the different needs and circumstances of different governments.

Something similar can be said about Recife's learning recovery policies. The research highlighted the importance of combining flexibility with clear guidance and support in order to find the right balance: giving schools the autonomy to adapt policies and practices to their specific contexts, while ensuring they are not left alone in the process. This balance, which was strengthened in 2023, appeared to be key to greater acceptance of the policies and improved results in Recife's municipal school system. At the same time, the diversity in how schools implemented the policies – and the variation in outcomes – suggests that further investigation is needed, as other factors influencing successful implementation remain to be identified.

In the context of Recife, this capacity to adapt was especially important given that the learning recovery policies were designed and implemented to deal with an unprecedented challenge which was the impact that the COVID-19 had on students' learning. If in other circumstances the complexity of the education system was such that it required constant learning, during a context of crisis it was even more necessary given the high uncertainty and lack of prior knowledge about how to deal with such unprecedented situations. Beyond that, during the process of adaptive evaluation Recife went through important changes in leadership, which also reflected in changes of the learning recovery policy between 2022 and 2023. The adaptive evaluation, by including systematic monitoring and reflection on the implementation, contributed to identifying changes that happened in schools as result of the policy changes.

Given its potential to systematically facilitate learning, the adaptive evaluation can be used by governments at different levels and to address challenges in diverse policy areas, contributing then to enhance state capability to address complex development problems, potentially in collaboration with other organizations, thereby contributing to more sustainable and impactful public policies.

References

Andrews, Matt; Pritchett, Lant; Woolcock, Michael (2017). Building State Capability: Evidence, Analysis, Action. Oxford: Oxford University Press.

Beach, D. (2020). Process tracing methods (pp. 699-719). Springer Fachmedien Wiesbaden.

Boudett, K. P., City, E. A., & Murnane, R. J. (Eds.). (2020). Data wise, revised and expanded edition: A step-by-step guide to using assessment results to improve teaching and learning. Harvard Education Press.

Burgos, M.; Bellato, C. (2019). "Gerencialismo e pós-gerencialismo: em busca de uma nova imaginação para as políticas educacionais no Brasil". Sociologia & Antropologia, v. 9, pp. 919-943.

Burgos, M. (2020). A avaliação externa e os novos sujeitos da educação. Pesquisa E Debate Em Educação, 10(1), pp. 1082–1102.

CAEd/UFJF. (2023). Plataforma de Avaliação e Monitoramento de Pernambuco. Resultados do SAEPE, from https://avaliacaoemonitoramentopernambuco.caeddigital.net/

Filho, O. (2022). Pontos fora da curva: por que algumas reformas educacionais no Brasil são mais efetivas do que outras e o que isso significa para o futuro da educação básica.

Fullan, M; Quinn, J. (2016) Coherence: the Right Drivers in Action for Schools, Districts and Systems. Thousand Oaks: Corwin Press.

Fullan, M. (2010). Positive pressure. In Second International Handbook of Educational Change. [s.l.]: Springer Science + Business Media (Springer International Handbooks of Education, 23), p. 119-130.

Gokhale, S. and Walton, M. (2022). Adaptive Evaluation: A Complexity-Based Approach to Systematic Learning for Innovation and Scaling in Development. Harvard University Center for International Development. Available at: https://www.hks.harvard.edu/sites/default/files/centers/cid/files/publications/faculty-working-papers/428_Gokhale_Walton.pdf [Accessed 13 November 2024]

Hargreaves, A. (2009). The fourth way of change: towards an age of inspiration and sustainability. In Change wards (pp. 11-44). Bloomington: Hawker Brownlow Education.

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP. (2022). Censo Escolar, from https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-escolar/resultados/2022

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP. (2023). Nível Socioeconômico (Inse), from https://www.gov.br/inep/pt-br/acesso-a-informacao/dados-abertos/indicadores-educacionais/nivel-socioeconomico

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP. (2023a). Indicador de Nível Socioeconômico do Saeb 2021 – Nota Técnica. INEP/MEC.

Palacios, C., & Bonamino, A. (2023). Brazil. How Two Municipalities Achieved Above-Average Results in Reading in the Early Years of Elementary School During the COVID-19 Pandemic. In Schools and Society During the COVID-19 Pandemic: How Education Systems Changed and the Road Ahead (pp. 27-42). Cham: Springer Nature, Switzerland.

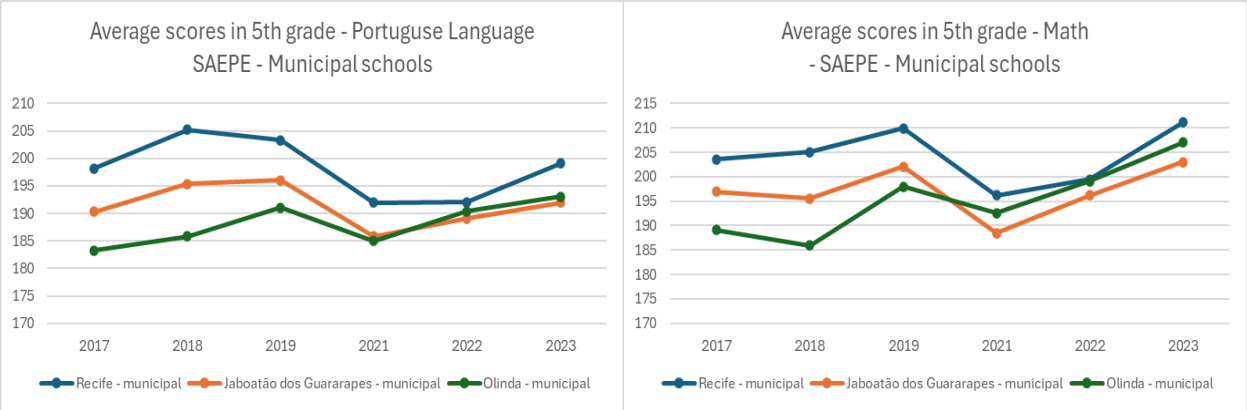
Pernambuco. Secretaria de Educação e Esportes. (2022). SAEPE - 2022. Revista da Gestão Escolar, 1, Juiz de Fora.

UNESCO. (2022). Database of duration of school closures. Retrieved from https://covid19.uis.unesco.org/wp-content/uploads/sites/11/2022/09/SDG-duration-of-school-closures-by-country.xlsx

Raimondo, E. (2020). Getting Practical With Causal Mechanisms: The application of Process-Tracing Under Real- World Evaluation Constraints. New Directions for Evaluation, 2020(167), 45-58

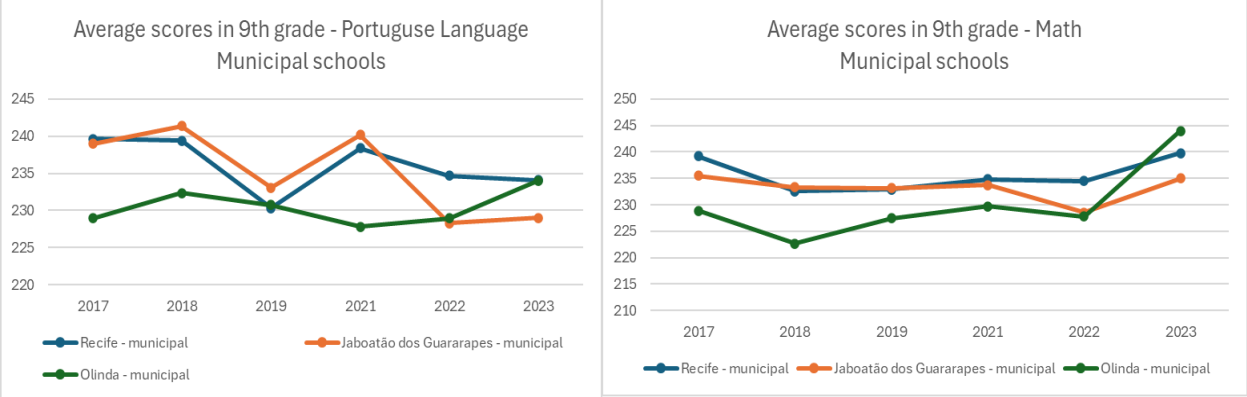
Annex A. Additional data

Figure A.1 – Proficiency in the 5th grade – Comparison between Recife, Olinda and Jaboatão dos Guararapes



Source: Caed/UFJF

Figure A.2 – Proficiency in the 9th grade – Comparison between Recife, Olinda and Jaboatão dos Guararapes

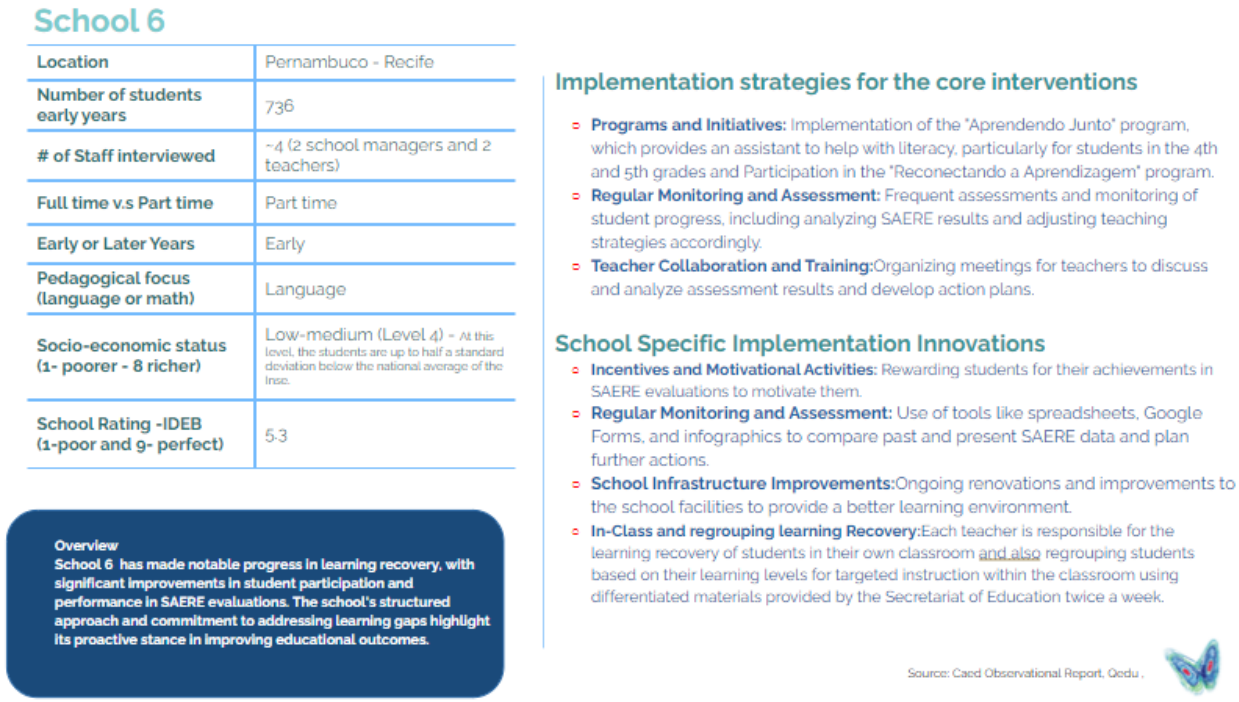


Source: Caed/UFJF

Annex B. Examples of detailed process tracing for three of the 9 schools

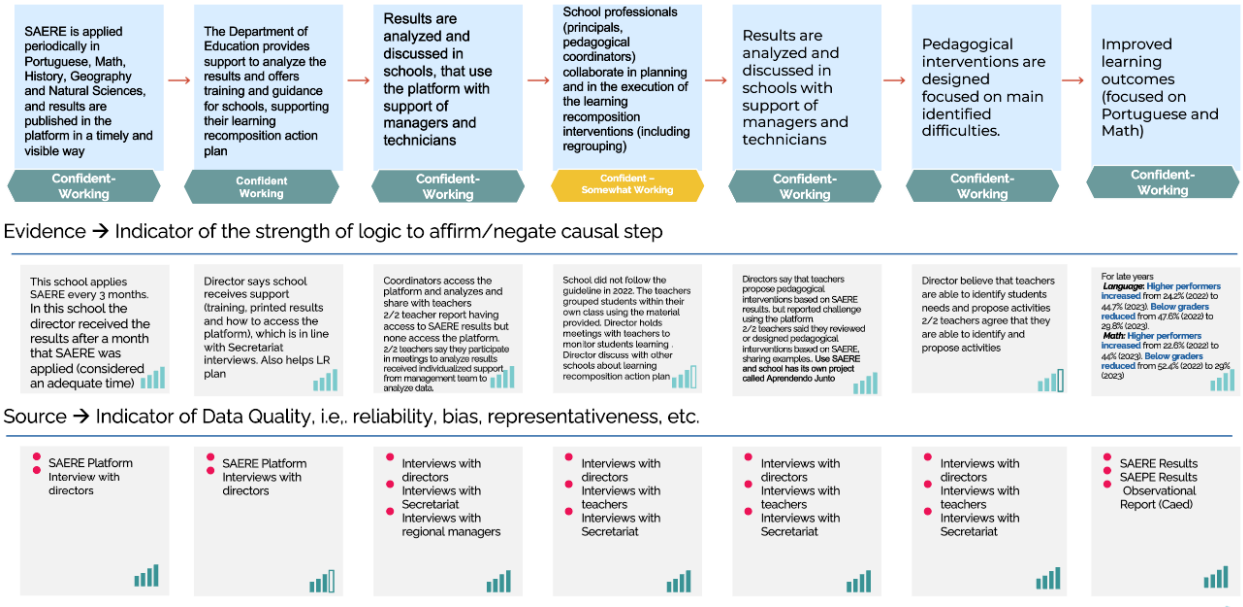
For each of the 9 schools that were monitored more closely, a detailed analysis was performed about their characteristics, the process tracing, and results in assessment over time. Below, we present examples of analyses performed for schools 6, 8 and 9. Similar analysis was done for the other six schools are available on request.

Figure B.1 - Analysis of school characteristics and interventions – School 6



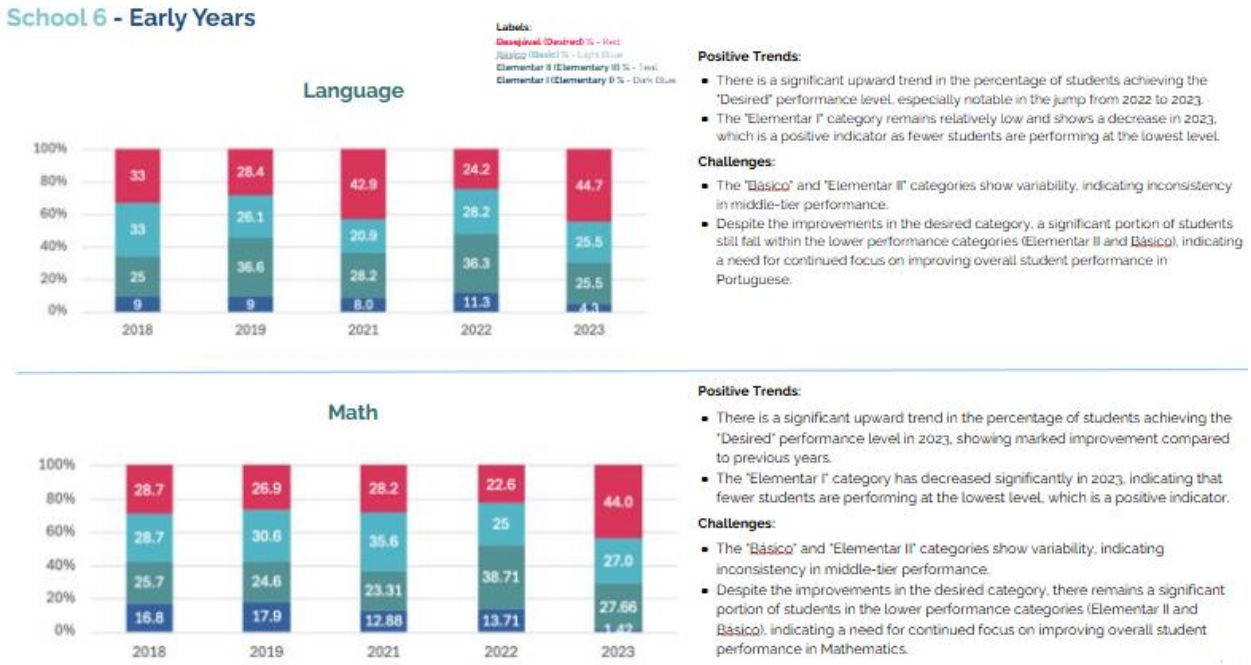
Source: Caed Observational Report, Oedu

Figure B.2 - Analysis of process tracing for School 6



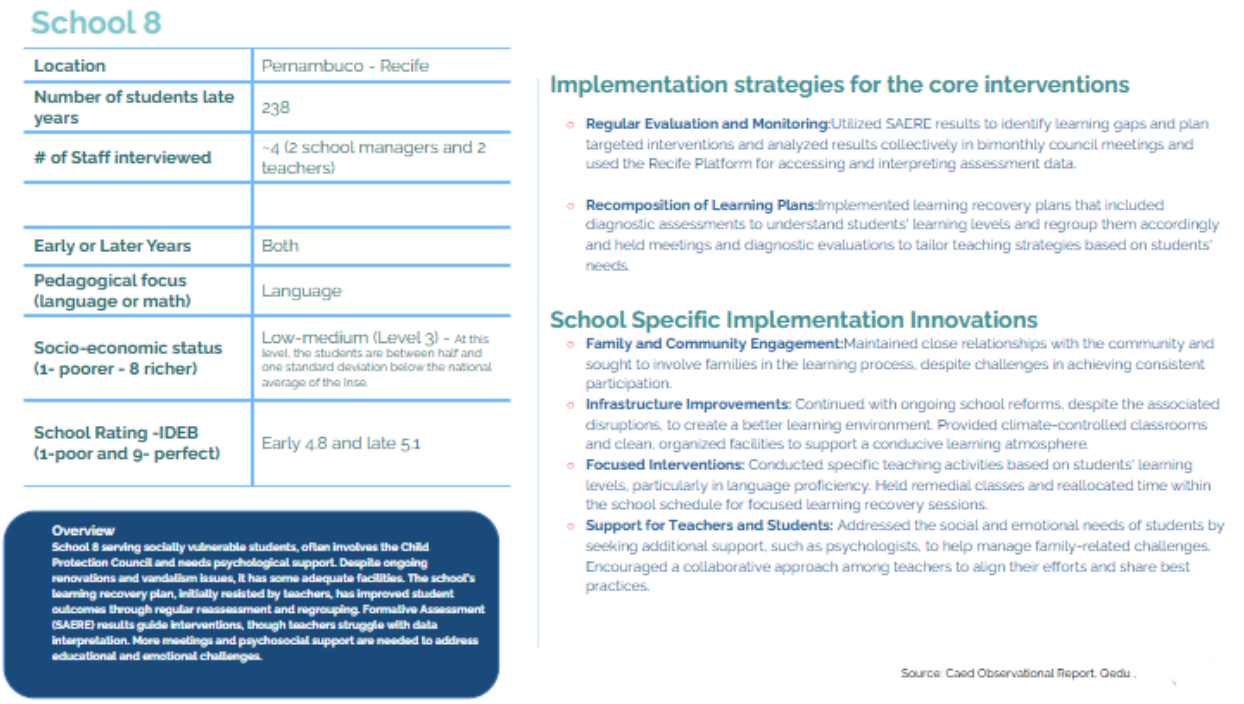
Source: Authors

Figure B.3. - Analysis of learning outcomes for school 6 – Early Years (1st to 5th grade)



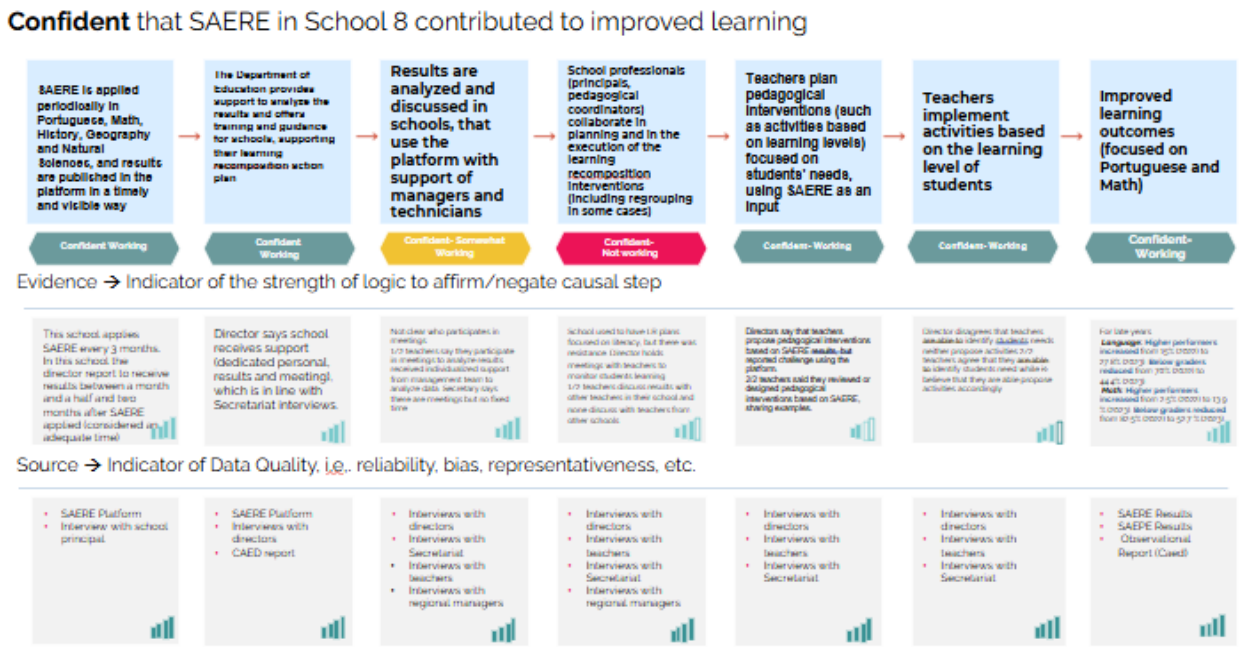
Source: Authors

Figure B.4 - Analysis of school characteristics and interventions – School 8



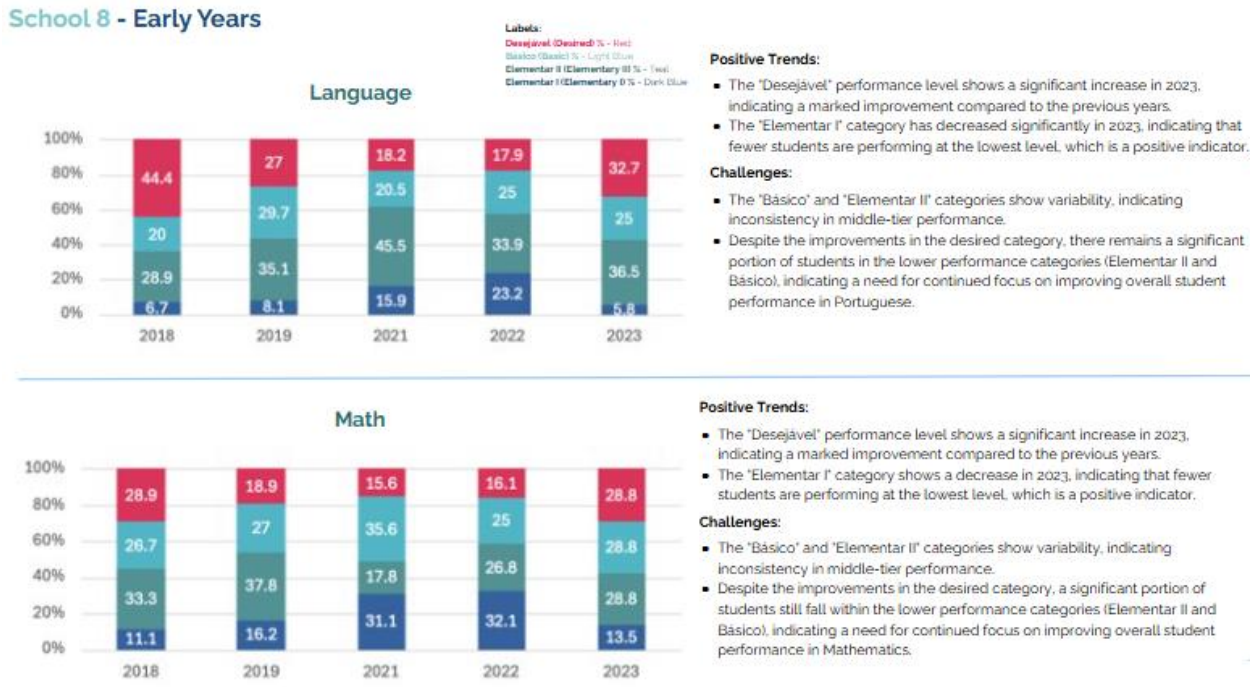
Source: Caed Observational Report, Oedu

Figure B.5 - Analysis of process tracing for School 8



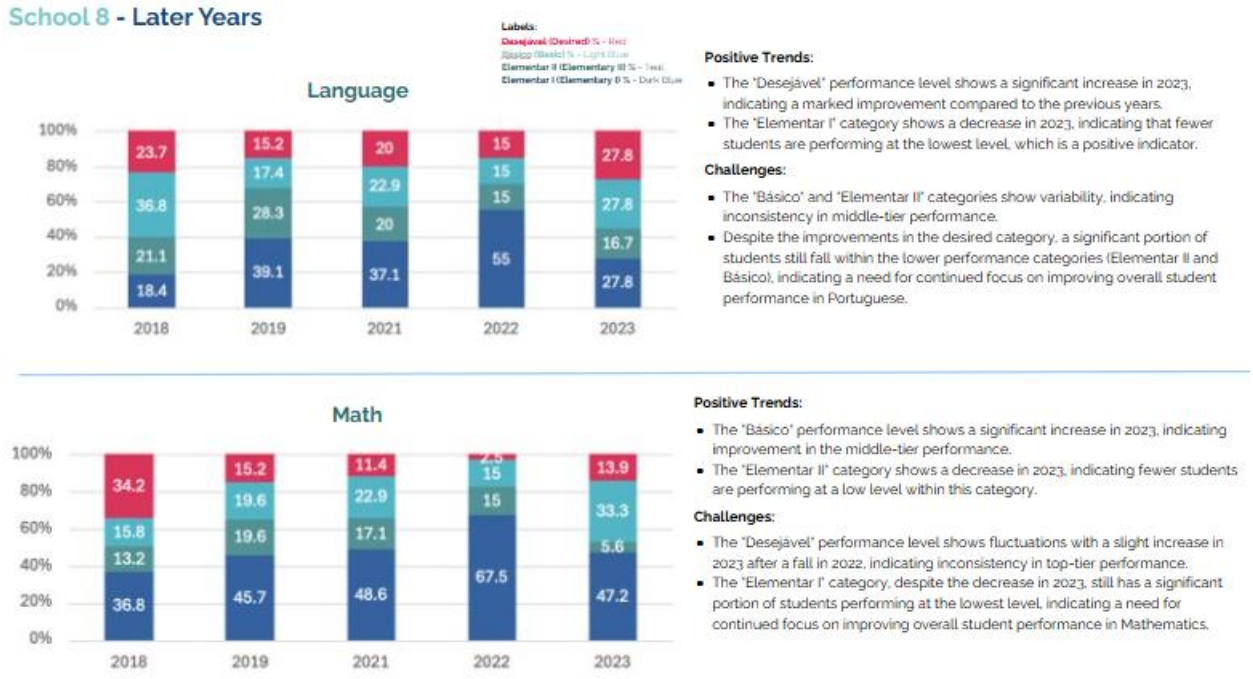
Source: Authors

Figure B.6. - Analysis of learning outcomes for school 8 – Early Years (1st to 5th grade)



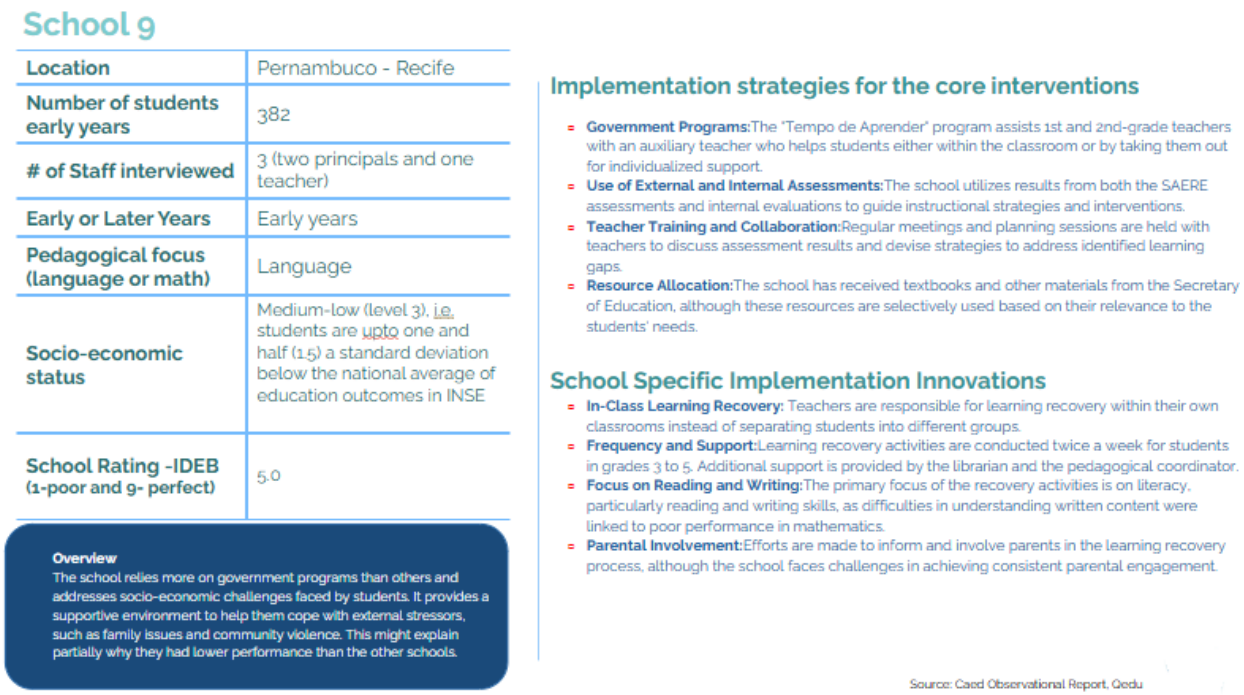
Source: Authors

Figure B.7. - Analysis of learning outcomes for school 8 – Later Years (6th to 9th grade)



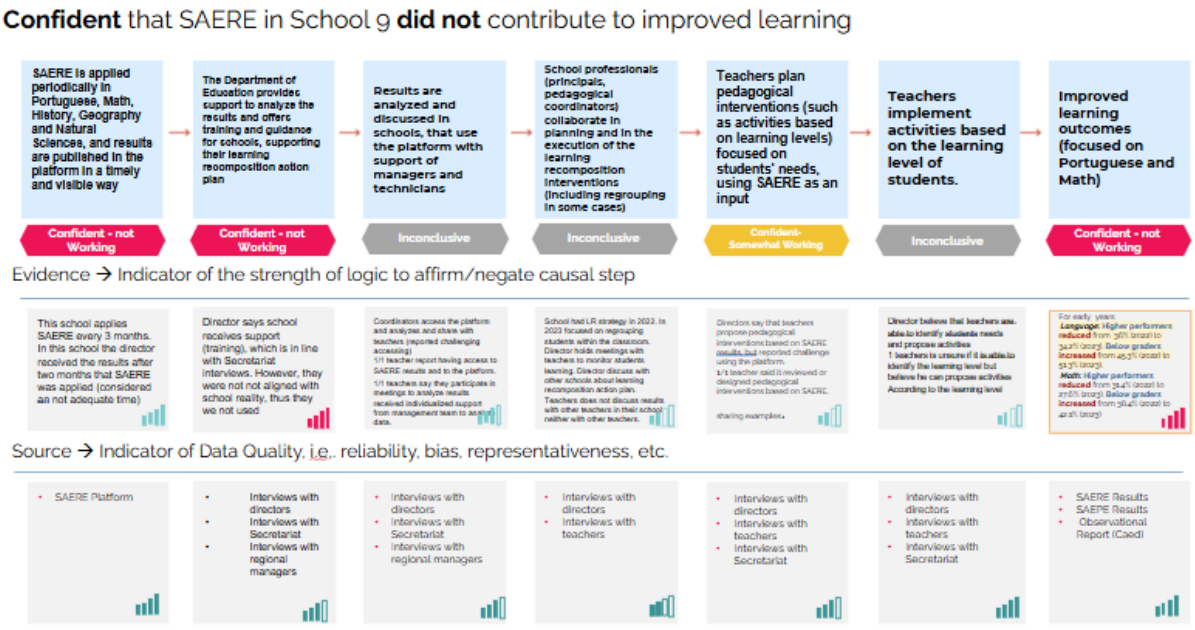
Source: Authors

Figure B.8 - Analysis of school characteristics and interventions – School 9



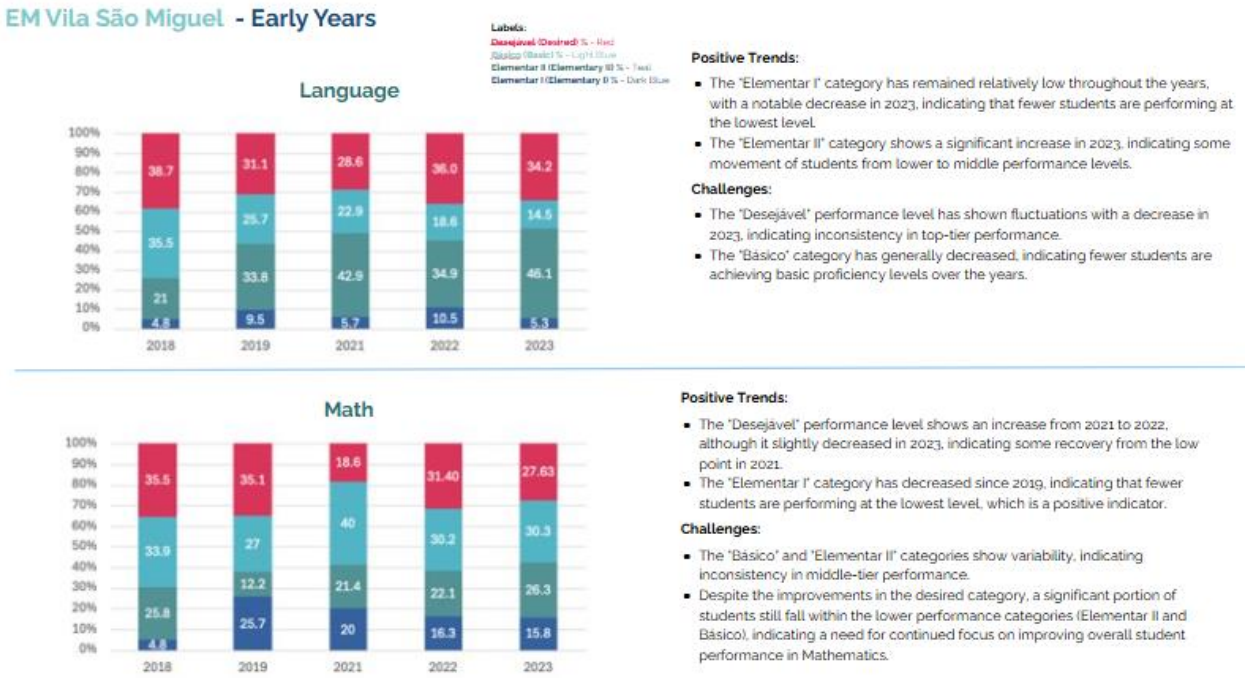
Source: Caed Observational Report, Oedu

Figure B.9 - Analysis of process tracing for School 9



Source: Authors

Figure B.10. - Analysis of learning outcomes for school 9 – Early Years (1st to 5th grade)



Source: Authors