

ISSN: 1130-3743 - e-ISSN: 2386-5660
DOI: <https://doi.org/10.14201/teri.31102>

A THEORY FOR SYSTEMIC IMPROVEMENT OF EDUCATION IN LATIN AMERICA

*Una teoría para la mejora sistémica de la educación en
América Latina*

Axel RIVAS
Universidad de San Andrés. Argentina.
arivas@udesa.edu.ar
<https://orcid.org/0000-0003-4714-3511>

Date of receipt: 12/01/2022
Date accepted: 03/03/2023
Date of online publication: 01/07/2023

How to cite this article / Cómo citar este artículo: Rivas, A. (2023). A Theory For Systemic Improvement of Education in Latin America [Una teoría para la mejora sistémica de la educación en América Latina]. *Teoría de la Educación. Revista Interuniversitaria*, 35(2), 99-120. <https://doi.org/10.14201/teri.31102>

ABSTRACT

The study of the improvement of educational systems has little development in the tradition of research in Latin America. The appearance of standardized assessments, which allow the comparison of student's achievement over time, has generated biased or simplistic interpretations of the causes of learning outcomes. The proposal for a theory of systemic improvement is based on these warnings and seeks to generate a framework that escapes evaluation biases with a pragmatic attitude that can interpret and use them to consolidate a vision of the complexity of educational systems.

The grounded theory that is presented is based on multiple previous research and is organized based on four combined analytical axes: 1-systemic governance,

2-the expansion of capacities of the education system actors, 3-the alignment of accountability and 4-coherence in curricular translation. These axes require entering the black box of the functioning of educational systems and noticing the multiple layers that mediate between policies and teaching practices. In order to advance in this direction, a series of methodological derivations are proposed that favor empirical testing of the theoretical proposal. Finally, the epistemological and political problem of defining the processes of "systemic improvement" and the need to use and at the same time go beyond standardized evaluations are addressed.

Keywords: systemic improvement; educational systems; comparative education; Latin America; educational policy; standardized assessments; quality of education; educational governance.

RESUMEN

El estudio de la mejora de los sistemas educativos tiene un escaso desarrollo en la tradición de la investigación en América Latina. La aparición de las evaluaciones estandarizadas, que permiten la comparación en el tiempo de los logros, ha generado interpretaciones sesgadas o simplistas sobre las causas de los resultados de aprendizajes. La propuesta de una teoría de la mejora sistémica parte de estas advertencias y busca generar un marco que escape a los sesgos de las evaluaciones con una actitud pragmática que pueda interpretarlas y usarlas para consolidar una visión de la complejidad de los sistemas educativos.

La teoría fundamentada que se presenta está basada en múltiples investigaciones previas y se organiza en base a cuatro ejes analíticos combinados: 1-la gobernanza sistémica, 2-la expansión de capacidades de los actores del sistema, 3-la alineación de la rendición de cuentas y 4-la coherencia en la traducción curricular. Estos ejes requieren entrar en la caja negra del funcionamiento de los sistemas educativos y advertir las múltiples capas que median entre las políticas y las prácticas de enseñanza. Para lograr avanzar en esta dirección se plantean una serie de derivaciones metodológicas que favorezcan testeos empíricos de la propuesta teórica. Finalmente, se aborda el problema epistemológico y político de la definición de los procesos de "mejora sistémica" y la necesidad de usar y al mismo tiempo ir más allá de las evaluaciones estandarizadas.

Palabras clave: mejora sistémica; sistemas educativos; educación comparada; América Latina; política educativa; evaluaciones estandarizadas; calidad de la educación; gobernanza educativa.

1. INTRODUCTION

In Latin America, standardized assessments that measure learning appeared in most countries in the early 1990s (Martínez Rizo, 2008). During these long 30 years, international comparative tests were also applied in the region: OECD's PISA for secondary level and UNESCO's ERCE for primary level. Latin America has been populated with measurements that make it possible to compare the evolution of educational results over time. However, there is almost no precedents of research on the improvement of education systems based on these data. Much less can we speak of a field of research on the subject. Precisely when it is possible to measure the evolution of the systems and to point out who is improving or worsening in the headlines of the newspapers, it does not seem to be the object of attention from the field of research.

This is the paradox of standardized tests: they shine a light on systems until they blind them. Tests are often misinterpreted or used politically. They generate discomfort in almost all systems. Even those that improve often do not know it because the news is the "low level", always detectable in a region with such large social debts (Ravela, 2003). Identifying cases of improvement seems to be an accomplice work with these standardized reflectors that rest on the systems. Doubts, fears, excuses and disinterest in "measurable improvement" appear, as if it were a mere power effect of those who use evidence to judge from the outside what happens within the systems.

This vacuum in the field of research is hastily filled by non-specialized interpreters who carry out superficial analyses or by the technocratic imprint of international organizations that seek to establish causal relationships in order to make policy recommendations.

The absence of in-depth and rigorous comparative studies on the improvement of education systems in the region is the starting point for this article. At its base is the search for a more elaborate understanding of how systems function, how they can improve and what meanings the concept of "improvement" has. To that end, I propose a theory for analyzing the improvement of educational systems in Latin America that can escape hasty conclusions and mistrust of measurement. A theoretical vision of systemic improvement is a necessary step in the face of the avalanche of non-specialized discourses that make naïve or incomplete hypotheses about test results. At the same time, it is a necessary link between empirical research and policy recommendations.

To address this purpose I propose to begin by identifying some of the obstacles facing the study of systemic improvement in the region.

First, we must confront the methodological limitations of the evidence. In previous studies we have shown that the PISA tests themselves have had methodological changes that make the results after the 2006 editions incomparable with the previous ones in Latin American countries (Rivas & Scasso, 2021). It is possible to notice thanks

to the transparency of the databases and methodological frameworks of the PISA test. However, national tests of the quality of education in the countries of the region often do not present open information and do not clarify to what extent the intertemporal comparison is valid (Fernández-Cano, 2016).

Secondly, the most well-known limitation of the tests is the one referred to the cutback of what educational system are and the possible incentives to focus teaching on what is measured. Many previous studies have analyzed the "teaching to the test" effects of performance-based policies (Verger et al., 2018, Lewis & Lingard, 2015). Assessment incentives can create biases in the curriculum and limit the view of learning to that which is measured. Excessive pressure for results can also generate a burnout effect on teachers and accelerate their exit from the education system (Gundlach et al., 2010).

Third, we must find a way to look at Latin American education in its own context. Regional analysis of education systems allows for a more equidistant comparison, highlighting cultural, economic and social aspects that Latin American countries share, beyond their obvious differences. The regional perspective in comparative education is especially appropriate to avoid focusing on individual countries or systems or jumping to comparisons of very different contexts.

In order to enter the Latin American context, this article proposes the development of a grounded theory of systemic improvement in education. Based on previous research (Rivas & Scasso, 2020; Rivas, 2021, Rivas, 2022), together with a large team of collaborators, we have developed conceptions that dialogue with the empirical evidence collected on the improvement of education systems at the national and subnational levels in Latin America. For two decades, we have compared hundreds of education systems in the region based on numerous variables and methodologies. The dialogue between this empirical evidence and international theories on systemic improvement feeds the theoretical proposal of this article.

2. AN ELUSIVE OBJECT OF STUDY

The analysis of the improvement of educational systems recognizes a long line of studies in the comparative tradition of education that has sought global success stories in order to copy "recipes" in other contexts. With the emergence of international standardized assessments this literature emerged more strongly by having a parameter for measuring learning (Barber & Mourshed, 2007; OECD, 2010; Tucker, 201; Steward, 2012; Creese et al., 2015; Schleicher, 2018; Crato, 2021).

But these were studies about "the best in the world," not aimed at those who managed to improve over time. Studies on the improvement of educational systems and subsystems recognize a longer trajectory of subnational comparison of districts and states in the United States, a country that combines strong decentralization with measurement of learning (Murphy & Hallinger, 1988; Snipes et al., 2002). District

success stories began to be analyzed: New York's District 2, Long Beach, California's Sanger District, and San Diego, among others.

In a more international perspective, the document that marks a before and after in this literature is the McKinsey consulting firm's report on 20 cases of systemic improvement in different parts of the world (Mourshed et al., 2010). The study found a series of stages of educational systems that are marked by evolutionary phases of universal policy prescriptions. This study was criticized for its soft criteria for comparison, which ignore the complexity of the historical and cultural contexts of the countries studied (Coffield, 2012).

As we analyzed in a systematic review of the literature on improving education systems (Barrenechea, Beech & Rivas, 2022), there are numerous other antecedents of individual case studies of improvement: Ontario in Canada, (Osmond & Campbell, 2018), Shanghai in China (Pang & Miao, 2017), London in England (McAlevy & Elwick, 2016), Poland (Zawistowska, 2014), Wales (Harris & Jones, 2017), and South Africa (Fleisch, 2016).

This literature revolves around different analytical axes, where the cultural change perspectives of educational systems predominate (Fullan, 2016). The influential study by Elmore (2004) characterizes systemic improvement processes as integrated, multidimensional, highly focused on common objectives and with coherence in interventions. Hopkins (2007) highlights that reforms that achieve systemic improvement build capacity at the school level with external support combined with endogenous growth of pedagogical autonomy.

This literature is predominantly concerned with developed countries. There are some classic precedents that analyze improvement in developing countries with scarce resources and installed capacities (Dalin, 1994). Other studies have focused on analyzing the factors associated with educational outcomes, especially in the new trends that focus on systematizing impact evaluations (Masino & Niño-Zarazúa, 2016).

In Latin America, the field of research on the improvement of educational achievement systems is limited. Carnoy's study (2007) analyzed the cases of Brazil, Chile and Cuba, although not based on test-based improvement but on the static result that showed Cuba with the best results in the region in UNESCO tests. Our own pioneering study developed some hypotheses after comparing over 15 years the educational policies and results of seven countries in the region (Rivas, 2015). Another antecedent is the comparative research of states and municipalities in Brazil that achieved improvements in their results (Carnoy et al., 2017). The recent RISE project study analyzed international cases of systemic improvement in education in developing countries and included some from Latin America (Crouch, 2020).

More specific studies have focused on individual case analysis: Peru (Cueto et al., 2016), Ceará (Vieira et al., 2019), Sobral (Cruz & Loureiro, 2020) and Pernambuco (Neto, 2017). Our comparative research on 12 cases of improvement at the

subnational level takes up several of these previous works in the first in-depth analysis of several cases of improvement in different countries in the region (Rivas, 2021).

The construction of theory on the systemic improvement of education has not yet found its way into a field with so few empirical antecedents. International lending agencies have developed some theoretical models based on new public management and the creation of incentive mechanisms based on market competition (Vegas & Petrow, 2008). Other studies complementary to that perspective sought the political economy keys that favor reforms based on decentralization and competitive pressure (Grindle, 2004).

UNESCO tests in the region led to papers on the factors associated with achievement in the tests (LLECE-UNESCO, 2013), although in-depth research on the contexts and policies of the countries that have achieved improvements in the region remains to be done.

The absence of a field of research on the improvement of education systems is contrasted in Latin America by the recent emergence of the field of effective schools, a topic that is close in theoretical and methodological terms. This field of research is abundant and has its anchors in key publications (Murillo, 2007, Bellei et al., 2014), academic journals and even international congresses. With an audience more focused on school managers, this literature has consolidated different theoretical frameworks that converge into a consolidated field of research. In contrast, the improvement of education systems is still an underexplored terrain, which does not seem to attract the attention of academic research despite the existence of multiple national and subnational cases for comparison after three decades of standardized assessments. Nor has it been the subject of theoretical debates or the search for different methodologies to address the questions left open by these assessments.

3. THEORETICAL CONDITIONS

Every theory assumes philosophical, political and epistemological references that should be made explicit. The field of study of the improvement of educational systems is crossed by disputes that can be traced on different models of scientific reasoning, disciplinary fields, and theoretical traditions.

To study the improvement of educational systems, I propose a first theoretical condition referring to an epistemological approach that allows us to escape the dispute between positivism and interpretative hermeneutics as the dominant paradigms of comparative education (Crossley & Watson, 2003). In the positions closest to positivism, which find their classic reference in Edmund King, it is proposed to assimilate the methods of the natural sciences to find generalizable formulas, predictive laws of the behavior of educational systems. The interpretative tradition is rooted in the comparative education of Brian Holmes, who analyzes history and

contexts in depth, seeking to broaden the understanding of systems and not the application of policy prescriptions.

The search for universal laws has an attachment to quantitative methods that standardized assessments have favored. The positivist imprint is connected with the practical interest of research for decision-making and with making "evidence-based" recommendations. In their analysis of the field of the study of school effectiveness, Fuller & Clarke (1994) called those who belong to this tradition the "policy mechanics". The solutions seem to have no boundaries: it's enough to find an educational practice that works to transfer it to a totally different context. In contrast, the "classroom culturalists" respond to interpretive paradigms where policies encounter different traditions and meanings that cannot be extracted from their contexts (Fuller & Clarke, 1994). In this second perspective, pedagogical practices, teaching materials, and socialization processes that students go through are inevitably traversed by the cultural meanings in which they are inserted, which limits the replicability of policies.

The classical epistemological dispute between generalization and specification can be circumvented from a third tradition, which we assume in our theoretical position: the pragmatic attitude. Lars Mjøset (2006) presents this theoretical current for the field of comparative education. The pragmatic attitude continues an empiricist tradition with a strong theoretical development. Theory is defined as a reflexive accumulation of empirical knowledge that performs a spiraling exercise of ascertainment and reformulation. Qualitative methodologies allow more flexibility for this pragmatic attitude that seeks "the systematic study of constellations of individual factors" (Mjøset, 2006, p. 350).

In addition to a pragmatic attitude, a holistic approach is proposed that is capable of analyzing the complexity of educational systems situated in specific cultural, social and political contexts. Simplistic views of educational outcomes are an epistemological obstacle. It is tempting to look for big factors to explain improvement: a change of government, increased educational funding or certain stellar policies. However, large-scale processes of change in teaching practices have a complexity that precludes linear cause-effect analysis. Our postulate is that without entering this "black box" of the processes of change in teaching practices, it is not possible to conceptualize systemic improvement.

Our theoretical perspective takes as reference previous studies on the complexity of educational systems (Burns & Köster, 2016; OECD, 2015; Snyder, 2013). In this perspective, outcomes cannot be explained via singular components of interventions but on the basis of interaction flows between subjects and educational institutions on a large scale. Educational systems are non-linear organic structures that evolve in multiple feedback layers of simultaneous interactions (Sabelli, 2006).

This approach combines knowledge from different disciplines. The field of political science approaches governance systems and power relations between central state bodies and local educational institutions. The field of didactics and curriculum allows us to conceptualize the structure of knowledge and teaching, their modes

and possibilities of change based on external or internal influences of the actors. At the same time, sociology favors reflections on the distribution of learning and the social effects of educational policies and structures. Even a more epistemological, philosophical and anthropological gaze opens doors to think about the meanings of measurable improvement in tests and how to conceptualize ideas about the most valuable learning visions for a society and its individuals.

The analysis of these processes can only be deepened in a localized way. Our theoretical hypothesis is that systemic improvement takes on specific features in regional contexts such as Latin America, with a broad shared tradition of ways of organizing education, capacities installed in the actors of the system and resources available in unequal societies with high levels of poverty. These conditions, beyond the enormous differences between countries, make it possible to situate the thresholds of public policy action in specific contexts.

4. A GROUNDED THEORY OF SYSTEMIC IMPROVEMENT OF EDUCATION

Following the pragmatic tradition of grounded theory (Glaser, 1992), the postulates proposed here are based on multiple previous studies (Barrenechea, Beech & Rivas, 2022). In particular, the most relevant has been the comparative study of 486 subnational educational systems in Argentina, Brazil, Chile, Colombia, Mexico and Peru (Rivas & Scasso, 2020). In that project we compared all educational indicators during the period 2004-19, identifying cases of improvement in each of the countries and ascertaining them from a consultation with experts. The selection of 12 cases of systemic improvement over time allowed us to study them in depth with interviews with key actors. This work was carried out by a team of 19 researchers from the six countries and led to the development of a grounded theory of systemic improvement for the Latin American context, which is presented in this article.

The theory proposes four major combined analytical axes:

1. systemic governance
2. expansion of system actors' capabilities
3. alignment of accountability
4. coherence in curriculum translation

Systemic governance refers to state capabilities to manage an improvement process. An extensive literature on state capabilities allows us to situate the first theoretical axis in the field of political science. Mann's classic study points out the importance of the infrastructural power of the state, as a fabric of legitimacy that allows it to penetrate the beliefs and practices of society within a territory (Mann, 1984). Unlike despotic power, infrastructural power involves relations of cooperation and trust between citizens and their government.

Studies on state capabilities highlight the need to prioritize among multiple demands, make efficient use of resources, coordinate conflicting objectives into a coherent whole and achieve institutional stability of policies, among other issues (Weaver & Rockman, 1993). In the educational field, some research has proposed a systematization of the dimensions that define the quality of governance of educational systems (López Rupérez et al., 2020) and state capacities to coordinate educational policies (IIEP-UNESCO, 2012).

The specific analysis of state capacities in Latin America has been a widely studied topic (Repetto, 2022; Grassi & Memoli, 2016; Soifer, 2012, Kurtz, 2013). Recent work by the Inter-American Development Bank focuses on the concept of trust as the backbone of the quality and sustainability of public policies in the region (Keefer & Scartascini, 2022). This coincides with the approach of Ehren & Baxter (2021) and the studies of Cerna (2014) on the importance of trust in the management of educational policies.

Our previously cited research has allowed us to elaborate a conceptualization of state capacities in the governance of education systems that managed to improve in Latin America. Based on the empirical evidence collected, we highlight the following characteristics of state capacities that influence the improvement of subnational education systems in the region:

- (1) to make education a political priority and shield it from short-term partisan battles
- (2) understanding the culture of the education system, listening to its voices and respecting teachers
- (3) having serious, reflective and flexible leadership to generate agreements and legitimacy in actions
- (4) to develop professional teams in policy management
- (5) generate articulations with different actors within and outside the educational system.

These characteristics are combined with an incremental approach to public policy, which seems especially relevant in the political and social instability of Latin America. Unlike studies that concentrate their focus on policy prescriptions to be implemented beyond contexts, such as the well-known McKinsey Report (Mourshed et al., 2010), what we found in our previous research is that improvement is constructed in a dynamic and flexible manner. As Oszlak (1980) pointed out, the attempt at excessive planning does not provide the necessary doses of realism of the unstable territories of actors in Latin America. Unlike *ex ante* approaches, the incremental tradition points to the importance of intermediate steps or "snowball" effects of policies that make their way through gaining legitimacy and continuity.

This is found in problem-solving iterations approaches (Andrews et al., 2017): it's more important to achieve step-wise progress and build confidence in the ability to solve issues relevant to system actors than to follow universal policy prescriptions. Thanks to these feedback loops, more trust is generated in the different spheres of government and a flexible path is built with a focus on improvement through different doors to enter the educational system. This approach builds what we have called in previous studies, an "education governance platform," without which it is not possible to effectively implement the complex policies that alter learning outcomes on a large scale (Rivas & Scasso, 2020).

Systemic governance generates combined effects in the symbolic and material dimensions of the space and time of educational systems. In the symbolic field, it entrenches a series of priorities, motivations and shared visions in the actors of the system. In terms of temporal sequence, it generates guarantees of policy continuity and sustainability.

The second axis, referring to the capacities of the system's actors, defines the conditions for the dynamics of improvement processes. Educational practices have the characteristic of being highly decentralized and interpersonal: they operate at the micro level of classrooms mediated by the personal intervention of thousands of individual teachers. Improvement cannot be forced or digitized from outside: participation, and involvement of individual and collective actors are necessary.

Burns & Koster (2016) develop an approach to the capabilities of an educational system at four levels: individual, institutional (at the school level), systemic and societal. The individual level is defined from the possibilities that the subjects of the system have to develop new ideas in the local context from existing knowledge. This expands from the institutional level where school management is key to generate learning environments and constant reflection.

This axis may involve the generation, expansion and liberation of capacities within the educational system, three different processes that can be combined with each other. What is relevant in this axis is the construction of a growing dynamic of learning, reflection and power of agency on the part of educators. Without this dynamic approach, improvement processes can happen in one first stage (perhaps more instrumental and technocratic) but not be sustained over time. Moving in this direction implies the construction of a certain ecosystem where learning processes are valued and effectively communicated at the local level so that they are not dispersed and, in this way, avoid dependence on policies that intervene from the top down in hierarchical terms.

In Latin America we find diverse contexts of interaction of educational actors. In some educational systems, such as Colombia, there is a strong tradition of autonomy of schools and teachers, while in other cases, such as Mexico, there are more rigid and structured centralized pedagogical regulation mechanisms (such as textbooks produced by the State and delivered to all students in the country). Beyond certain

shared characteristics in the region, this dimension varies widely both in initial teacher training circuits and in school regulation and control practices.

The third axis refers to accountability processes that express the alignment and control of education systems. Accountability indicates the existence of a set of shared, clearly defined, communicable and achievable objectives that are regularly monitored and produce a certain systemic consistency between the practices, motivations and expectations of educational actors and institutions.

Accountability depends on the creation of a data dashboard that indicates achievements at the level of educational institutions. In all the studied cases of systemic improvement, this systemic and detailed information base appears as a necessary condition for action planning and for the shared allocation of responsibilities. This is what we called in previous studies the "multifocal approach" to improvement: a constant information flow mechanism that informs the achievements of each school on multiple indicators to define an improvement plan, external assistance and/or incentives to push the various actors towards shared goals.

Previous studies have particularly focused on accountability schemes as the core of systemic improvement. Pritchett (2015) developed a theoretical model on systemic coherence. Coherence is built when there is a balance between autonomy, capabilities, and accountability. This process relates to clear information flows that define purposes. These purposes should be coherent at the scale of the whole system, demarcating a shared direction (Pritchett, 2015).

Our theoretical hypothesis here is that accountability processes in Latin America are inseparable from systemic governance: they need solid, reliable and enduring governance institutions to settle on actors in a way that is neither excessively rigid nor stifling. The cases of systemic improvement we have previously studied show the importance of accountability systems: in Ceará, Pernambuco (Brazilian states) and Puebla (Mexican state) we found aligned systems of objectives, incentives and constant measurement. However, these systems were accompanied by large doses of trust in government institutions and processes of constant support to schools with more adverse conditions. These are not systems of exclusionary competition but combine "positive pressure" for results with preventive support mechanisms based on collaboration and coordination between schools.

The fourth axis opens the black box of improvement processes and refers to the structure of school knowledge translation. In the words of Basil Bernstein (2000) we refer to the processes of pedagogical recontextualization that condense the systemic organization and management of the curriculum.

Curriculum governance refers to studies on the pedagogical flow that travels from policies to classrooms. The pioneering work of Cohen & Spillane (1992) pointed out the importance of combining consistency, specificity and clarity with levels of curriculum prescription, hierarchization and sequencing. Perhaps the most emblematic study of this current theory is the book "Why Schools Matter", which studied the flow of curriculum policy in the 54 countries participating in

the TIMMS test. There it was highlighted that the countries with the best results have a curriculum structure with more focus, clarity, articulation and rigor (Schimdt et al., 2001).

The hypothesis behind this fourth axis of the theory is that systemic improvement cannot occur without understanding the internal dynamics of school knowledge management. In some cases of small subnational educational systems, improvement processes can occur relatively independently of the curricular governance structure. But in these cases, improvement will always be limited by a non-controllable external factor that can be enhanced by the agencies that have the greatest capacity for curricular regulation (in many countries these attributions are in national hands, while local governments have limited curricular powers).

The capacity for curricular and pedagogical adaptation requires a combination of technical knowledge, dialogue with multiple stakeholders and continuity over time. These three factors are scarce in Latin America. Perhaps this is one of the reasons why it is so difficult to find systemic improvement processes in the region. Previous studies in the region have highlighted the low level of clarity, focus and rigor in curricular instruments (Valverde, 2009). A comparative research on the educational systems of Cuba, Chile and Brazil highlighted how Cuba achieved better educational results based on a strong integration between the curriculum, teacher training and teaching materials (Carnoy, 2007).

The cases we have found in our own research on subnational education systems in Latin America confirm these hypotheses. The cases of Ceará and Pernambuco in Brazil, Córdoba in Argentina and Puebla in Mexico had an outstanding capacity for curriculum management, with strong technical competencies and coordinated devices that produce materials and guides in dialogue with teachers and schools (Rivas, 2021). These are examples of the use of curricular and pedagogical policy channels to work together to achieve better results.

Again, this axis is strengthened by the development of the previous analytical axes. For example, curricular policies are also combined with the legitimacy of government institutions. Previous comparative studies indicate that countries with better results in standardized assessments have a reliable curricular governance system that offers teachers content that they can interpret, understand and apply, with constant feedback and calibration processes (Schimdt & Prawat, 2006).

5. METHODOLOGICAL STRATEGIES

The theoretical framework for the analysis of the improvement of educational systems in Latin America has a direct empirical derivation. This section presents some methodological strategies to apply the theoretical framework in future studies.

5.1. *Multiplying sources of data on improvement processes and analyzing the intertemporal validity of each source*

The first methodological strategy takes up one of the central obstacles of studies on the improvement of educational systems: the fragile methodological validity of the intertemporal comparison of standardized assessments and other educational indicators.

There are different methodological strategies to face this obstacle. The first involves methodological analysis of the intertemporal validity of standardized assessments. An analysis of the methodological documents of the tests is a first step, but often in Latin America the tests do not explicitly define these technical aspects and it is necessary to conduct interviews with the official specialized agents to better understand the possible biases and methodological problems of the tests. Unfortunately, standardized tests are instruments with a high margin of error and can end up generating studies on cases of improvement that in reality only crystallize some methodological flaw in the application of the tests.

A second complementary strategy is to multiply the sources of data to measure processes of improvement. We recommend comparing different standardized assessments: national and international tests (most Latin American countries have their own evaluations and participate in international tests). Another way of multiplying sources is to appeal to a broader vision of the data ecosystem that countries have: data on access to education, students' trajectories, educational equity, among others. In this way, it is possible to select cases of improvement in various indicators that are consistent over time, avoiding large jumps from one year to the next in standardized assessments that may be the result of statistical problems rather than real improvements.

Other more sophisticated quantitative strategies is to compare different variables associated with the tests and in particular the value-added effect between different test editions, controlling for other factors such as the social context of the students (Carnoy et al. 2017).

Finally, another way to escape the pitfalls of standardized assessments is to broaden the perspective on what systemic improvement means. For example, interviews can be conducted with experts in the education systems analyzed to find possible improvement processes that are not captured in the evaluations, or system actors can be consulted on educational aspects that are not evaluated in a standardized manner.

5.2. *Comparison of cases using multiple quantitative and qualitative variables*

In order to generate hypotheses on the processes of systemic improvement, it's key to use the comparative method of description, interpretation, juxtaposition and simultaneous comparison (Beraday, 1964). An individual case that achieves improvements in its own test has notable limitations that can be overcome by the comparative method analyzing multiple cases that participate in international or subnational comparative tests. Many of the singular explanations can be falsified by comparing similar cases.

Quantitative strategies can be useful if comparable data exist to measure associated factors over time in a set of cases. This allows multiple regression analysis to isolate variables.

The study of macro-structural factors that affect educational outcomes is a necessary complementary strategy to avoid the possible biases of the available databases. For example, it is key to study differences in standardized test scores in different learning areas and years assessed to confront different types of educational policies implemented.

The falsification principle requires comparing multiple variables with different types of available data. If the improvement process was concentrated in reading but not in mathematics, are there possible hypotheses of differential policy interventions? If the improvement was greater in primary than in secondary education, was it because there were differences in policies, subsystem governance, or factors exogenous to education that affected one age group more than the other? If the improvement was over a period of time and then stopped, were there specific policies that were discontinued? If the improvement was in one edition of the test, what is the time frame of the effects of the policies, considering that students begin their schooling long before they are assessed and carry those learnings longitudinally?

To investigate these major questions, it is important to map educational policies in their historical context. Document analysis and interviews with key actors are the best known methodologies for this task of reconstructing policy trajectories (Gale, 2001). It's key to construct a systematic map of interventions over a period of time that should go back to the moment when the students participating in the first edition of the tests being compared began their educational journey up to the final year of those same tests. The sequencing of the socio-economic context and of each of the relevant policy strategies will make it possible to construct different scenarios and contrast hypotheses that often appear a priori and that do not survive the systematic comparison of variables in time and space.

5.3. *Analyzing the reception of educational policies and the culture of the system*

Policy mapping and macro-structural variables are vital but limited tools for understanding the complexity of education systems. Even questionnaires complementary to tests, which provide relevant information on the actors and institutions of the system, have obvious methodological limitations. That is why it's so important to implement a third methodological strategy to grasp the complexity of an educational system. Conducting interviews with teachers, students and actors of the educational community is a key step, but it requires a lot of resources and time in order to achieve representative samples. It should be noted here that even if they do not correspond to a representative sample, these interviews can be important to take the temperature of the system and to learn from the testimonies some visions that allow for the sophistication or nuance of macro-political arguments.

Conducting interviews or focus groups with school principals can be a valuable strategy to capture the representations of the actors in the system on a more limited scale. A further step, always desirable but very costly to practice, is the observation or filming of classes or the ethnographic work of participant observation of the daily life of educational institutions. This approach makes it possible to analyze the practices in their natural environment, unmediated by the representations of the testimonies of the actors interviewed.

The analysis of the interior of the educational system favors the reconstruction of the paths, translations and reinterpretations of educational policies. This makes it possible to review often naïve or incomplete hypotheses about the effects of policies on educational outcomes. In turn, the vision of the actors in the system (including students) is a recommendable way to analyze what improvement means beyond measurements. Many interviews and observations can escape the prison of quantitative data to interpret how an educational system generates multiple effects and learning that cannot be captured by standardized systems.

5.4. *Curriculum flow analysis*

A methodological strategy complementary to the previous ones is the study of the flow of curriculum policy. This strategy makes it possible to open the black box of translations of the structure of school knowledge. Some international studies have analyzed the impact of the prescribed curriculum, textbooks and teacher planning on the results of standardized assessments (Schmidt et al., 2001; Westbury & Hsu, 1996). These lines of research have not generally been applied in Latin America and much remains to be known about the processes of curriculum translation, its evolution over time and the comparison of contexts and countries.

The central purpose of this strategy is to compare models of curriculum organization, prescription, coherence and extension in order to understand how they affect learning processes.

6. DISCUSSION

The four analytical axes of systemic improvement are pieces that fit together in a jigsaw puzzle. They cannot be viewed in isolation. It is in combination that they generate what we might call the "combustion" necessary for continuous sustainable improvement. This does not imply that systems improve only when these four axes are present, since some may be more developed than others and may be tuned in different ways in each context to the existing traditions of that educational system.

The proposed theoretical framework is intended to provide a framework for the development of empirical studies. It comes from and returns to research in comparative systems. Grounded theory is built on the basis of previous research that allows the development of categories based on concrete evidence and a framework that can pass through the verification of multiple cases of improvement.

It is a theoretical model that relies especially on the previous studies compiled by Ehren & Baxter (2021). These scholars developed an analytical framework of systemic reforms based on three backbones: trust, accountability and capabilities. These three dimensions are complemented in our theoretical framework by a fourth one, referring to the process of curriculum translation, which gives more emphasis to the ways in which macro policies are linked to local practices with a focus on the system's curriculum coherence and alignment.

In turn, our theory seeks a specific design for the Latin American context. Based on the study of improvement cases in the region and the adaptation of the categories to the particularities of the context, this article favors a comparative view of educational systems based on the juxtaposition of similarity and proximity factors. Can this same theoretical framework be useful for analyzing systemic improvement in other contexts? Of course, but starting from the conception of grounded theory, based on multiple previous empirical studies, it is more appropriate to present the theoretical proposal adjusted to the context from which it starts.

The dialogue of this theoretical model with some relevant antecedents in the field of educational policy is inevitable. There is a central anchor in the study of systemic improvement which is the well-known McKinsey Report (Mourshed et al., 2010). Our previous studies find many similarities with some of the central findings of this report. We see the flow that occurs between centralized regulation for low development contexts and the extension of local capabilities to achieve consolidate improvement processes. However, our approach differs in a central axis, which is the reading of contexts and the analysis of governance modalities.

Unlike the universal recipes proposed by the McKinsey Report, our theory argues that improvement is possible under conditions of flexibility and with incremental approaches. It's not so decisive what policies are implemented but how the system is governed: it's more important to achieve concrete progress, generate trust, build legitimacy and open doors than to apply a specific textbook

recipe. Latin American education systems are fragile and uncertain: what generates improvement processes is dynamic, contextualized and strongly supported by the resolution of concrete problems experienced by the system's actors. This more territorial, situated and incremental logic is combined with the other axes analyzed: capacity building; the clear definition of objectives and the coherence of accountability; and the systemic alignment of curricular frameworks that feed pedagogical practices.

Finally, it is worth noting the blind spot of our theory. One of the axes that explain the low theoretical and empirical density of the academic debate on the improvement of educational systems in Latin America is the meaning acquired by standardized measurements of quality. Reduced to biased cut-offs on the purposes of education, tests are an iron cage for education systems. When they are strong markers of the way forward -through high-stake assessments- they have the risk of reducing non-assessed areas and teaching to the test.

Can our theory be capable of escaping the magnetism of the measurable? There are two ways to answer this question. The first is empirical and requires studies that analyze whether standardized tests correlate well with non-measurable learning. We know little in this area, but it is possible to advance from research in measuring in broader ways and in analyzing that which is not measured with other methodological strategies.

The second answer is anchored in the political positioning of our theoretical proposal. It is not necessary to fall into a "black-white" polarization. Standardized tests may well have or generate biases, but they also have a value in themselves. It's enough to look at the most basic: whether students can read and understand a text is a central component of their right to education and this can indeed be measured. This pragmatic attitude, as we have defined it, does not fall into the technocratic view of assessments as mirrors of education systems. It is precisely the theoretical layer that allows us to conceptualize the complexity of what is measured and, above all, how the practices of an entire system are modified so that its results change.

Moving away from simplistic analytical biases that propose linear hypotheses to explain the results is a necessary step to review what systemic improvement is and what it means. We hope to contribute to this search with a theory that is both applicable and open to other visions of what are the goals of educational systems in diverse contexts and in a changing and uncertain world.

REFERENCES

- Andrews, M., Lant P., & Woolcock, M. (2017). *Building state capability: Evidence, analysis, action*. Oxford University Press.
- Barber, M., & Mourshed, M. (2007). *How the world's best-performing school systems come out on top*. McKinsey & Company.
- Barrenechea, I., Beech, J., & Rivas, A. (2022). How can education systems improve? A systematic literature review. *Journal of Educational Change*. <https://doi.org/10.1007/s10833-022-09453-7>
- Bellei, C., Contreras, C., Vanni, X., y Valenzuela, J. (2014). *Lo aprendí en la escuela: ¿Cómo se logran procesos de mejoramiento escolar?* Lom Ediciones.
- Bernstein, B. (2000). *Pedagogy, symbolic control and identity: Theory, research, critique*. Rowman & Littlefield.
- Bereday, G. (1964). *Comparative Method in Education*. Holt, Rinehart & Winston.
- Burns, T., & Köster, F. (2016). *Governing Education in a Complex World*. OECD Publishing.
- Carnoy, M. (2007). *Cuba's Academic Advantage*. Stanford University Press.
- Carnoy, M., Marotta, L., Lozano, P., Khavenson, T., Recch, F., Guimarães, F., & Carnauba, F. (2017). Intranational Comparative Education: What State Differences in Student Achievement Can Teach Us about Improving Education—the Case of Brazil. *Comparative Education Review*, 61(4), 726-759. <https://doi.org/10.1086/693981>
- Cerna, L. (2014). Trust: what it is and why it matters for governance and education. *OECD Education Working Papers*, no. 108 (November). <http://dx.doi.org/10.1787/5jxswcg0t6wl-en>
- Coffield, F. (2012). Why the McKinsey reports will not improve school systems. *Journal of Education Policy*, 27(1), 131-149. <https://doi.org/10.1080/02680939.2011.623243>
- Cohen, D., & Spillane, J. (1992). Chapter 1: Policy and practice: The relations between governance and instruction. *Review of Research in Education*, 18(1), 3-49. <https://doi.org/10.3102/0091732X018001003>
- Crato, N. (2021). *In Improving a Country's Education*. Springer.
- Creese, B., Gonzales, A., & Isaacs, T. (2015). *Aligned instructional systems: Cross-jurisdiction benchmarking report*. Center for International Education Benchmarking.
- Crossley, M., & Watson, K. (2003). *Comparative and international research in education. Globalisation, context and difference*. Routledge Falmer.
- Crouch, L. (2020). Systems Implications for Core Instructional Support Lessons from Sobral (Brazil), Puebla (Mexico), and Kenya. RISE Insight Series. 2020/020. https://doi.org/10.35489/BSG-RISE-RI_2020/020
- Cruz, L., & Loureiro, A. (2020). *Alcançando um Nível de Educação de Excelência em Condições Socioeconômicas Adversas: O Caso de Sobral*. World Bank Group.
- Cueto, S., León, J., & Muñoz, I. G. (2016). Conductas, estrategias y rendimiento en lectura en PISA: análisis para el Perú. *REICE: Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 14(3), 5-31. <https://doi.org/10.15366/reice2016.14.3.001>
- Dalin, P. (1994). *How Schools Improve. An International Report*. Cassell.

- Ehren, M., & Baxter, J. (2021). *Trust, Accountability and Capacity in Education System Reform. Global Perspectives in Comparative Education*. Routledge.
- Elmore, R. (2004). *School Reform from the Inside Out: Policy, Practice, and Performance*. Harvard Education Press.
- Fernandez Cano, A. (2016). A Methodological Critique of the PISA Evaluations. *Relieve*, 22(1). <http://dx.doi.org/10.7203/relieve.22.1.8806>
- Fleisch, B. (2016). System-wide improvement at the instructional core: Changing reading teaching in South Africa. *Journal of Educational Change*, 17(4), 437-451. <https://doi.org/10.1007/s10833-016-9282-8>
- Fullan, M. (2016). The elusive nature of whole system improvement in education. *Journal of Educational Change*, 17(4), 539- 544. <https://doi.org/10.1007/s10833-016-9289-1>
- Fuller, B., & y Clarke, P. (1994). Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules and pedagogy. *Review of Educational Research*, 64(1), 119-157. <https://doi.org/10.3102/00346543064001119>
- Gale, T. (2001). Critical policy sociology: Historiography, archaeology and genealogy as methods of policy analysis. *Journal of Education Policy*, 16(5), 379-393. <https://doi.org/10.1080/02680930110071002>
- Glaser, B. G. (1992). *Basics of grounded theory analysis*. Sociology Press.
- Grassi, D., & Memoli, V. (2016). Political determinants of state capacity in Latin America. *World Development*, 88, 94-106. H
- Grindle, M. (2004). *Despite the Odds. The Contentious Politics of Education Reform*. Princeton University Press.
- Gundlach, G. T., & Cannon, J. P. (2010). 'Trust but verify?' The performance implications of verification strategies in trusting relationships. *Journal of the Academy of Marketing Science*, 38, 399-417. <https://doi.org/10.1007/s11747-009-0180-y>
- Harris, A., & Jones, M. (2017). Professional learning communities: A strategy for school and system improvement? *Wales Journal of Education*, 19(1), 16-38. <https://doi.org/10.16922/wje.19.1.2>
- Hopkins, D. (2007). *Every school a great school: Realizing the potential of system leadership*. McGraw-Hill Education.
- IIEP-UNESCO. (2012). *Strengthening of Education System*. IIEP-UNESCO.
- Keefer, P., y Scartascini, C. (2022). *La clave de la cohesión social y el crecimiento en América Latina y el Caribe*. BID.
- Kurtz, M. (2013). *Latin American State Building in Comparative Perspective: Social Foundations of Institutional Order*. Cambridge University Press.
- Lewis, S., & Lingard, B. (2015). The Multiple Effects of International Large-Scale. Assessment on Education Policy and Research. *Discourse: Studies in the Cultural Politics of Education*, 36(5), 621-637. <https://doi.org/10.1080/01596306.2015.1039765>
- LLECE-UNESCO. (2013). *Factores asociados al aprendizaje en el SERCE: Análisis de los factores latentes y su vínculo con los resultados académicos de los niños*. OREALC-UNESCO Santiago.

- López Rupérez, F., Expósito-Casas, E., y García, I. (2020). Un marco analítico para la evaluación de la calidad de la gobernanza de los sistemas educativos. *Revista Iberoamericana de Educación*, 83(1), 53-76. <https://doi.org/10.35362/rie8313672>
- Mann, M. (1984). The autonomous power of the state: its origins, mechanisms and results. *Archives Européennes de Sociologie*, 25(2), 185-213. <https://www.jstor.org/stable/23999270>
- Martínez Rizo, F. (2008). La evaluación de aprendizajes en América Latina, *Cuadernos de Investigación*, 32, INEE.
- Masino, S., & Niño-Zarazúa, M. (2016). What works to improve the quality of student learning in developing countries? *International Journal of Educational Development*, 48, 53-65. <https://doi.org/10.1016/j.ijedudev.2015.11.012>
- McAleavy, T., & Elwick, A. (2016). *School Improvement in London: A Global Perspective*. Education Development Trust.
- Mjøset, L. (2006). No fear of comparisons or context: on the foundations of historical sociology. *Comparative Education*, 42(3), 337-362. <https://doi.org/10.1080/03050060600875584>
- Mourshed, M., Barber, M., & Chijioke, C. (2010). *How the World's Most Improved School Systems Keep Getting Better*. McKinsey & Company.
- Murillo, F. (Coord.). (2007), *Investigación Iberoamericana sobre eficacia escolar*. Convenio Andrés Bello.
- Murphy, J., & Hallinger, P. (1988). Characteristics of instructionally effective school districts. *The Journal of educational research*, 81(3), 175-181. <https://doi.org/10.1080/00220671.1988.10885819>
- Neto, R. M., Medeiros, H. A., Mello, F. M., de Oliveira, F. M., Lira, M. H., & Paiva, F. (2017). Ensino médio na rede estadual de Pernambuco: educação integral e pacto de gestão por resultados. *Cadernos Cenpec*, [S.l.], 6(2), junio. <http://dx.doi.org/10.18676/cadernoscenpec.v6i2.363>
- OECD. (2010). *Strong Performers and Successful Reformers in Education. Lessons from PISA for the United States*. OECD Publishing.
- OECD. (2015). *Governing Complex Education Systems – Framework for Case Studies*. OECD Publishing.
- Osmond-Johnson, P., & Campbell, C. (2018). Transforming an education system through professional learning: developing educational change at scale in Ontario. *Educational Research for Policy and Practice*, 17(3), 241-256.
- Oszlak, O. (1980). Políticas públicas y regímenes políticos: reflexiones a partir de algunas experiencias latinoamericanas. *Estudios CEDES*, 3(2). <http://repositorio.cedes.org/handle/123456789/3470>
- Pang, N. S., & Miao, Z. (2017). The Roles of Teacher Leadership in Shanghai Education Success. *Bulgarian Comparative Education Society*, 15. <https://eric.ed.gov/?id=ED574221>
- Pritchett, L. (2015). Creating education systems coherent for learning outcomes: Making the transition from schooling to learning. RISE-WP-15/005. <https://riseprogramme.org/publications/creating-education-systems-coherent-learning-outcomes>
- Ravela, P. (2003). *¿Cómo aparecen los resultados de las evaluaciones educativas en la prensa?* PREAL.

- Repetto, F. (2022). *Capacidad estatal: requisito para el mejoramiento de la política social en América Latina*. Banco Interamericano de Desarrollo.
- Rivas, A. (2015). *América Latina después de PISA. Lecciones aprendidas de la educación en siete países (2000-2015)*, CIPPEC, Natura e Instituto Natura.
- Rivas, A. (2021). *La mejora sistémica de la educación. Estudio de casos en América Latina*. Fundación Santillana.
- Rivas, A. (2022). The long road to systemic improvement in education: a comparative multi-level case study in federal countries. *Comparative Education*, 59(1), 77-98. <https://doi.org/10.1080/03050068.2022.2133860>
- Rivas, A., y Scasso, M. (2020). *Las llaves de la educación. Estudio comparado sobre la mejora de los sistemas educativos subnacionales en América Latina*. Fundación Santillana.
- Rivas, A., & Scasso, M. (2021). Low stakes, high risks: the problem of intertemporal validity of PISA in Latin America. *Journal of Education Policy*, 36(2), 279-302. <https://doi.org/10.1080/02680939.2019.1696987>
- Sabelli, N. (2006). Complexity, technology, science, and education. *The Journal of the Learning Sciences*, 15(1), 5-9. https://doi.org/10.1207/s15327809jls1501_3
- Schleicher, A. (2018). *How to build a 21st-century school system*. OECD Publishing.
- Schmidt, W., McKnight, C., Houang, R., Wang, H., Wiley, D., Cogan, L., y Wolfe, R. (2001). *Why Schools Matter: A Cross-National Comparison of Curriculum and Learning*. Jossey-Bass.
- Schmidt, W.H., & Prawat, R.S. (2006). Curriculum coherence and national control of education: Issue or non-issue? *Journal of Curriculum Studies*, 38, 641-658. <https://doi.org/10.1080/00220270600682804>
- Snipes, J., Doolittle, F., & Herlihy, C. (2002). *Foundations for Success: Case Studies of How Urban School Systems Improve Student Achievement*. Council of the Great City Schools.
- Snyder, S. (2013). *The Simple, the complicated, and the complex: educational reform through the lens of complexity theory*. OECD Education Working Papers, no.96.
- Soifer, H. (2012). Measuring state capacity in contemporary Latin America. *Revista de Ciencia Política* 32(3). <https://doi.org/10.4067/S0718-090X2012000300004>
- Steward, V. (2012). *A World-Class Education: Learning from International Models of Excellence and Innovation*. ASCD.
- Tucker, M. (2011). *Surpassing Shanghai. An Agenda for American Education Built on the World's Leading Systems*. Harvard Education Press.
- Valverde, G. (2009). Estándares y evaluación. En S. Schwartzmann y C. Cox (Eds.), *Políticas Educativas y Cohesión Social en América Latina* (pp. 57-68). CIEPLAN.
- Vegas, E., & Petrow, J. (2008). *Raising Student Learning in Latin America. The Challenge for the 21st Century*. Banco Mundial.
- Verger, A, Parcerisa, L., & Fontdevila, C. (2018). The Growth and Spread of Large-Scale Assessments and Test-Based Accountabilities: A Political Sociology of Global Education Reforms. *Educational Review*, 71(1), 1-26. <http://dx.doi.org/10.1080/00131911.2019.1522045>
- Vieira, S. L., Plank, D. N., y Vidal, E. M. (2019). Política Educacional no Ceará: processos estratégicos. *Educação e Realidade*, 44(4). <https://seer.ufrgs.br/index.php/educacaorealidade/article/view/87353>

- Weaver, K., y Rockman, B. (1993). *Do institutions matter?: government capabilities in the United States and abroad*. Brookings Institution Press.
- Westbury, I. y Hsu, C. (1996). Structures of curriculum governance and classroom mathematics. *Educational Evaluation Policy Analysis*, 18, 123–139. <https://doi.org/10.2307/1164552>
- Zawistowska, A. (2014). The Black Box of the Educational Reforms in Poland: What Caused the Improvement in the PISA Scores of Polish Students? *Polish Sociological Review*, 187(3), 333-350.