

School Autonomy with Accountability in Latin America (1980-2023):
Politics, Institutions, and Ideas in the Diffusion of a Global Education Reform

Tomás Esper

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2024

© 2024

Tomás Esper

All Rights Reserved

Abstract

School Autonomy with Accountability in Latin America (1980-2023):
Politics, Institutions, and Ideas in the Diffusion of a Global Education Reform

Tomás Esper

As time goes by, different policy ideas dominate the global scene and circulate among countries rendering different policy reforms. Since the 2000s, ideas promoting standardization, outcomes-based governance, and competition have gained momentum as the dominant paradigm for educational improvement, epitomized in the global model of School Autonomy with Accountability (SAWA). SAWA promotes greater autonomy in exchange of tight-accountability measures through learning standards, large-scale assessments, and routinized inspections. Globally, Latin America has one of the longest and most intimate relationships with SAWA reforms, under the direct influence of the World Bank and OECD. From school-based management to large-scale assessments, SAWA policies have widely spread in the region, despite failing to fulfil their promise of educational improvement.

This dissertation examines SAWA in Latin America from a comparative, cross-scalar, and historical perspective, analyzing its regional diffusion and its local imbrications in the cases of Argentina and Colombia. In doing so, it asks the following questions: (1) To what extent have SAWA policies disseminated in Latin America and the Caribbean since the 1980s? (2) Why have Argentina and Colombia followed a different trajectory regarding SAWA policies?

The research design combines two different strands. The first explores the dissemination of the SAWA package across Latin America since the 1980s. It empirically traces the de jure

adoption of 17 policies associated with the SAWA reforms across 23 Latin American countries from 1980 until 2023. The second offers an in-depth, comparative, and historical examination of the causes and mechanisms driving divergent processes of adoption, resistance, and institutionalization of SAWA policies in Argentina and Colombia. This strand combines qualitative methods, using interviews (n=93) and policy documents as data sources, with social network analysis of documents references. This dissertation follows the three-article format, where the different papers unpack the underlying reasons driving SAWA policies' trajectory in Argentina and Colombia.

The regional analysis demonstrates an uneven spread of SAWA policies, ruling out ideas of policy convergence across Latin America. First, it shows that policies associated with centralized control over education, such as standardized testing and national curricula, have been more widely disseminated than those link with school competition or market-like mechanisms in education.

This research demonstrates that policy change in Latin America's education governance occurred through layering, conversion, and drift mechanisms, as seen in Colombia and Argentina. In Colombia, the gradual accumulation of accountability reforms and coalitions' strategic use of contextual shifts have established a governance model deeply rooted in new public management principles. In Argentina, the evolution of decentralization reforms has rendered a weakened centralized authority from the national government, favoring provincial autonomy. Consequently, sub-national developments, such as the re-emergence of standardized testing, became more relevant for education governance than national ones. These findings illustrate how incremental policy adjustments can drive substantial transformations over time.

The study also reveals the role of policy coalitions in shaping the adoption and trajectory of SAWA reforms. In Colombia, a coalition aligned with SAWA's ideals successfully navigated resistance to advance these reforms, while in Argentina, an opposing coalition successfully blocked similar initiatives. In both cases, policies that challenge established interests and power dynamics face increased resistance, often requiring broad cross-sectoral support to be adopted and institutionalized.

Ultimately, this research advances policy transfer as a dynamic cross-scalar process extending beyond initial adoption. The role of Bogotá and Buenos Aires as champions of accountability reforms exemplify this, each playing a pivotal role in shaping national and regional policy trajectories over time.

Table of Contents

List of Tables and Figures.....	iv
Acknowledgments.....	vi
Dedication.....	viii
1. Introduction.....	1
1.1. The rise of a global school reform	5
1.2. SAWA in Latin America	13
2. Policy transfer across time and space: An interpretative framework	15
2.1. Institutions, ideas, and interest groups as engines of policy change.....	18
2.1.1. Unpacking institutional change.....	19
2.1.2. Ideas and paradigms in policy change	21
2.1.3. Interest and agendas through policy coalitions	22
3. Research design	23
3.1. Macro-level strand (RS1): Researching SAWA diffusion.....	23
3.1.1. Data collection process	24
3.1.2. Analyzing SAWA diffusion and selecting cases	28
3.2. Meso-level strand (RS2): A comparative sequential design.....	32
3.2.1. The vertical, horizontal and transversal axes in cross-case and within-case analysis.....	34

3.2.2.	Case selection: Unpacking causal mechanisms	35
3.2.3.	The cases of Argentina and Colombia.	36
3.2.4.	Data collection: Interviews and document analysis.	39
4.	Results.....	50
4.1.	Article 1: Changing paradigms: a historical analysis of school autonomy and accountability policies in Colombia.....	50
4.2.	Article 2: Not everything learned is transferred: Unravelling knowledge use and lesson-drawing behind school autonomy with accountability reforms in Colombia.....	96
4.3.	Article 3: The politics of standardized assessments: A multi-scalar analysis of a global reform in Argentina.....	145
5.	Conclusion and agenda for further research	200
5.1.	Comparing across key mediating factors: institutions, ideas, and coalitions	200
5.2.	Global determinants of education reform: the pressure and influence of International Organizations.....	202
5.2.1.	Argentina’s pendular stance ahead International Organizations	203
5.2.2.	The rising role of the OECD in Colombia	205
5.2.3.	Looking beyond the nation-state: cross-scalar dynamics of global policies circulation.	208
5.3.	The characteristics and effects of governance decentralization in education	209
5.3.1.	The shifting dynamics of educational decentralization in Argentina	210

5.3.2. Decentralization in Colombia: controlling schools at a distance.....	212
5.4. Antagonistic ideas in constant dispute	214
5.5. Coalitions and interest groups: victories, defeats, and unfinished businesses	216
5.6. Intended research contributions	219
5.7. A developing research agenda	222
References.....	224
Appendix A. Country Coding Sheet for the SAWA Reform Database	248

List of Tables and Figures

Introduction

Table 1. School Autonomy with Accountability: its defining logic, instruments and tools

Table 2. Sample of SAWA in Latin America database

Table 3. Adoption of SAWA policies in Latin America (1980-2023)

Table 4. Country-level characteristics of Argentina and Colombia

Table 5. Interviewed key informants per country.

Figure 1. The ‘School Autonomy with Accountability’ mechanism.

Figure 2. Aggregate diffusion of SAWA elements in Latin America (1980-2023)

Figure 3. Costa Rica coding sheet. Questions 22-24.

Article 1

Table 1: Evolution of ‘SABER’ assessment over the years

Table 2: SAWA instruments trajectory and policy changes in Colombia (2002-2018)

Figure 1: The “Educational Revolution” quality improvement cycle

Article 2

Table 1: SAWA instruments trajectory and policy changes in Colombia (2002-2018)

Table 2. Reform waves and the number of analyzed documents and references

Table 3. Most cited authors in Colombia’s policy documents, order by document appearances

Figure 1. A streamlined model of evidence use and policy transfer

Figure 2. Network of authors cited in Colombia’s education policy documents (2002-2022).

Figure 3: *Network of authors cited in Colombia’s policy documents by reform period.*

Article 3

Table 1. The four LSAs – Argentina, Buenos Aires, Córdoba & Mendoza

Table 2: Policy Coalitions in their three-tier belief system

Table 3: Policy Coalitions and Policy Streams

Figure 1. Causal Mechanism for Policy Coalition Influence on LSAs

Figure 2. The adoption, institutionalization and re-emergence of LSAs in Argentina

Conclusion

Table 6. A Comparison of SAWA policies in Argentina and Colombia (1990-2023)

List of Abbreviations

LSAs	Large Scale Learning Assessments
OECD	Organization for Economic Cooperation and Development
PLSAs	Provincial Large-Scale Assessments
PREAL	Program for Educational Reform in Latin America
PISA	Program for International Students Assessments
SAWA	School Autonomy with Accountability
SBM	School-Based Management
MOE	Minister of Education
NLSAs	National Large-Scale Assessments
NPM	New Public Management

Acknowledgments

First and foremost, I would like to express my gratitude to Gita Steiner-Khamsi for the immense privilege of being one of her students and for trusting me from the very beginning. This privilege is none other than counting with her unwavering support, the intellectual challenges and curiosity that characterize her –pushing me to go further–, her respectful advice and mentorship, and the warmth that defines Gita. This dissertation would not have been written without her guidance throughout these years.

Next, I want to thank Oren Pizmony-Levy for his unconditional support and trust and for encouraging me to go the extra mile, learn new things, and step out of my comfort zone. Thank you, Oren, for the countless hours of conversations and your invaluable guidance over these years. I am also grateful to the other members of my committee, each of whom has significantly influenced my academic and professional growth. First, I thank Regina Cortina for the opportunity to serve as editor of *Current Issues in Comparative Education* and for her support and guidance during that year. Regina's work has been a continuous source of inspiration. To Andreas Wimmer, for first welcoming me into his class and offering me groundbreaking and intellectually fascinating approaches to analyzing and understanding the social and political phenomena affecting our world. I am also profoundly grateful for the many exchanges that enriched me both as a researcher and a person. Lastly, I am enormously thankful for Antoni Verger's generosity and warmth since we first met in June of 2019. Toni has been and continues to be a source of inspiration and admiration; I am forever grateful to him for treating me as one of his students throughout these years.

I also wish to acknowledge the other professors and colleagues who have been essential in my journey. To Mauro Moschetti, Daniel Friedrich, Radhika Gorur, Aparna Anand, Garnett

Russell, Jorge Baxter, and Felicitas Acosta, from whom I have learned so much and whom I have constantly chased with my questions, requests for feedback, and new ideas.

My gratitude also extends to the institutions that have supported me in carrying out this research: the U.S. Fulbright Program, the Tim Morris Award from the Education Development Trust, the Institute of Latin American Studies at Columbia University, the Center for the History of Education, and the Department of Education Policy and Social Analysis.

Finally, to my colleagues and friends with whom I have exchanged ideas, shared conversations, and navigated these years together: Iván Matovich, Chris Henderson, Arnela Čolić, Darren Rabinowitz, Mia Chin, Edgar Quilabert, and Sara Pan. Thank you for all your help over these years and for making #TheLatinAmericanWay a reality.

Dedication

Above all, this dissertation is dedicated to my mother, my family and my friends. I could easily match the length of this work if I had to thank my mom for what she has done for me since I was a child. In 2000, when I was just ten years old, we visited Columbia University on one of our trips to New York. I never imagined that I would one day study here; she, perhaps, dared to dream it. For this reason, I believe she is even happier about this dissertation than me. Her happiness and proud justifies the hard work and all the years spent far from home. *Gracias, vieja.*

Next, I want to dedicate this work to my family. To Karina, who raised me since I was five shortly after my dad passed away. To my dad's memory, always an inspiration to do better. To my cousins, who taught me almost every I know about love and friendship: Nacho, Naty, Iva, Vero, Apu, Mati, Sebi, Juan & Feli.

Finally, to all my friends. For believing in me, for supporting me, for putting up with my complaints, for buying me a drink when I was broke, and for coming to visit me. To my friends in New York, with whom we built a family far from home: Diego, Irene, Alejandra, Teresa, Aldo, Marina, Pablo, Daria, Tomás, Josefina, Bijon, Nikola, Alvarito, Saiful, Vincent, Gergana, Momori, Mechi, Meli, Gustavo, Arti— I hope I do not forget anyone. To my friends in Barcelona, with whom I built another family across the ocean: Josefina, Julián, Juan, Mica, Nadia, Manu, and Emi. Thank you all for the support, laughs, and friendship.

And, of course, to my friends in Argentina and around the world who have always been there for me, no matter what I do—this dissertation is for and because of you: Bukito, Dano, Fran, Guido, Cofi, Ari, Bruno, Conde, Rola, Valen, Frank, Nuni, Pili, Melu, Jesi, Bel, Coni, Cami, Marto, Facu, Ale, Guide, Anita, Esteban, Gonza, Norman, Marian, and the list goes on. Being far away was the hardest part! Thank you for making the distance feel a little smaller.

1. Introduction

As time goes by, different policy ideas dominate the global scene. During their hegemonic periods, countries are exposed to them and are pressed by international actors to follow them. In education, manpower planning dominated the 1960s and 1970s, while neoliberalism seemed inescapable during the 1980s and 1990s. Since the 2000s, ideas promoting standardization, outcomes-based governance, and competition have gained momentum as the dominant paradigm for educational improvement, epitomized in the global model of School Autonomy with Accountability (Verger et al., 2019). Known as “SAWA,” the dominant global reform of the last two decades has merged New Public Management and neoliberal principles in its theory of change (Steiner-Khamsi, et al., 2024; Woessmann et al., 2009). On the one hand, SAWA promotes transferring decision-making from central levels to schools on curricula, staffing, and budgeting while holding them accountable through learning standards, large-scale assessments, and routinized inspections (De Grauwe, 2005). On the other hand, it stimulates improvement by mixing monetary or symbolic incentives for reaching pre-defined performance targets with school competition by fostering school choice or enabling private management of state schools (Demas & Arcia, 2015). As a result, SAWA promises to boost students’ performance, progressively eliminate underperforming schools, and create a system more responsive to parental demands (Greany & Higham, 2018). However, not only countries adopting SAWA policies have shown mixed effects on student performance (Hanushek et al., 2013; OECD, 2011), but it has also triggered side effects like test cheating (Berliner & Glass, 2022), the narrowing down of curriculum (Au, 2007), inequalities among parents who can benefit from choice dynamics and those who don’t, or high stress on students and teachers, including higher teachers turnover rates (Jerrim & Sims, 2022). Despite this, since the 2000s, a soaring number of countries have implemented SAWA policies in Europe (Högberg &

Lindgren, 2021; Mentini & Levatino, 2023; Pagès, 2021; Wiborg, 2015, among others), Asia (Patrinos et al., 2015; Teng et al., 2020), or North America, and Latin America (Díaz Ríos, 2020; Maroy & Pons, 2019; Rivas, 2021; Watson & Supovitz, 2001). Hence, the reasons behind SAWA's global spread remain a scholarly puzzle.

In education governance, autonomy, and accountability are the two sides of a coin: the intensity, consequences, and surveillance methods of accountability systems tend to accommodate the degree and type of autonomy granted to schools. Consequently, different SAWA versions' have consolidated globally. In countries characterized by market-based reforms to the state, like England, Sweden, or Australia, autonomy entailed privatizing school management with liberalizing curricula and staffing practices. In compensation, schools are under strict test-based accountability systems with high stakes on teachers' and students' careers or even reaching school closure (Cox, 2012; Greany & Higham, 2018; Wiborg, 2015). However, because such models often increase student segregation and teachers' turnovers (Parcerisa et al., 2022), other countries prefer lower-stakes approaches to the accountability system.

Rendering 'quality assurance' objectives, accountability policies like standardized testing or school data aim to inform educational improvement decisions but avoid direct consequences based on performance to prevent undesired side-effects (Parcerisa et al., 2022). Good examples of this are Colombia, Canada, and Brazil in the Americas (Maroy & Pons, 2019; Mentini, 2023; Rivas, 2021) or European countries such as Norway or the Netherlands (Camphuijsen & Parcerisa, 2023; Ydesen et al., 2023). Thirdly, a subset of countries have granted certain autonomy over school curricula and enacted soft accountability systems where assessments are inconsequential, and inspections portray something similar to a bureaucratic ceremony. These countries tend to maintain a neo-weberian state bureaucracy, like Argentina, Spain, Italy or France, to name a few

(Maroy & Pons, 2019; Mentini & Levatino, 2023; Pagès, 2021; Rivas & Sanchez, 2022). Although these ideal types do not contain all of SAWA's vernacular expressions, they showcase one of the effects of traveling reforms: when it moves, it morphs (Cowen's (Cowen, 2009, p. 315).

As with every complex reform, its global circulation not only renders different versions among adopting venues but also produces the selective borrowing of certain reform ideas, elements, or policies (Gilardi, 2010). Undoubtedly, policy diffusion entails two inseparable sets of questions. One deals with drivers and mechanisms behind the spread of policy models, the role of transnational agents in generating and disseminating them, the effects of temporal and spatial loci, and policies' intrinsic characteristics that explain why certain go viral while others do not (Blatter et al., 2022; Wimmer, 2021). Research centered on this first group has shown the influence of economic coercion, competition, learning from others, and the emulation of global norms in diffusing reforms globally (Hossain, 2022b; Simmons & Elkins, 2004). Similarly, scholars have thoroughly demonstrated the role of International Organizations (IOs) in generating and transferring global scripts (Pizmony-Levy, 2011; Ramirez, 2012) or the association of hegemonic ideas over certain periods coincident with the rise of aligned reforms – characteristic of neoliberalism and pro-market reforms (Henisz et al., 2005; Meseguer, 2004).

The other group of questions relates to why countries adopt such global models, what aspects are withheld and which are not, how this process unfolds, what is the role of key agents, interest groups, and institutions in shaping the transfer of new policies on the ground (Dolowitz & Marsh, 2000; Porto de Oliveira & Osorio Gonnet, 2023). This second strand of policy transfer has frequently concentrated on unpacking global processes within specific contexts, emphasizing the mediatory role of ideas, interests, and institutions in this process (Morais De Sá e Silva, 2017; Steiner-Khamsi, 2021; Verger et al., 2016). This dissertation aims to bridge both traditions by

examining from a comparative, cross-scalar, and historical perspective the diffusion of the SAWA reform package in Latin America and its local imbrications in the cases of Argentina and Colombia. In particular, it asks the following questions:

1. To what extent have SAWA policies disseminated in Latin America and the Caribbean since the 1980s?
 - a. What policies spread more widely, and which had less traction?
2. Why have Argentina and Colombia followed a different trajectory regarding SAWA policies?
 - a. To what extent have global actors or forces influenced SAWA's trajectory in each country?
 - b. How have countries' institutional legacies and local interest groups shaped the adoption, translation, and evolution of SAWA policies in each case?

To answer these questions, this dissertation is organized into two research strands. The first one explores the dissemination of the SAWA package across Latin America since the 1980s, coincident with the peak of the neoliberal order in education (Bromley et al., 2023). It empirically traces the de jure adoption of 17 policies associated with the SAWA reforms across 23 Latin American countries from 1980 until 2023. The second strand offers an in-depth, comparative, and historical examination of the causes and mechanisms driving divergent processes of adoption, resistance, and institutionalization of SAWA policies in Argentina and Colombia. These two countries have followed Washington Consensus policies during the 1990s, both in education and in the public administration more broadly, including state downsizing reforms, fiscal cuts, and a shift towards a regulatory state (Bonal, 2002; Burki & Perry, 1998). However, the implementation and the effects of such reforms, including educational decentralization or the formation of an evaluatory state in education, produced divergent pathways in each country (Benveniste, 2002;

Díaz Ríos, 2024; Falletti, 2010). The legacies of the 1990s reforms and the economic crisis both countries faced during the early 2000s created different political and institutional scenarios. In Argentina, a center-left government dominated the political scene for 16 of the last 20 years, where IOs were portrayed as responsible for past crises, constraining their domestic influence. In contrast, Colombia continued its center-right electoral tradition, which was only interrupted in 2022. During these years, it strengthened its historical ties with the World Bank and the OECD, securing its membership in 2020. To what extent did political-ideological differences among these countries shape the receptiveness towards the SAWA agenda? How did the epochal changes reconfigure the power relations among coalitions struggling to influence education policy? What institutional features were decisive for policy transfer, and which were rendered inconsequential? These, among others, are the questions answered throughout this dissertation.

1.1. The rise of a global school reform

The origins of SAWA can be traced back to the 1980s when the OECD promoted New Public Management (NPM) reforms in state bureaucracies to improve governance efficiency (Morgan, 2007). SAWA locates its privileged unit of change in schools, aiming to break down education systems into smaller compartments with wider decision-making but steered at a distance and under pressure to achieve performance targets (Verger et al., 2019). The OECD refined the ideas and policies to instrument SAWA reforms over the years, from its International School Improvement Project (Verger et al., 2019) to institutionalizing measures on school autonomy and types of accountability systems in PISA questionnaires (OECD, 2011). As with any global reform, school autonomy with accountability has increasingly become deterritorialized and detached from any authorship fingerprints. SAWA is a set of context-resilient principles adaptable to diverse school

systems worldwide and capable of yielding the promised educational improvement. Likewise, over the years, the OECD would not simply recommend SAWA policies at large but provide tailored advice to its country members on how to calibrate these policies to obtain better results (OECD, 2004, 2016; Verger et al., 2019).

In parallel, the World Bank has also contributed significantly towards consolidating SAWA as the dominant model for school reform. Devoted to advancing education decentralization reforms at sub-national and school levels since the 1980s (Mobarak, 2022), during the 1990s, the Bank centered part of its agenda on school-based management (SBM) reforms – undoubtedly SAWA’s predecessor. SBM was grounded in similar principles, namely a critique of public-sector bureaucracy in support of devolved decision-making to increase democratic participation in school governance via parental involvement (Cardenas, 2009; De Grauwe, 2005). Later, in the 2000s, the World Bank would shift towards the triple ‘A’ formula: school autonomy, assessments, and accountability (Demas & Arcia, 2015). These two dimensions, school autonomy, and accountability, were institutionalized within the Bank’s portfolio – and crystalized into its SABER initiative (Demas & Arcia, 2015).

The role of IOs in crafting, advancing, and disseminating education reforms is one of the most documented phenomena of modern times (Bonal, 2002; Hossain, 2022a; Meyer et al., 1997; Mundy et al., 2016). In particular, the World Bank and the OECD, whose central mission spins around economic development, have embraced and disseminated neoliberal policies since the late 1970s (Heyneman, 2003; Ydesen et al., 2019). IOs have used multiple mechanisms to spread reforms, including aid conditionalities, technical assistance, pilot experiences, knowledge generation and brokerage, normative dissemination of best practices, etc. (Edwards et al., 2023; Niemann & Martens, 2018; Zapp, 2017). Scholars have shown that aid conditionalities in World

Bank loans have fostered education decentralization (Hossain, 2022a) or the normative role of IOs in promoting environmental education into a country's curriculum (Pizmony-Levy, 2011). More recently, scholarly attention has shifted towards IOs' role as 'policy brokers', seizing their privileged position as data and knowledge generators to nudge countries' decisions into their preferred policy agendas (Baek, 2022; Niemann & Martens, 2018; Santos & Centeno, 2023; Seitzer et al., 2023). As we shall demonstrate, however, IOs' influencing mechanisms are always mediated, resisted, or leveraged by agents among those countries pursuing policy reforms or under pressure to adopt them.

SAWA is a complex reform package formed by (i) a policy goal, (ii) an underlying theory of change, (iii) a subset of policy instruments, and (vi) a bundle of specific policies to operationalize it. Allegedly, SAWA's goal is to improve education systems' performance, which is measured mainly in student learning outcomes. To do so, SAWA envisions a streamlined and de-bureaucratized governance structure with increased horizontal and vertical accountability from parents and authorities (De Grauwe, 2005; Högberg & Lindgren, 2021). Accountability mechanisms would signal underperforming schools, which would either improve by learning from competitors' best practices or be withdrawn for their continuous failure (Greany & Higham, 2018; Steiner-Khamsi et al., 2024). This idealized functioning of education systems is to be achieved through a complex theory of change composed of four interlocking elements: (i) school autonomy, (ii) standardization & accountability, (iii) performance incentives, and (iv) competition (shown in Figure 1). However, like a food recipe, reform packages require all ingredients to be present and mixed correctly to obtain the promised results.

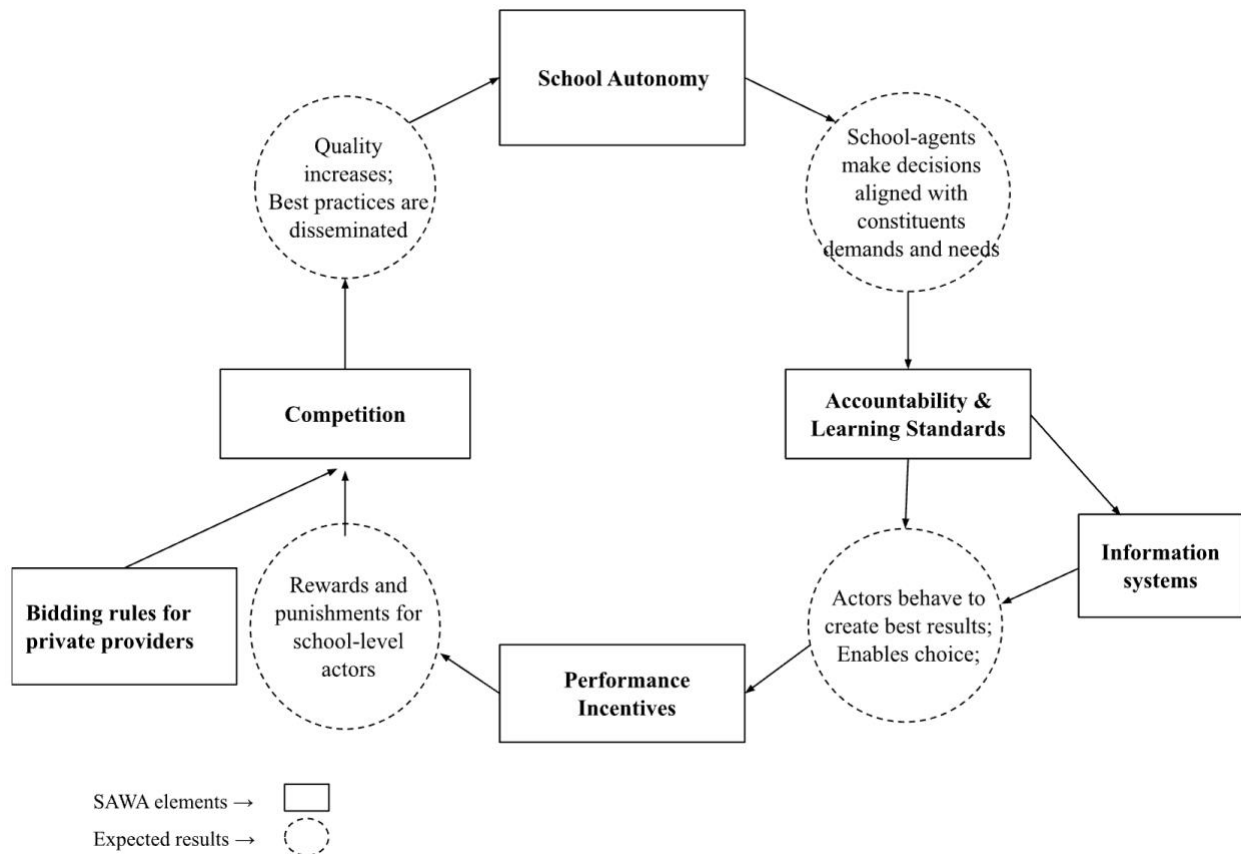
The first of SAWA elements is school autonomy, which has yet become another buzzword in education reform. In the context of SAWA policies, it embodies the transfer of decision-making

from central levels –national or provincial– to school-level agents –principals or teachers, but also school boards–. School autonomy represents the most radical form of educational decentralization, the reason why often these reforms are closely associated (Chambers-Ju, 2006; Lastra-Anadón & Mukherjee, 2019)., Several questions arise when dealing with autonomy: What is devolved? To whom? Under what circumstances? (Frostenson, 2015; Lundström, 2015). Schools might have autonomy over curricula, personnel, budget, and how the school is organized (Demas & Arcia, 2015; De Grauge, 2005). Besides, the authority might fall on principals’ shoulders or school boards, as occurred in Central America under school-based-management reforms (Ganimian, 2016).

Typical to managerial or neoliberal tenets, the devolution of authority is balanced by a double movement composed of standardization and re-centralization through distance monitoring procedures (Ball, 2003; Ozga, 2009). Standardization is frequently operationalized through school curricula, learning standards, or routine inspection procedures for schools. These policies hold school agents accountable, where data gathering and monitoring are deployed with large-scale learning assessments, school monitoring systems, and inspection reports (Högberg & Lindgren, 2021; Milner & Ydesen, 2024; Wiborg, 2015). The emphasis on greater autonomy and accountability uncovers the state restructuring. The traditional role of service provider, input-oriented, and responsible education governance, often labeled the ‘teaching state,’ was replaced by the regulatory or evaluatory state, guided by outcomes, regulations, and sanctions (Maroy & Pons, 2019).

The third element of SAWA reforms is performance incentives. These take the form of economic rewards like school bonuses when schools achieve performance targets. For example, Colombia adopted a school bonus between 2015 and 2017 for those schools reaching the annual

minimum improvement goal (Baxter & León Cadavid, 2021; Esper, 2024). Another example is salary bonuses, which target teachers directly when students reach certain metrics in standardized assessments, for instance. In a different realm, SAWA reforms also create reputational incentives – or stakes – for teachers, principals and schools, as their performance in standardized tests can be seen as a positive or negative sign for their surrounding community (Bunar & Ambrose, 2016).



Source: Elaborated based on Verger (2012).

Figure 1. The 'School Autonomy with Accountability' mechanism.

Arguably, the fourth element of the SAWA model is closer to neoliberal principles than those of managerialism: competition. Competition among schools is seen as the primary driver for improvement (Woessman et al., 2009). Several policies incentivize this logic. For instance, when results in standardized tests are publicly available, schools are often placed in rankings. Similarly, per-capita financing reforms might trigger disputes for more students as the budget is allocated based on the number. Therefore, better performance in official indicators will likely attract more parents to schools. Nonetheless, school choice is the key condition for these policies to produce their desired effect. Only when families can select where to send their children, particularly in the public or state sector, will prior policies increase school competition (Steiner-Khamsi et al., 2024).

Furthermore, certain policies fall closer to the realm of educational privatization or marketization, although they also increase competition in the education system. One is the establishment of supply-side subsidies, like those granted to catholic schools in Argentina (Verger, Moschetti, et al., 2020), or demand-side subsidies to families, exemplified by Chile's voucher system (Zancajo, 2019). The second is the transfer of school management to private providers. This de-regulation de-regulation of school provision is also closely coupled with autonomy, as these schools often are endowed with significant decision-making over most school dimensions. Epitomized in U.S. charter schools, a similar model mushroomed in Colombia during the late 1990s, labeled 'Concession Schools' (Edwards et al., 2017).

The SAWA package includes different policy instruments that operationalize its goals and elements. As discussed in greater detail in section 5.1.1, policy instruments are similar to institutions in that of constraining agents' behavior, rendering power relations among its affected actors, and being subject to change and contestation (Capano & Howlett, 2020). The choice of different instruments reveals an explicit theorization of the relationship between the governing and the governed, or, in other words, the power relations among them (Lascoumes & Le Galès, 2007). Typical instruments include regulations and laws, economic and fiscal ones, information or communication-based, or agreement/incentive-based (Lascoumes & Le Galès, 2007). In this sense, standard-setting and performance management have become privileged instruments in policy reforms (Le Galès, 2016). In Table 1, I apply the existing taxonomy of policy instruments to the SAWA reform package.

Table 1. School autonomy with accountability: Its logic, instruments, and tools

Element	Definition	Type of Instruments	Techniques & Tools
1. Autonomy & Decentralization	Deregulation or devolution of authority over decision-making from centrally administrative units to lower management units	Legislative and regulatory	<p><i>Decentralization</i></p> <p>1. Decentralization of education governance</p> <p><i>Autonomy</i></p> <p>2. Schools (principals or boards) make decisions about budget allocation</p> <p>3. Schools make decisions about curriculum or learning contents</p> <p>4. Schools (principals or boards) make decisions about staffing</p>
2. Accountability & Standardization	Central monitoring of performance through standardized metrics and indicators	<p>Legislation and regulations;</p> <p>Information systems;</p> <p>Standards;</p>	<p><i>Standardization</i></p> <p>5. National learning standards/curriculum for students</p> <p>6. Standardized national exams in primary education (ISCED level 1)</p> <p>7. Standardized national exams in secondary education (ISCED level 2 or 3)</p> <p>8. Standardized testing at subnational level</p> <p><i>High-stakes accountability</i></p> <p>9. Standardized exams to exit secondary education level (ISCED level 2 or 3)</p> <p>10. Exam results are publicly available at the school level</p>
Competition & Marketization	Creation of quasi-market conditions as a driver for improvement and reduction of inefficiencies	Incentive-based	<p><i>Competitive incentives</i></p> <p>11. School choice</p> <p>12. Per-capita funding</p> <p>13. School rankings or quality indexes</p> <p><i>Marketization of the education sector</i></p> <p>14. Private management of public schools (i.e. Charter schools)</p> <p>15. State subsidies to private schools (demand or supply-side)</p>
Performance-Based incentives	Establishment of a rewards and sanctions system to nudge actors towards desired goals	<p>Economic and fiscal;</p> <p>Standards;</p>	<p><i>Performance incentives</i></p> <p>16. Teacher bonuses linked to students' performance on standardized assessments</p> <p>17. School bonuses linked to students performance in standardized assessments</p>

Source: Self-elaborated.

As is well known, countries rarely adopt the whole package but rather ‘selectively borrow’ specific policies or even ideas from global reforms (Gilardi, 2010; Dolowitz & Marsh, 2000). Hence, by dissecting SAWA into its composing elements and associated policies, this dissertation advances a more nuanced understanding of traveling reform packages while considering the intrinsic features of policies and their likelihood to be resisted or adopted elsewhere.

1.2. SAWA in Latin America

Among world regions, Latin America has one of the longest and most intimate relationships with SAWA reforms. This history is rooted in IOs’ preference for Latin America to deploy and experiment with their policy reforms, from structural adjustment programs, privatization of public services, or government decentralization (Borges, 2018; Fernández, 2007; Meseguer, 2004; Morais De Sá e Silva, 2017, p. 201). The World Bank has been pivotal in advancing pro-privatization policies or large-scale learning assessments (LSAs) through conditional loans, pilot projects, or technical assistance, among other strategies (Edwards Jr et al., 2021; Hossain, 2023). In particular, the World Bank has been at the forefront of SBM and accountability reforms in Latin America (Barrera-Osorio et al., 2009; Demas & Arcia, 2015). For instance, the loans supported the adoption of SBM in Honduras, El Salvador, Guatemala, and Nicaragua during the early 1990s (Cardenas, 2009; Di Gropello, 2006). This innovative experience promoted decentralized decision-making by transferring families in rural or suburban areas significant authority over schools’ governance. However, most of these transformations were abandoned a few decades later (Ganimian, 2016). Similarly, the adoption of LSAs was significantly influenced by the World Bank’s agenda through loans and grants all over the region (Hossain, 2023)

The OECD is the other leading international actor advancing SAWA reforms in Latin America, a more recent influence compared to the World Bank. (Díaz Ríos & Urbano-Canal, 2023). Led by its Program for International Students Assessments (PISA), the OECD became a critical regional player in the field of education (Acosta, 2019) (see Chapter 3 for a detailed discussion). OECD's global educational dominance has solidified through a three-fold strategy that combines idea generation, knowledge mobilization, and policy brokerage (Niemann & Martens, 2018; Steiner-Khamsi, Martens, et al., 2024). These mechanisms render a soft-governance influencing strategy compared to typical coercive methods from other IOs. OECD's rise in Latin America took off with PISA but has grown with a large number of countries entering –Mexico (1994), Chile (2011), Colombia (2020), Costa Rica (2021)– or under the accession process –Brazil, Peru, and Argentina–.¹ Undoubtedly, PISA's paramount attention endowed the OECD with policy legitimacy and a unique policy-relevant platform in the region (Acosta, 2019; Avelar & Ball, 2019; Rivas & Scasso, 2019).

Despite the World Bank and OECD's promotion of SAWA in Latin America, their policies have had an uneven trajectory over the years and across different countries (Rivas, 2021). Put differently, SAWA policies have undergone moments of high receptiveness, resistance, and even abandonment after adoption. For instance, Chile is an early adopter of SAWA and a paradigmatic case of market mechanisms in education (Cox, 2012; Zancajo, 2019). After decades of pro-market reforms, it has recently enacted changes to mitigate the adverse effects of these policies, like ameliorating the stakes of the national assessment (SIMCE) or limiting private schools' autonomy as they perpetuated the system's inequalities (Falabella & Ramos Zincke, 2019). Similarly, Mexico

¹ Furthermore, Ecuador, Guatemala, Honduras, Panama, and Paraguay participated in PISA for Development between 2014 and 2016, and 16 Latin America joined the 2022 PISA cycle

adopted and abandoned high-stakes teachers' evaluations across two presidential terms (Moreno Salto, 2023). Hence, after decades of global circulation, a systematical examination of SAWA policies is long overdue. The extent to which Latin American countries adopted SAWA policies or whether they have relied on similar policies to govern their school system is what we explore next.

2. Policy transfer across time and space: An interpretative framework

It is time now to lay out the three analytical dimensions of this project and the interpretive framework used to make sense of our findings. First, the study centers on the temporal dimension inherent to policy transfer. The research tradition of transfer studies often asks why do countries adopt a certain reform? What has made it possible? Or, why have reform attempts failed in the past but succeeded years later? Such questions touch on the timing behind policy transfer, an aspect that received paramount scholarly attention (Marsh & Sharman, 2009; Porto de Oliveira & Osorio Gonnet, 2023). In other words, these questions require unpacking why, at specific moments, the different elements required for policy change –political support, the urge to change things, and a promise to solve such problems– made possible a given reform (Steiner-Khamsi, 2021). Likewise, diffusion studies also pay close attention to the pace at which policies are disseminated and their underlying causes (Dolowitz & Marsh, 2000). Dozens have shown the association between the rise of a liberal order and the spread of its landmark institutions, namely mass schooling, human rights regulations, etc. (Boli et al., 1985, p. 198; Buckner & Russell, 2013; Paglayan, 2024). Similarly, scholars have shown the linkages between neoliberal hegemony and the worldwide diffusion of pro-market reforms (Henisz et al., 2005; Weyland, 2007).

However, aspects linked to the institutionalization of policy reforms have received less attention in transfer studies (Steiner-Khamsi, Appius, et al., 2024). When it comes to reform packages, the sequence in which its different policies are adopted matters, as well as whether just some of them are incorporated while others are not. For instance, Henisz et al. (2005) have shown that in the spread of market-oriented reforms, most countries in their data set have adopted at least one policy, either privatization of public companies, the deregulation of the specific subsector, etc, but just a handful adopted the whole package (Henisz et al., 2005). This has several implications. On the one hand, it shifts our attention to the underlying causes behind the selective adoption, which will be discussed later. On the other hand, it also impacts the achievement of the reform goals, as its theory of change requires the interactive effect of the complete policy bundle to obtain its promised results (Henisz et al., 2005). In the case of SAWA, one could expect systems with a high degree of autonomy to show stronger accountability systems with potentially high stakes. Contrarily, if school choice is restricted, it will be rare to see the adoption of rankings or other competitive incentives.

A second relevant aspect concerns the lifespan or trajectory of a reform over time. When policies remain in place, one should wonder why this is the case and how they evolve to endure. Every new policy, particularly controversial ones, impacts the power relation among its affected agents (Mahoney & Thelen, 2010). Therefore, those who benefit will seek to maintain or increase their power returns, while losers will fight to dilute its effects or reverse its implementation. As institutional analysis has demonstrated, a policy can remain in place, but it might lose its power if its context has shifted (Hacker et al., 2015). Consequently, we must also pay attention to how policies ‘age’ and what it means for the parties involved (Morais de Sá e Silva & Porto de Oliveira, 2023).

The second dimension concerns the spaces over which we analyze phenomena. This study investigates SAWA spread following a threefold differentiation of the spatial dimension: (i) comparing against a global script, (ii) comparing across different contexts, (iii) and comparing cross-scalarly (Steiner-Khamsi, Appius, et al., 2024). First, it analyzes the diffusion of the SAWA ‘script’ from a cross-national and historical perspective, exploring what SAWA policies were selectively borrowed across the region, providing a cross-national overview of its dissemination over more than 40 years. Then, the study examines the cross-scalar interplay of policy actors, entrepreneurs, consultants, and so forth, driving the adoption –and resistance– of SAWA policies in Argentina and Colombia. In these cases, the articles emphasize different spatial and temporal aspects operating in the transfer process. For instance, Chapter 3 shows how Bogotá’s experience in the late 1990s influenced the national trajectory of SAWA policies in Colombia during the 2000s. Similarly, Chapter 5 explores how the changes to Argentina’s national assessments gave rise to subnational standardized testing among provinces or their rationale for participating in PISA as independent samples. The article provides a vertical comparison by exploring how cross-scalar dynamics and political factors shaped the adoption of standardized assessments within three provinces.

From a different angle, Chapter 4 explores the role of the OECD in influencing SAWA policies in Colombia, with particular attention to lesson drawing as a mechanism behind policy transfer. In sum, the transnational, national, and sub-national scales are jointly analyzed in each case to explain in detail, why and how SAWA policies were adopted, resisted, or institutionalized. Finally, the conclusion section discusses the cross-national comparison between Argentina and Colombia, seeking to answer one of this project’s guiding questions: why have two structurally similar cases yielded different results? Differently put, what is it in each case that explains what

we observe? (Simmons & Rush Smith, 2021). Therefore, compares how the interplay of existing institutional legacies, agents' interests, and policy ideas among the two countries explain the SAWA trajectory over time.

2.1. Institutions, ideas, and interest groups as engines of policy change

To analyze SAWA dissemination over time and space, I integrate macro-sociological processes with local institutional and political factors in a two-tier interpretative framework. On the one hand, I steer macro-level processes like economic coercion or global norm scripting as engines of diffusion. On the other hand, countries' institutional history and configurations, their norms and ways of operating, and their political forces mediate and translate global influence.

Building on traditional political science literature, the interpretative framework is grounded on the three 'I' of political analysis: ideas, interest, and institutions. To explain how I conceive these three elements, I will use an analogy from the field of cognitive psychology. Aaron Beck, the father of modern cognitive therapy, outlined a mental scheme formed by three elements: emotions, beliefs, and behaviors, which represented the baseline triangle of how our mind works. In Beck's view, although the three elements can influence one another, cognitive beliefs or ideas have more leverage in shaping emotions and behavior. In an analogous way, the three 'I' in our framework do not hold the same explanatory power. The approach taken here stresses actors' *agency* as carriers of policy ideas and responsible for coalition building to mobilize interests as more influential in explaining policy change than constraints or ritualized norms produced by the institutional environment or external pressures from international agents.

2.1.1. Unpacking institutional change

Among different institutional branches like rational choice theory, neo-institutionalism, or historical institutionalism, a common ground is their view of institutions as a set of formal or informal norms or rules with remarkable stability over time which constrain and guide agents' behaviors, shape their beliefs, and create power structure among actors (Hall, 2010; Meyer & Rowan, 1977). Institutions such as political systems, laws, or schools are not only historically constructed, but also, once in place, they produce feedback effects that help them maintain their stability over time (Pierson & Skocpol, 2002; Thelen, 2004). In this sense, power is inherently linked to institutions, as agents are in constant dispute to increase, maintain, or alter the existing payoffs.

Two ideas derive from this conceptualization. First, changing institutions is a complex endeavor, as it affects agents' power relations (Béland & Powell, 2016). Second and paradoxically, change is built into institutions because of the power struggles embedded within them (Stefes, 2019). Therefore, policy continuity can derive from continuing returns to power among interest groups, or, on the contrary, policy abandonment might be driven by a shift in coalition interests and alliances (Falleti, 2010). As discussed in Chapter 3, institutional change should not be seen only as a result of large external shocks, instead, in Hall's (1993) terms, can also derive from more subtle adjustments over time (Mahoney & Thelen, 2010).

Following the policy instruments literature, we conceptualize policies as institutions. Policy instruments like laws, standards, or incentives share key institutional features. Instruments provide a bounded frame for action, they limit what is feasible and what is not, and have power effects over agents, and are subject to change (Lascoumes & Le Galès, 2007). As such, instruments have both intended and unintended effects on policy actors and a group of constituents who

advocate for adopting the preferred choice at the expense of alternatives (Simons & Voß, 2018). Similarly, instruments, such as performance measurement and incentive-based ones, interact with each other, creating complex ‘mixes’ whose aggregated effects differ from what stand-alone ones would produce (Capano & Howlett, 2020). To illustrate this, standardized assessments joined by economic rewards for improved results are likely to nudge school agents towards enhancing testing scores, even at the expense of other subjects. Opposingly, if tests are inconsequential and results are only shown at aggregate levels, schools will likely see them as a bureaucratic requirement rather than anything affecting their core activities (Hallett, 2010).

Given their historical roots, institutions produce enduring effects over time, deemed as legacies. In its simplest terms, legacies mean that past factors have consequences in the present. Scholarship on institutional legacies is extensive, as not all institutions produce legacies, nor does every historical period shape the present in the same ways (see Wimmer, 2023 for a detailed discussion). We adopt a narrower definition of institutional legacies, both in their time-scope and effects, circumscribed to two main factors. First, we conceptualize legacies as influencing agents’ beliefs, attitudes, and interests (Díaz Ríos, 2024). Second, legacies are also consequential in shaping power relations among policy actors (Falletti, 2010; Béland, 2010). For instance, in Argentina, private subsidies to state schools have been in order since 1949, and among provinces, they represent between 7 to 20% of their public expenditure in education (Moschetti et al., 2023). Despite harsh critiques of this mechanism due to their segregative and socially regressive effects, they have remained central in Argentina as eliminating them touches on the interests of powerful groups, such as the catholic church (Moschetti et al., 2023).

Along this line of thought, policies seen as compatible with existing legacies –i.e., increasing the number of eligible schools for subsidies– are less likely to face resistance from

opposing groups. In contrast, policies that confront existing norms and shift power configurations among stakeholders will likely face greater opposition (see Chapter 5 for a detailed discussion). This does not mean that legacies directly obstruct policy change, but rather, they trigger different resistances – through institutionalized veto points, resource mobilizations, and so forth – requiring more significant transformations to be adopted and sustained.

2.1.2. Ideas and paradigms in policy change

Undoubtedly, ideas are another key driver behind policy change. Different policy studies theories have conceptualized and tried empirically operationalizing how ideational drivers influence the policy process (Béland, 2009, 2010; Hogan & Howlett, 2015). Peter Hall's (1993) landmark paper on policy paradigms and policy change has probably epitomized the notion that paradigmatic or fundamental policy change only occurs when the governing ideas underpinning instrument selection and their settings have also shifted. Policy ideas provide a framework to observe and understand the world, guiding agents' actions, decisions, and beliefs. Neoliberalism has been one of the dominant paradigms over the last four decades, articulating that supply and demand market dynamics should be extended across every societal dimension as the best and most efficient way to regulate human interactions (Ball, 2009; Harvey, 2005). More recently, ideas on 'evidence-based policymaking' or 'knowledge-based' governance have pressed decision-makers to ground their decisions in available research and data (Baek, 2022; Niemann & Martens, 2018; Pawson, 2006). Ideas carry with them value judgments, beliefs, and cosmo-views to which people and institutions adhere or repel. They also project power relations and seek to affect interests (Béland, 2010). Hence,, one must always historicize their origins and, more importantly, understand and identify who sustains them. Put differently, our focus should be on the agents, interest groups, and

organizations that articulate and circulate discourses and ideas as they reveal their stance ahead of key policy matters.

2.1.3. Interest and agendas through policy coalitions

The last piece of the interpretive framework concerns interest groups. In public policy, these groups within different policy subsystems organize themselves into coalitions (Díaz Ríos, 2024; Sabatier & Jenkins-smith, 1993; Sabatier & Weible, 2007). According to the Advocacy Coalition Framework, agents with multiple institutional affiliations will tend to group and form policy coalitions around shared beliefs to influence the political arena. Coalitions compete to influence the policy agenda and advance their preferred policies (Henry et al., 2014; Jenkins-Smith et al., 2019). In doing so, coalitions mobilize resources and engage in different activities to frame policy problems (Sabatier & Weible, 2007). From a conflict-based approach, the policy-borrowing and lending framework has demonstrated that agents reference international norms or policies as golden standards to create the urge for reform (Seiner-Khamsi, 2021). This practice, referred to as externalization, points to the use of global scripts to influence local policy reforms, although it does not imply a real connection between policies on the ground and the global models.

In our cases, relevant interest groups are members or consultants from IOs, teachers' unions, political parties, national elites, public servants, or private sector organizations. In Argentina and Colombia, actors from these groups have organized into coalitions, seeking to advance their beliefs and interests, either resulting in maintaining the 'status quo' or leading to policy change. Over the years, coalitions faced shifting environmental conditions –right or left-leaning governments, economic crisis, pressure from IOs during membership processes, etc– creating opportunities for action. Therefore, SAWA reforms will be supported or resisted by agents

perceiving themselves as potential winners or losers to the proposed changes (Díaz Ríos, 2024). Hence, to a large extent, the evolution of SAWA policies will be explained by how different coalitions navigated shifts in the societal mood or whether they could mobilize resources and political support in favor of their preferred policies. Policy changes in Argentina and Colombia, rather than groundbreaking overhauls, occurred through a careful mix of policy layering, institutional drift derived from environmental alterations, and the reinterpretation of existing norms leading to policy conversion (Hacker et al., 2015; Mahoney & Thelen, 2010).

3. Research design

To study SAWA diffusion across Latin America, this study is organized into two research strands to provide a comprehensive overview of regional trends combined with an in-depth analysis of two countries. The first strand (RS1) adopts a macro-level perspective, and it focuses on the regional spread of SAWA policies from 1980 until 2023. The second strand (RS2), which I call the ‘meso-level’ strand, explores, from a comparative and historical perspective, the adoption and evolution of SAWA policies in Argentina and Colombia since the 1990s. The cases are derived from the first strand, following a most similar–most different case selection criteria (Seawright & Gerring, 2008). The following sections outline the data collection strategies and analytical procedures used for each strand.

3.1. Macro-level strand (RS1): Researching SAWA diffusion

The macro-level strand (RS1) focused on identifying the adoption of each of the four SAWA elements –autonomy, accountability, competition, and performance incentives– through a set of specific policies across Latin America. To fulfill this research objective, RS1 entailed the creation

of a unique data set on the de-jure adoption of SAWA policies. The database build-up was inspired by similar projects looking, cross-nationally and historically, into education reforms globally. Among them, I analyzed the methodological approach used in the construction of the “World Education Reform Database” (Bromley et al., 2022) and in the “Education Policies and Systems across Modern History” dataset (del Río et al., 2024). Following these projects, the SAWA database build-up involved the following steps: (i) defining the variables of interest, (ii) defining the sample, (iii) identifying data sources, (vi) creating a standardized coding protocol, (v) transferring the coded data it into the dataset. Next, I outline the main methodological decisions.

3.1.1. Data collection process

First, based on the literature review on SAWA policies, I revised and refined its elements, resulting in n=17 policies to capture how the reform ideas are operationalized (detailed in Table 1). Those 17 policies represented an outcome variable, whose adoption I traced in each country. As stated, I focused on the de-jure or regulatory adoption of policies, not their factual implementation. Therefore, if a decree or law stated that schools have the right to modify the national curricula, I would have coded that this country adopted curricula autonomy, regardless of its implementation. Each policy constitutes a binary measure (0/1) in the dataset, coded one the year the policy was adopted (i.e., 1993). Second, I defined my sample of 33 Latin American (n=22) and Caribbean countries (n=11) according to the Economic Commission for Latin America and the Caribbean. Due to time and data availability constraints, the final sample was reduced to n=23 countries, shown in Table 2. Thirdly, data combined a number of different primary and secondary sources, using a similar approach to del Río et al. (2024, see p. 8-9). Sources were (a) country-level policy documents, legislation, and sectoral plans; (b) international organization publications; and

academic publications in peer-reviewed journals or Ph.D. dissertations in sociology, politics, or history of education. Each of the sources was downloaded and coded following the same pattern: the UN country code, the year, and the author, i.e., ARG_1993_Congreso_de_la_Nacion. Likewise, the complete citation list was included at the top of each country coding sheet.

To collect the data, I created a country coding sheet totaling 57 questions to capture the 17 SAWA policies and some additional background information for the country. For each of the policies of interest, the protocol followed a similar sequence of questions, as shown in Figure 3. While the question for policy adoption had only a binary answer (yes/no), subsequent questions were asked to provide enough data from the original sources that supported the coder's answer, as shown in Figure 3. The coding sheet template can be found in Appendix A. As part of the data collection process, I worked with two research assistants, Courtney Martinez, and Huilan Wu, who were responsible for coding 10 countries (5 each). To ensure inter-coder reliability, I organized two sessions dedicated to explaining the project's goals, the policy reform's contents, and the rationale for creating the database. Following that, we selected Chile, and each person coded it individually. Then we compared our answers to reach a standard criterion. Furthermore, the team had several available resources to tap into whenever they had doubts, including the definition of each of the policies and the rule of thumb for deciding on the inclusion criteria. Coders focus on countries over which they have mastery of the language, either Spanish, English or Portuguese, respectively. Before coding, the Principal Investigator (PI) collected the policy documents to answer the protocol, yet when information was not found, coders would do additional desk research by looking into additional sources. Each of the documents used is listed at the top of the coding sheet, and a copy is stored on the shared folder. When working on countries, the PI held weekly meetings individually to answer doubts and pose questions on the coding protocol. Once

coders have completed the coding sheet, the PI reviews them and checks on each question to ensure common criteria.

Despite the coding process centered on the de-jure adoption of our policy bundle, identifying the adoption date could be troublesome. This challenge arose for two reasons: (a) because documents do not include such information or (b) because there was conflicting information between different sources. A recurrent example of case (a) situations was school choice policies because often countries do not have specific legislation promoting or openly allowing for school choice, but rather, it was a standard practice where families were free to choose among their preferred public schools without state interference (i.e., Argentina or Venezuela). In contrast, case (b) situations that render conflicting information might arise when documents discuss the beginning of the implementation date versus adoption or if different sources mentioned distinct dates. In the first case, we stuck to the de jure adoption and not the first implementation. In the second case, the rule of thumb was to keep the date indicated by the primary source, such as laws, official policy documents, government decrees, etc. In the few cases where the discrepancy couldn't be solved or whether the data couldn't be found, I asked country experts to settle the discrepancy.

Table 2. Sample of SAWA in Latin America database

#	Countries	Region	Organization	Compulsory Schooling	Type of Governance
1	Antigua and Barbuda	Caribbean	Unitary	12	Centralized
2	Argentina	Latin America	Federal	14	Decentralized
3	Bahamas	Caribbean	Unitary	6	Decentralized
4	Belize	Latin America	Unitary	8	Decentralized
5	Bolivia	Latin America	Unitary	14	Decentralized
6	Brazil	Latin America	Federal	14	Decentralized
7	Chile	Latin America	Unitary	13	Decentralized
8	Colombia	Latin America	Unitary	10	Decentralized
9	Costa Rica	Latin America	Unitary	11	Centralized
10	Cuba	Caribbean	Unitary	9	Centralized
11	Dominican Republic	Latin America	Unitary	15	Decentralized
12	Ecuador	Latin America	Unitary	13	Decentralized
13	El Salvador	Latin America	Unitary	12	Centralized
14	Guatemala	Latin America	Unitary	12	Decentralized
15	Guyana	Latin America	Unitary	11	Decentralized
16	Haiti	Latin America	Unitary	6	Centralized
17	Honduras	Latin America	Unitary	12	Decentralized
18	Mexico	Latin America	Federal	15	Decentralized
19	Nicaragua	Latin America	Unitary	7	Decentralized
20	Paraguay	Latin America	Unitary	13	Centralized
21	Peru	Latin America	Unitary	14	Decentralized
22	Uruguay	Latin America	Unitary	14	Centralized
23	Venezuela	Latin America	Federal	15	Decentralized

Source: self-elaborated.

Once the 23 countries were coded, I transferred the information into the database. The database is organized in a panel structure, with an individual observation for the 44 years (1980-2023) on every variable, including country, year, years of compulsory schooling, decentralization law, national assessments, etc. The different variables appeared on the columns and were coded

zero (0) if the policy was not adopted in such a year and coded one (1) in the year of adoption. A necessary clarification is that if the adoption year was before the investigated period, it fell out of the dataset. For instance, Argentina established state subsidies to private schools in 1949, and Colombia has had its national large-scale assessment at the end of secondary education in place since 1968. These policies then are not considered in the analysis.

National Large Scale Assessments (NLSAs) for Lower Secondary Education (ISCED Level 2)	
22. Has “the country” ever adopted NLSAs for lower secondary education (ISCED level 2)?	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
Please insert the direct quote with the page number that helped you answer.	<p>Standardized assessment activities in Costa Rica began in 1986 and were originally carried out by the Research Institute for the Improvement of Costa Rican Education (IIMEC) at the University of Costa Rica, in response to a request from the Ministry of Education for a formative assessment of basic education.... The agreement between the ministry and IIMEC expired in 1996, and, the following year, the National Center for Educational Assessment was established as an agency of the Higher Council on Education. The center was dissolved in 1998, and assessment activities are now managed by the Division for the Control and Macro-Assessment of the Education System. (Ferrer, 2016, p. 81)</p> <p>Costa Rica monitors national learning outcomes through a standardised assessment of 6th and 9th grade students in a sample of schools. The assessment has been applied twice in primary education (2007 and 2010) and three times in lower secondary education (2008, 2012, 2014). It tests students' knowledge of the curriculum in core subjects, and also surveys schools' socio-demographic characteristics, resources and climate. (OECD, 2017, p. 102)</p>
23. What year has the assessment been adopted?	1986
24. What is the assessment name?	Bachillerato National Exam FARO la Prueba Nacional Estandarizada / National Standardized Exams (PNE)

11

Figure 2. Costa Rica coding sheet. Questions 22-24.

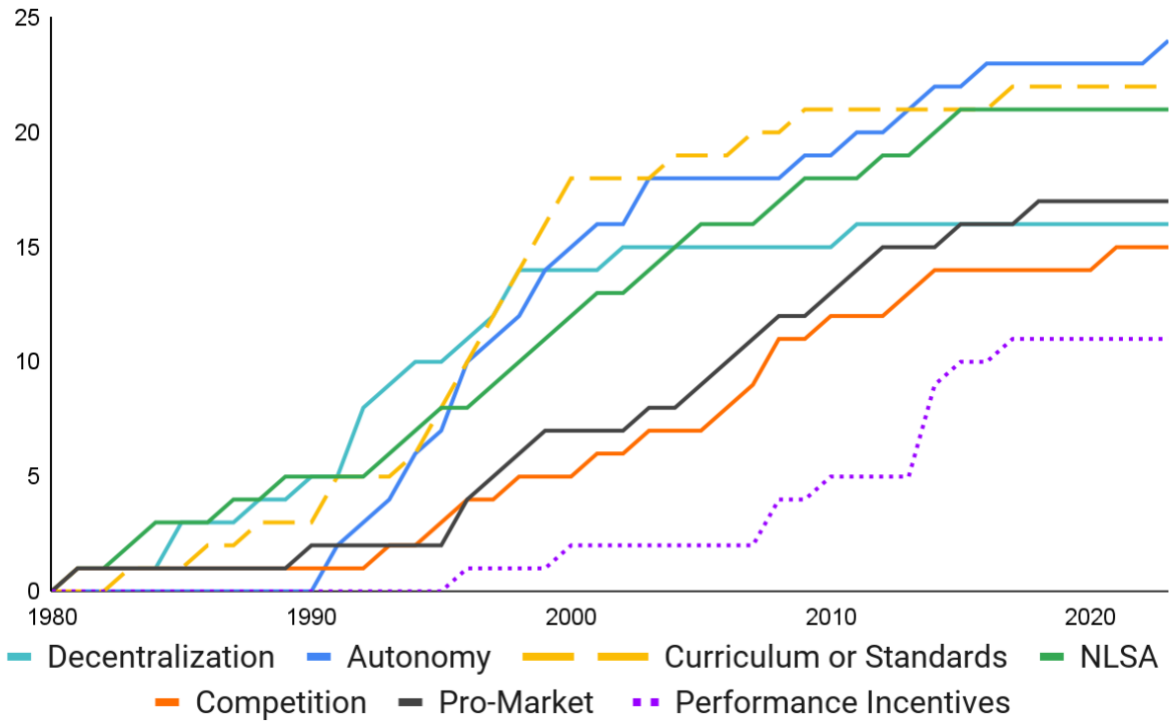
3.1.2. Analyzing SAWA diffusion and selecting cases

After consolidating the database, I have explored the main patterns derived from SAWA dissemination across Latin America to select the two case studies. Whereas a complete analysis of diffusion results is part of a forthcoming publication,² we can foreshadow some key patterns that illumined case selection: certain policies became universally adopted while others have had less traction. Clearly, large-scale assessments, learning standards, and curriculum autonomy have been widely adopted region-wide, whereas those policies fostering market-like dynamics had a modest spread. Table 3 shows the spread of the four SAWA elements of autonomy, standardization and accountability, competition and marketization, and performance incentives. The six diffusion lines disaggregate the elements even further, showing the spread of decentralization reforms independent from autonomy policies taken in aggregate terms –any dimension including curricula, budgeting or staffing. Likewise, learning standards or curricula, in yellow, are separate from national scale assessments, and so forth.

The regional spread of SAWA elements signals some interesting trends. In the first place, policies linked to managerial principles of increased autonomy, standardization and distance monitoring have been more widely adopted than those associated with neoliberalism, that is, marketization and competition. Every country in our sample (n=23) has introduced a curriculum or learning standards since the 1980s, while most (n=21) have at least one national standardized assessment. Not only that, but these policies rapidly expanded during the 1990s, reaching a ‘burn-out’ face in the early 2000s. In contrast, policies fostering competition as a driver for improvement or incentivizing agents with economic rewards are less ‘popular’ and grew almost a decade later.

² Esper, T. (Forthcoming). *Examining the diffusion of a global reform package (1980-2023): trends, drivers and explanations from Latin America*

Arguably, these trends might relate to the fact that performance-based accountability took off during the 2000s (Verger et al., 2024) or even the intrinsic features of such policies, creating more resistance among adopting countries.



Source: self-elaboration.

Note: Plotted lines show the earliest country-adoption of at least one policy linked with each SAWA elements. If a country adopted more than one policy tied to a given element, its first adoption is the one represented. For instance, Chile's NLSA correspond to 1982 when it implemented primary and lower secondary assessments, despite it also adopted a standardized tests for upper secondary education in 1988

Figure 3. Aggregate diffusion of SAWA elements in Latin America (1980-2023)

Undoubtedly, these initial trends on SAWA diffusion open several questions requiring further investigation on why that is the case. Some explanations are presented in conclusion, while others belong to ongoing work. Nonetheless, the two selected cases for in-depth exploration are exemplar of SAWA uneven diffusion. On the one hand, Colombia has practically adopted the 'complete' package since 2000. As shown in Table 4, between 1990 and 2015, Colombia implemented 14 of the 17 SAWA policies, ranking at the top of all Latin American countries. In

contrast, SAWA had less traction in Argentina, adopting just seven of all investigated policies, particularly reluctant to pro-competitive and performance-driven reforms.

Table 3. Adoption of SAWA policies in Latin America (1980-2023)

#	Countries	Autonomy				Accountability & Standardization				Competition & Deregulation				Performance Incentives		Total	
		Decentra- lization	Finance	Curricula	Staff	National Curric/ Learning Stand.	LSA Primary	LSA Low Sec	LSA Upp Sec	School Choice	Per capita finance	Rankings	Private Mgmt of Public Sch	Private Subs	Perf. Payment		Teacher Bonus
1	Antigua and Barbuda	-	-	-	-	2008	2008	1998	1998	2008	-	-	-	2008	-	-	7
2	Argentina	1992	-	1996	-	1993	1993	1993	1994	1962*	-	-	-	1947*	-	-	8
3	Bahamas	1992	-	-	2001	1982	1984	1994	1993	-	-	-	-	1996	-	-	7
4	Belize	-	2003	2009	2003	1999	1980	-	-	-	2010	-	2010	2003	2010	-	10
5	Bolivia	1994	-	-	-	1994	1980	1998	2000	-	-	-	-	-	-	2000	7
6	Brazil	1988	1996	1996	-	1997	1995	1990	1996	-	1996	2003	2012	-	2017	-	12
7	Chile	1981	1992	1991	-	1990	1982	1982	1988	1981	1981	1995	-	1981	2008	1996	12
8	Colombia	1988	1999	1994	1999	2006	2002	2002	1968*	-	2001	2014	1999	-	2014	2015	14
9	Costa Rica	-	-	-	-	2016	1986	1986	1988	-	-	-	-	2005	-	-	6
10	Cuba	-	-	-	-	1987	1996	2003	2004	-	-	-	-	-	-	-	5
11	Dominican Republic	1997	2000	1994	-	1994	-	1992	1993	-	2008	-	2018	1997	-	-	10
12	Ecuador	1998	-	2011	-	1996	1996	1996	2013	2008	-	2003	-	2011	-	-	10
13	El Salvador	1996	1991	1991	1991	1996	1993	1993	1997	1996	-	-	-	1996	-	-	11
14	Guatemala	1985	1996	1995	1996	1995	1997	1997	2004	-	-	-	-	2015	-	-	10
15	Guyana	1985	2014	-	2014	1990	2007	-	-	-	-	2021	-	-	-	-	7
16	Haiti	-	-	-	-	1998	-	-	-	-	2007	-	-	2007	-	-	3
17	Honduras	2011	1999	1999	1999	2003	1997	1997	2011	1998	-	-	-	-	2014	-	11
18	Mexico	1992	2013	2016	-	1998	1992	1999	2011	-	-	2008	-	-	-	2008	10
19	Nicaragua	1993	1993	1993	1993	1999	1996	2009	2009	2006	1993	-	-	2006	-	-	10
20	Paraguay	-	-	1998	-	1998	1995	1999	1996	-	1998	-	-	1998	-	-	8
21	Peru	2002	2003	1996	-	1997	1996	1998	1998	-	-	-	1981	-	2014	2014	11
22	Uruguay	-	-	2023	-	1995	1996	1998	1998	-	-	-	-	-	-	-	6
23	Venezuela	1990	-	1997	-	1985	-	-	-	-	-	-	-	1990	-	-	5
Total		16	12	16	8	23	18	18	18	8	8	6	5	13	6	5	200

Source: self-elaboration based on RS1 data.

3.2. Meso-level strand (RS2): A comparative sequential design

Complementary to the first strand, the RS2 adopts a comparative sequential design to explore the circulation of SAWA policies in Argentina and Colombia. The rationale for choosing Argentina

and Colombia was straightforward: countries shared structural similarities in key explanatory variables (shown in Table 4), yet as shown in Table 4, they portrayed variations in the policy ‘outcome’. Furthermore, these differences became even more noteworthy when comparing SAWA policies over time. Therefore, to make sense of results and the historical variations among the two countries, I followed a comparative sequential design.

The comparative sequential design draws from comparative historical analysis (CHA) and process tracing methods to provide an analytical and historical narrative of the underlying causes driving to specific outcomes (Mahoney & Thelen, 2015). Following CHA, which poses empirical questions about social phenomena that extend over long periods (Mahoney & Rueschemeyer, 2003), the comparative sequential design operates similarly but within shorter periods. Two other features are shared with CHA. First, the conceptualization of social phenomena as historically determined while offering a detailed account of long-term processes and their effects in producing social outcomes (Mahoney & Thelen, 2015). Second, the comparative examination of puzzling trajectories of arguably similar phenomena. For instance, why have structurally analogous cases rendered different outcomes, or why have structurally different cases resulted in similar results (Mahoney & Rueschemeyer, 2003)?

What characterizes the comparative sequential method is combining cross-case and within-case analysis. In doing so, the prime is in the sequence of events, how they are chained together, and in what ways they affect each other. In this sense, it accounts for the feedback effects of prior reforms, such as decentralization, affecting later ones (Falleti & Mahoney, 2015). Then, by evaluating relatively shorter periods compared to CHA, process tracing methods allow for a detailed analysis of the agents, entities and activities linking different events and producing the observed results (Beach & Pedersen, 2019). Hence, the comparative sequential analysis lies in

between the macro-analysis characteristic of CHA and the micro-level process-tracing orientation of agents' activities to examine the policy processes (Falleti, 2010). Following Sabatier's (2018) advice for understanding policy change through a coalition lens, studies should cover about a decade or more. Thus, we explore SAWA evolution over the last three decades.

3.2.1. The vertical, horizontal and transversal axes in cross-case and within-case analysis

In policy studies, scholars have termed the three analytical dimensions of the comparative sequential method as the vertical, horizontal, and transversal axes (Bartlett & Vavrus, 2016). The vertical axis references the different scales of intervention of the multiple agents involved in policy transfer, typically referred to as the global, the national, and the subnational (Bartlett & Vavrus, 2016). This heuristic distinction helps us elucidate the multiple spatialities in which policies, discourses, and agents circulate. For instance, using the vertical axis, one could create multiple cases within a single one by comparing the national developments with subnational ones and their mutually influencing dynamics (Beach & Pedersen, 2019). The horizontal axis refers to the comparison across similar scales and different cases, from comparing two provinces in the same country or well two nation-states (Bartlett & Vavrus, 2016). Finally, the transversal axis refers to the historical dimension, as is crucial in understanding why policies are adopted or resisted at certain times, how agents deal with them once in place, and how they change over time, affecting agents' relations on the ground, and so forth (Bartlett & Vavrus, 2016).

In sum, the RS2 takes a comparative sequential design to analyze the adoption and evolution of SAWA in Argentina and Colombia during the last three decades. In doing so, I pay attention to how similar phenomena influence each case differently—such as IOs' pressure to adopt

SAWA policies or the enduring consequences of governance decentralization—combined with the role of agents and different coalitions in advancing their agents to influence the policy process.

3.2.2. Case selection: Unpacking causal mechanisms

Case selection in qualitative comparative research requires a precise delimitation of the research objectives and a detailed understanding of each case context (Mahoney & Goertz, 2004; Seawright & Gerring, 2008). The two cases are structurally similar in their main variables yet they diverge in their outcomes (Seawright & Gerring, 2008), in this case, the adoption of SAWA policies. In particular, case selection criteria centered on explaining and un-blackboxing the underlying causes driving the differences in the trajectory –adoption, rejection, and life-span– of SAWA policies. The study focus is twofold in this sense. First, it aims to unpack the causal forces explaining why and policy transfer occurs or why sometimes it fails. Then, it examines an often-overlooked portion of the transfer process: the temporal evolution of adopted policies. Therefore, from all Latin American countries, I selected two causally homogenous cases that not only shared structural similarities –governance structure, population, presence and strength of unions, etc–, but where the explanatory drivers and mechanism explaining SAWA trajectory could be present (Beach & Pedersen, 2019; Collier, 2011; Mahoney & Goertz, 2004).

The comparative sequential design aims to shed light on the causal mechanisms connecting events leading to an outcome (Falleti, 2010). Mechanisms are theorized as a set of relatively stable processes that connect events and produce the observed results. Examples include normative pressure, policy learning, or institutional drift and conversion. Thus, to select the cases, researchers need deep contextual knowledge to assess whether the causal relationship between X and Y could exist (Beach & Pedersen, 2018, p. 18; Steiner-Khamsi & Morais de Sa e Silva, 2024). Second, a

mechanistic explanation of policy change is grounded in an ontological determinism, meaning that things happen for a reason, and a probabilistic epistemology, stressing that where we can never be fully certain of our knowledge of the world, we can be confident of the casual relationships based on the evidence we produced (Beach & Pedersen, 2019; Thelen & Mahoney, 2015). Finally, comparison between structurally similar or casually homogenous cases helps us elucidate what factors linking causes and outcomes explain the difference in the observed results. Hence, comparisons enable us to (1) find causes to social phenomena; (2) compare the unfolding of hypothesized mechanisms across causally homogeneous cases; (3) engage in confirming sufficient or necessary conditions for a given outcome using cross-case evidence (Beach & Pedersen, 2016; Falletti, 2010).

3.2.3. The cases of Argentina and Colombia.

As stated, cases were selected due to their difference regarding SAWA adoption, despite their structural similarities (presented in Table 4). To start with, both countries shared a common pathway during the 1990s and early 2000s but seemed to take part after that. In the case of Argentina, after a profound economic crisis leading to hyperinflation in the late 1980s, its government embraced neoliberal economic policies to stabilize the economy and restructure the state. During the presidency of Carlos Menem (1989-1999), the country followed the Washington Consensus economic recipes and succeeded in controlling inflation at the expense of increased poverty. In education, Argentina decentralized school governance from the national sphere to its provinces (Falletti, 2010). More importantly, the 1990s reforms transformed the national government's role from its historical 'teaching state' with a large monopoly over education towards the 'evaluatory state,' which is more fragmented due to the growing privatization (Narodowski &

Moschetti, 2015). At the time, Argentina took in school autonomy over educational curricula and its national large-scale assessment (Derqui, 2001), although it remained reluctant to market-like policies in education (Beech & Barrenechea, 2011).

Interestingly, despite other Latin American countries moving towards SAWA policies since the 2000s, Argentina has remained reticent to this agenda. For the most part, it has been governed by left-leaning political coalitions that, at least discursively, opposed IOs' neoliberal policies as part of their domestic political strategy (Beech, 2019). However, the country has participated in PISA since 2000 and recently aimed at joining the OECD, now under a second attempt after an initial rejection in 2018. Along the same line, after losing centrality and intensity in the first decade of the new century, Argentina's accountability system in education has been shaken by several reform attempts aimed at strengthening its influence over school governance (Diker et al., 2023; Rodríguez et al., 2018). Therefore, we find a vivid dispute of opposing groups struggling to shift Argentina's education policies into new avenues. Arguably, while only a few recent reform attempts succeeded or remained, this phenomenon poses new questions: what enables or constrains reform adoption and their institutionalization over time? Is it possible that certain structural elements, norms, or beliefs ruling Argentina's education system in the past might have changed? In other words, what are the consequences of past reforms over present ones?

Colombia also underwent decentralization reforms during the 1990s aligned with the Washington Consensus (Azóca & Giraldo, 1994). However, the most prevailing difference with Argentina is that right-wing governments have ruled Colombia for more than 40 years.³ Another stark difference is that school autonomy was established in Colombia in 1994, representing a

³ This hegemony has been broken in 2022, when Gustavo Petro was elected president.

triumph of teacher unions' fight for a pedagogical and professional revindication ahead of successive detrimental reforms (Suárez, 2023). Undoubtedly, the 1990s were a time of fierce confrontation between two powerful groups. On one side, national elites promoted neoliberal reforms and sought to diminish teachers' union power by atomizing their representativeness through school municipalization projects (Azóca & Giraldo, 1994; Rodríguez Guarín, 2023). On the other side, unions and minor left-wing political groups represented what was known as the "*Movimiento Pedagógico*" [Pedagogical Movement], a grassroots project that envisioned teachers as pedagogical leaders for social liberation (Suárez, 2023). The clash between these two groups transpired into the 1990s legislative reforms in Colombia. Decentralization was limited to '*Departamentos*' instead of Municipalities, seen as a threat to unions' membership homogeneity, and the 1994 General Education Law eliminated the national curricula at the expense of school autonomy, demanded by teachers.

Comparable to Argentina, Colombia faced a harsh economic and social crisis in 2001. However, instead of rendering a center-left government, the reforms gave rise to a 'neoliberal counter-reform' (Rodríguez Guarín, 2023). Since then, Colombia has begun a bumpy but steady pathway toward limiting school autonomy by installing a managerial paradigm in the education sector. Hence, when comparing Argentina and Colombia, we should account for the differences in educational decentralization processes, governments' ideology, IOs' ability to influence them, and the varying power configurations among interest groups to understand education reforms.

Table 4. Country-level characteristics of Argentina and Colombia

Indicator	Argentina	Colombia
Country Development Indicators		
<i>Population size (2021)</i>	45.3 million	51.1 million
<i>GDP per capita in USD (2021)</i>	10,535.12	6,0104,14
<i>Human Development Index (2021)</i>	0.842	0.752
Education		
<i>Government organization</i>	Decentralized, provincial	Decentralized, municipal (certified territorial authorities)
<i>Investment in education as % of GDP (2021)</i>	5,1	4,9
<i>Years of free and compulsory schooling</i>	14	12
Participation in International Large Scale Assessments		
<i>PISA 2022 (OECD) aggregate ranking (81 countries)</i>	66	64
<i>First participation cycle</i>	2003	2006
<i>ERCE (UNESCO) 2019 ranking (16 countries)</i>	14	12
<i>First participation cycle</i>	1997	1997

Source: self-elaboration based on World Bank (2024).

3.2.4. Data collection: Interviews and document analysis.

To answer why and how Argentina and Colombia diverged in their adoption of SAWA policies, I followed a qualitative approach to data collection, combining semi-structured interviews and policy document analysis. The following chapters (3, 4, and 5) present the methodological strategies, data collection procedures, and data analysis in great detail. Therefore, this section provides a summary.

Data sources combined semi-structured interviews with key informants and policy documents analysis for both countries. The rationale for interviews was to gain a understanding of agents' rationale, interests, and views behind their policy decisions, as well as how they navigated external pressures and political turmoils (Saldaña, 2021, p. 2). Likewise, document analysis served a twofold purpose. First, media outlets and newspaper articles allowed me to gain a better contextual knowledge of the political and social climate in each country. Then, policy documents portrayed the declared goals and intentions and the choice of instruments to achieve them.

Data collection for the case study occurred in three phases: the first between June and August of 2022, the second during December 2022, and the last in April 2023. I conducted n=93 semi-structured interviews with key informants and decision-makers, broken down in Table 4. Most interviews were in-person in Argentina and Bogotá, with some exceptions via Zoom. Then, I reviewed more than sixty documents (n=62), combining media outlets, laws, decrees, national sector plans, and policy briefs. In the case of Argentina, I conducted a thematic coding analysis of policy documents to gain contextual knowledge of the policy reform goals and instruments. In the case of Colombia, in addition to the thematic analysis, I coded the documents' citations to explore what and whose knowledge was used in adopting policy reforms (detailed in Chapter 4). As a result, I studied the citations using network analysis methods to unpack the knowledge networks shaping Colombia's policy discourse. Interestingly, the OECD and the World Bank ranked at the top of Colombia's citations in policy documents, a result used during interviews to further investigate the role of these IOs in Colombia's education reforms.

Table 5. Interviewed key informants per country. 4

Organization and role	Argentina	Colombia
Minister of Education		
<i>Top-ranked officials (i.e., ministers, vice ministers, national directors)</i>	11	9
<i>Advisors or high-level staff</i>	3	6
Education Quality or Assessment Agency		
<i>Directors or high-level officials</i>	7	5
<i>Advisors</i>	2	2
National Teachers' Union		
<i>High-level members</i>	4	5
<i>Affiliated teachers</i>	2	4
Provincial or local education authorities		
High-level provincial officials	6	-
High-level municipal officials	-	5
International Organization Officials		
<i>Country-level official</i>	1	3
<i>IO staff member</i>	1	1
<i>IO consultant</i>	2	2
Private sector		
<i>Private or charter school chain administrators</i>	-	2
<i>Leaders of philanthropic foundations</i>	2	3
<i>Leaders of private sector coalitions</i>	1	1
<i>Leaders of NGOs working on education</i>	3	-
Total	45	48

⁴ IRB (protocol number 23-327): Participants holding high-rank positions (i.e. former ministers, vice-ministers, etc.) were informed of the risks of participating as their anonymity will not be guaranteed given the small number of people having hold those roles.

Dissertation organization This dissertation followed the three-articles format. Therefore, chapter 4 contains the three articles presented in sub-sections 4.1, 4.2 and 4.3 respectively, followed by a conclusion section in Chapter 5. The research articles unpack the underlying reasons driving SAWA policies' trajectory in Argentina and Colombia. Each of them proposes a different focus of attention. In section 4.1, article 1 is titled "Changing paradigms: a historical analysis of school autonomy and accountability policies in Colombia," published in a Special Issue on SAWA reforms at the Nordic Journal of Studies in Educational Policy (Esper, 2024). The paper concentrates on the evolution of SAWA in Colombia in light of shifting policy paradigms in the country's education sector. Then, Section 4.2 presents Article 2, titled "*Not everything learned is transferred: Unraveling knowledge use and lesson-drawing behind school autonomy with accountability reforms in Colombia,*" currently under review at Globalisation, Societies & Education. This paper explores the influence of knowledge brokers in Colombia's adoption and institutionalization of SAWA policies. Finally, Chapter 4 includes the last article, "*The politics of standardized assessments: A multi-scalar analysis of a global reform in Argentina*" under review at the Journal of Education Policy. This paper centers on the evolution of different large-scale learning assessments in Argentina from a multi-scalar perspective. Standardized assessments, a pivotal policy of the SAWA package, have been at the cornerstone of political disputes in Argentina, being subject to multiple changes over time in response to shifting policy goals. Finally, chapter 5 is the conclusion section, connecting the dots among papers, answering the project research questions, and outlining a research agenda for the future.

In section 4.1, the first of the three research articles explores the trajectory of autonomy and accountability policies in Colombia from a historical and political perspective. In the

Colombian case, it is impossible to understand the country's turn towards managerial policies since the 2000s without unpacking the effects of past reforms. Arguably, when observing the evolution of Colombia's education system from a distance, the magnitude and consequentiality of the 1990s reform immediately called our attention. In 1991, Colombia adopted its modern constitution and shortly after passed several laws to restructure its governance organization. As imaginable, Colombia was not immune to neoliberal ideas and the Washington Consensus policies. However, the neoliberal policy repertoire and its local allies confronted fierce opposition from different groups in the country, where teachers' unions stood out among them. This resulted in a new institutional layout that merged two opposing paradigms. On the one hand, neoliberal views foster decentralization, state retrenchment, and more significant involvement of private actors (Azóca & Giraldo, 1994; Edwards, Moschetti, et al., 2023). On the other hand, democratic values promote teachers' autonomy as pedagogical leaders in the context of inclusive education for all (Rodríguez Guarín, 2023). As a result, Colombia embarked on an educational decentralization process, including significant teacher autonomy on curriculum matters, which was unparalleled in the region.

The 1990s reforms embodied an ongoing confrontation of these two paradigms represented by opposing coalitions that marked modern education policies in Colombia. However, since the early 2000s, the country has slowly departed from the democratic and inclusive tenets present in the 1990s reforms towards increased managerialism in school governance. Considering these developments, the paper asks two research questions: (1) how have school autonomy and decentralization policies shaped the trajectory of SAWA in Colombia? (2) how did the dispute between political elites and teachers' unions influence the adoption and trajectories of SAWA

policies? To answer these questions, the paper adopts a qualitative methodology combining interviews (n=48) with key informants and stakeholders and document analysis.

Since the 2000s, Colombia has deployed a new education reform agenda around three axes: standardization, performance monitoring, and improvement incentives. This three-legged model was operationalized by several new policies. First, Colombia launched a new large-scale assessment program in 2002. The assessment moved from a peripheral role towards a core policy in about a decade, enabling distance monitoring in a highly decentralized system. In parallel, the government installed learning standards, which, together with assessments, constrained teachers' autonomy. Finally, from 2010 onwards, economic and symbolic rewards were aimed at increasing school competition to drive quality improvements. However, while standards and assessments for accountability purposes consolidated, pro-competition policies were rapidly abandoned. As shown in the article, these policies reanimated the content between national elites and teachers' unions. In each case, the relative strength and ability of the different players combined with the institutional levers derived from the 1990s reforms drove the success or failure of the proposed reforms.

The main argument conveyed in the paper is that the adoption and consolidation of SAWA policies as a governance model responded to a dominant managerial paradigm endorsed by national elites since the 2000s. The installment of a new governance model was not preceded by a radical crisis or external shock. Instead, successive and piecemeal changes render incremental but significant institutional transformations through policy layering, drift, and conversion mechanisms (Hacker et al., 2015). The accumulation of these 'minor' shifts and their consolidation over time resulted in the institutionalization of a new school governance paradigm. Thus, the paper shows that policy change is not only more likely to occur in an incremental fashion, but also that this often called 'minor' shift can render significant institutional and political effects over time.

In section 4.2, the second article deals with two mechanisms behind policy transfer: the normative pressure from global scripts and the role of policy learning or lesson-drawing. As argued, SAWA “script” has been actively diffused by the OECD and the World Bank. However, it is not the only game in town. Knowledge-based regulation or evidence-based policymaking is another global mandate lingering over countries worldwide, also driven by IOs (Seitzer et al., 2023). In this matter, the OECD has worn two hats: it champions SAWA reforms globally (Verger et al., 2019) and uses knowledge dissemination and policy brokerage to influence countries’ policy decisions (Niemann & Martens, 2018). Interestingly enough, Colombia’s principal policy planning authority calls itself an ‘evidence-based’ organization, and the country recently joined the OECD in 2020. Hence, in light of Colombia’s adoption of SAWA policies and its OECD membership (2011-2018), this paper wonders: (i) how and whose evidence was used in shaping Colombia’s education policy between 2002 and 2022? (ii) To what extent has lesson-drawing driven policy transfer in Colombia’s SAWA policies? (iii) In what way does evidence use lead to policy learning as a driver of policy transfer?

To answer these questions, the study adopts a mixed-methods design. First, it uses network analysis to explore the knowledge networks in Colombia’s education sector by studying citations in policy documents (n=25) between 2002-2022. Then, it unpacks network findings with in-depth interviews with national and OECD decision-makers. The article offers three relevant findings. First, coincident with its accession process, the OECD emerged as Colombia’s most cited knowledge source in policy documents. Second, Colombia’s use of citations seemed to conform with an environmental requirement of ‘knowledge-based’ policymaking but decoupled from the sources and knowledge used by decision-makers regarding SAWA policies. Finally, the paper stresses the importance of contextual compatibility in adopting policy recommendations. In other

words, while the OECD advocated for specific reforms, such as the re-introduction of national curricula, the potential domestic backlash from unions prevented decision-makers from following the OECD's advice despite coinciding with it.

Then, the third article in section 4.3 centers on the Argentine case, zooming into the adoption and institutionalization of large-scale assessments. Large-scale assessments (LSAs) became a global policy mandate symbolizing educational accountability and quality monitoring for countries (Furuta, 2022). The global 'testing fever' sparked the rise of national, international, and subnational tests (Addey et al., 2017; Bonal et al., 2023; Kamens & Benavot, 2011; Piattoeva & Vasileva, 2023), representing somewhat of a race toward educational monitoring and improvement. Hence, why countries adopted LSAs and how they translated them has been a critical *focus* of scholarly attention in recent years. Nonetheless, after the adoption, LSAs' trajectories have become far from homogeneous. Certain countries, such as India, Jamaica, and Guatemala, have abandoned international testing programs.⁵ Others have tended to increase LSAs' centrality for school governance, attaching economic rewards or higher stakes to their results, while a smaller group has prioritized de-escalating test consequentiality (Verger et al., 2024). Surprisingly, the evolution of LSAs ahead of global and domestic drivers and their mixing with other policies has been less explored. What explains the continuity of LSAs and how they change to endure is what this paper seeks to answer.

Similar to Colombia, Argentina faced profound governance reforms during the 1990s. Aligned with Washington Consensus ideas, the neoliberal government of Carlos Menem (1989) advanced educational decentralization and expenditure cuts (Burki & Perry, 1998). The national

⁵ See PISA participating countries over the years at:
<https://www.oecd.org/en/about/programmes/pisa/pisa-participants.html>

government rendered ‘school-less’, uptaking a new role as an ‘evaluatory state’ (Narodowski, Moschetti, et al., 2016). Hence, in 1993 Argentina launched a national standardized testing program, ongoing until today. However, the national LSAs continuity suffered continuous shifts in its centrality, uses, and goals for school governance, pivoting between dismantling and strengthening attempts (Diker et al., 2023). Beyond its national test, Argentina joined international programs early on, including PISA since its first round in 2000. However, as with national tests, PISA participation and its effects were drawn into ongoing political contestation, being used for every imaginable purpose but school improvement.

Historically, Argentina’s education system was highly centralized in the National Ministry of Education, despite the country’s federal organization (Narodowski, Gottau, et al., 2016). Hence, endowed with power and autonomy after decentralization, provinces started deploying their own educational agenda. During the 1990s, a handful of provinces adopted their own standardized test, abandoned shortly after (DINIECE, 2003). However, in light of the global rise of educational accountability and testing since the 2000s, the country has witnessed a resurgence of educational testing at provincial levels. On the one hand, three provinces have joined PISA as independent samples. This is far from expected in an impoverished country such as Argentina, with a consolidated trajectory in PISA participation. On the other hand, more than ten provinces have re-launched provincial LSAs in the last five years. Why that happened and what implications it has remains under investigation.

The article examines LSAs’ adoption, repurposing, and re-emergence in Argentina. First, LSAs are the core policy in operationalizing accountability models in education. LSAs have not simply spread globally but have taken different shapes and roles within countries’ accountability systems (Högberg & Lindgren, 2021; Verger et al., 2018). Second, Argentina’s LSAs have been

subject to fierce debates and contentious politics among coalitions over the years, aiming to advance contrasting policy agendas. Moreover, the country has had a pendular stance ahead of international organizations and neoliberal global discourses, where different coalitions have gotten closer or further to promote policy changes. Hence, a historical analysis of LSAs' evolution sheds light on how cross-scalar political dynamics have shaped education policies. Thirdly, LSAs are not limited to the national test but also encompass international testing programs and sub-national standardized assessments. Therefore, focusing on LSAs across different sites, the study examines the circulation, translation, and institutionalization of a global reform across time and space. Two research questions guide this paper: (1) what has driven the subsequent shifts in LSAs in Argentina at national and sub-national levels? How did policy coalitions and cross-scalar dynamics shape the changes of LSAs ahead of different global scripts over time?

In this case, the paper combines the policy coalition framework and mechanism-based explanations of policy change to explain the trajectory of LSAs at the national and sub-national levels. The study looks at the adoption, circulation, and evolution of different LSAs, including PISA, Argentina's national large-scale assessment, and three sub-national LSAs – those of Buenos Aires, Córdoba, and Mendoza. Data for this study consists of semi-structured interviews (n=45) and secondary sources, including policy documents and briefs, legislation, and academic literature. Then, using process-tracing methods, the article provides a thorough historical account of how cross-scalar interactions between national and provincial developments shaped the adoption, repurposing, and re-emergence of the different assessments in the country.

The conclusion in Chapter 5 is organized into three main sections. The first revisits the key components of the theoretical framework—policy paradigms, institutions, and coalitions—and examines their impact on educational policies in Argentina and Colombia from the 1990s to the

present. It also explores how international organizations and global discourses have shaped SAWA policies, highlighting the greater influence of Colombia's ties to the OECD due to ideological alignment compared to Argentina. The second section compares the adoption, institutionalization, rejection, or abandonment of SAWA policies in both countries. After discussing the study's contributions, this section closes with a proposed future research agenda

4. Results

This chapter includes the three research articles with an introductory page outlining their submission details.

4.1. Article 1: Changing paradigms: a historical analysis of school autonomy and accountability policies in Colombia

Journal: Nordic Journal of Studies in Educational Policy

Article Status: Published on September 5, 2024. Open Access.

URL: <https://doi.org/10.1080/20020317.2024.2396555>

Article's Table of Content

Introduction

The many faces of school autonomy with accountability

An institutional lens for education reforms

Education in Colombia: a historical battleground for opposing paradigms

Data and methods

The 'Bogotá' laboratory (1998–2001): the local rehearsal of SAWA policies

A dominant paradigm: standards, assessments and performance management (2002–2010)

Fine-tuning and consolidation: NSLAs at the heart of the new governance (2010–2014)

“The harder you push, the harder it resists”: a reform pendulum (2014–2018)

Discussion: critical junctures and institutional displacement

Concluding thoughts: a vernacular expression of a global reform

References

Changing paradigms: A historical analysis of school autonomy and accountability policies in Colombia

Abstract:

Spreading like wildfire, school autonomy with accountability (SAWA) policies have transformed education worldwide. Rooted in new public management, SAWA is a ‘policy bundle’ that links greater decision-making at the school level with standardization and monitoring practices like large-scale assessments and performance indicators. As with every traveling reform, SAWA varies among countries, influenced by prior policies, the government's ideology, and veto players’ influence. However, studies often focus on individual elements of the SAWA bundle, leaving gaps in our understanding of why and how certain aspects are adopted while others are not. Thus, this paper addresses this gap by analyzing SAWA in Colombia, examining how policy paradigms and institutional legacies have shaped its trajectory – from its adoption to the present. Data combines policy document analysis and interviews with key informants (n=48). The study shows the 1991-1994 reform period as a critical juncture in Colombia, consolidating a decentralized governance system dominated by a democratic educational paradigm. However, a counter-reform in 2001 produced significant effects under an opposing sign. Eager to regain political and administrative control in education, Colombia’s national elites incrementally deployed large-scale assessments, learning standards, and system-wide performance monitoring to govern a highly decentralized system. Over the years, education standardization and outcomes-based policies intensified, articulating a ‘quality assurance’ discourse around SAWA policies. While market-like policies incentivizing competition and higher stakes were rapidly abandoned, SAWA policies represent a managerial turn in Colombia’s education.

Keywords: historical institutionalism, education reform, New Public Management, national assessments, policy transfer

Introduction

In the last decades, dozens of countries worldwide have adopted school autonomy with accountability (SAWA) policies as their governance model for education. The SAWA reform consists of a ‘policy bundle’ that promises to improve the quality of education everywhere (Steiner-Khamsi et al., 2024; Verger et al., 2019a). As two sides of a coin, autonomy transfers decision-making to schools, while accountability policies create multiple data points for centralized monitoring or parental choice (Verger et al., 2019a; Lennert da Silva, 2022). However, adopting, countries often cherry-pick policies suitable to their goals and contexts. For instance, England, Australia, and Chile have merged standardized testing with pro-market reforms, creating high-autonomous-high-stakes accountability systems (Parcerisa et al., 2022; Wilkins et al., 2024). Others, such as Italy or Spain, have advanced the autonomy agenda as a driver for innovation and school improvement, often combined with a low-stakes accountability model (Mentini & Levatino, 2023; Pagès, 2021). Therefore, SAWA layout depends on how and what policies from the bundle are mixed and merged with each country’s agendas, pre existing norms, and institutional legacies.

As in every reform, its underlying ideas are crucial drivers to produce change. Policy paradigms offer a cognitive and normative framework that guides the adoption of policy goals, instruments, and settings (Béland, 2007; Hall, 1993). In the case of SAWA, it is based on New Public Management (NPM) principles of devolved authority, competitive dynamics, standardization, and outcomes-based governance (Gunter et al., 2016; Klijn, 2012; Verger et al., 2019a). However, these paradigms never translate purely into actual policies. Rather, we observe blends of multiple paradigms, where, typically, one set of ideas is more salient in guiding policies than others (Hogan & Howlett, 2015). Likewise, the receptivity of ideas or the ability of local constituents to materialize them into policies transpires between domestic political conditions and

existing institutional norms (Díaz Ríos, 2020; Pagès, 2021). Thus, in each case, we must unpack how SAWA ideas are translated and which factors delimit the adoption of its different elements on the ground.

In this context, Latin America has not remained immune to SAWA spread while countries display significant variations on what policies and to what extent were adopted (Rivas, 2021). For instance, test-based accountability –understood as the use of standardized assessments attached to rewards or sanctions for school governance– prospered in Chile, Mexico or Brazil (Moreno Salto, 2023; Rivas, 2021), while this agenda found stronger pushback in Argentina or Colombia (Díaz Ríos, 2020). Similarly, in Central American countries autonomy as school-based management policies thrived and perished between the 1990s and 2010s (Ganimian, 2016), while it remained limited in southern cone countries (Cardenas, 2009). Nonetheless, as most studies analyze only a subset of SAWA policies, they neglect what explains selective borrowing processes and limit our understanding of how policies in the SAWA interact with each other affecting school governance and power relationships among actors involved. This paper addresses this gap through the analysis of SAWA in Colombia, explaining how competing policy paradigms and existing institutional legacies have shaped its trajectory – from its adoption to the present.

In its own fashion, Colombia has progressively moved towards its version of SAWA policies since the 2000s (Baxter & León Cadavid, 2021; Díaz Ríos, 2020). Like most Latin American countries in the 1990s, Colombia underwent a complex decentralization process, with the particularity of also adopting school autonomy over the curriculum (Azóca & Giraldo, 1994; Cobo, 2022). The emerging institutions and governance organizations from this process portrayed a historical confrontation between two policy paradigms in Colombia's education sector (Giraldo-Paredes & De La Cruz-Giraldo, 2016; Rodríguez Guarín, 2023). On one side, a neoliberal and

managerial paradigm, endorsed by international organizations and political elites, promotes a minimal state limited to quality oversight and supportive of private involvement (Cruz, 2012; Giraldo-Paredes & De La Cruz-Giraldo, 2016). On the other side, teacher unions and left-wing coalitions support education as a universal social right with the democratic involvement of teachers as key agents of change (Rodríguez, 2018; Suárez, 2023). This contend between government elites and teacher representatives to impose their policy preferences would sign Colombia's education in the coming years.

To analyze the adoption and evolution of SAWA in Colombia, this paper delves into three interrelated literatures: historical institutionalism, power distributional perspectives, and policy instruments (Hacker et al., 2015; Le Galès, 2016). In particular, it investigates: (1) How have school autonomy and decentralization policies shaped the trajectory of SAWA in Colombia? (2) How did the dispute between political elites and teachers' unions influence the adoption and trajectories of SAWA policies? To answer these questions, I follow a qualitative approach combining semi-structured interviews (n=48) with key informants and policy document analysis. By dissecting the Colombian case, the study offers two main contributions. First, it advances the literature on policy trajectories by accounting for the role of ideational drivers and political forces behind policy change. Second, it expands the literature on policy transfer by explaining how global reforms are adopted and translated in the Latin American context, where SAWA continues to spread.

The many faces of school autonomy with accountability

As a trojan horse, SAWA global circulation reanimated NPM ideas of standardization and outcomes-based governance across education systems worldwide (Högberg & Lindgren, 2021;

Verger et al., 2019a). NPM has become an umbrella term containing two schools of thought. One is of public choice theory, which advocates for market-like and competitive environments (Bevir, 2009). The other, business managerialism, supportive of transferring private methods into the public sector (Bevir, 2009). Beyond these technical or economic nuances, NPM ideological essence lies in reducing the role of the state into a target-setter and controller (Gunter et al., 2016). In this line, SAWA theory of change can be defined as follows. First, autonomous school agents have freedom to choose their preferred strategy to meet centrally defined and monitored academic targets. Then, school results become publicly available to enable choice and parental accountability, while improvement is incentivized through performance rewards and other competitive dynamics (Steiner-Khamsi et al., 2024; Verger et al., 2019a). However, the domains and extents of school autonomy –including its budget, staffing, curricula, and organization– (Lennert da Silva, 2022) and the types of accountability regimes –with higher or lower stakes or hierarchical vs. horizontal types– (Högberg & Lindgren, 2021) vary widely depending on the purposes, uses, and predominant institutional legacies where SAWA is ingrained.

Despite the coupling between autonomy and accountability is never linear, it is possible to identify predominant models across different contexts (Parcerisa et al., 2022). In Anglo-American countries, where NPM advocates have advanced pro-market reforms to the state, school autonomy is rooted in managerialism. Typical of market models, the state acts as a target setter and system oversight, where principals emulate middle managers having to achieve externally defined goals (Lundström, 2015). Thus, in deregulated school systems like England or the U.S., high-stakes test-based accountability models enable distance monitoring, performance sanctions, and school choice (Högberg & Lindgren, 2021; Wilkins et al., 2024). Nonetheless, this does not mean a perfect correlation. For instance, in Chile autonomy is limited, while schools are subject to both

administrative accountability and pressure from high-stakes testing and pro-choice policies (Díaz Ríos, 2020).

In other cases, school autonomy stems from professional training models positioning teachers as subject-matter experts entitled to make teaching and content-related decisions, typical of Nordic countries (Frostenson, 2015; Voisin & Dumay, 2020). In such cases, accountability regimes are often low-stakes and follow a ‘quality assurance’ rationale, similar to decentralized systems across Latin America (Parcerisa et al., 2022; Rivas, 2021). The quality assurance model offers a weakly articulated theory of change among instruments, processes, and desired improvements, but it relies on distance monitoring through performance measurement to govern schools (Verger et al., 2019b). This model avoids sanctions linked to performance (Díaz Ríos, 2020), although it is frequent to observe reputational consequences for schools and teachers when a certain degree of choice is allowed or test results are published (Bunar & Ambrose, 2016). While these examples do not exhaust all SAWA alternatives (see Parcerisa et al., 2022), they show the complex and non-linear nature of autonomy and accountability policies in education.

An institutional lens for education reforms

To analyze the trajectory of SAWA policies in Colombia, I draw on institutional theory by cross-fertilizing three approaches: historical institutionalism, power distributional perspectives, and policy instruments scholarship. These interlinked analytical frames provide a comprehensive conceptual toolkit to examine the evolution of policies and their effects over time (Mahoney & Thelen, 2010; Le Galès, 2011; Stefes, 2019). Their common ground lies in exploring how and why institutions change to endure. Institutional change is at odds with the widespread view of institutions as “relatively enduring features of political and social life (rules, norms, procedures)

that structure behavior” (Mahoney & Thelen, 2010, p. 4). The analysis of temporal continuity has been central to institutional scholars, explaining the role of lock-in effects, feedback loops, or isomorphism as major processes behind invariance (Stefes, 2019). However, institutional theory has provided thorough accounts of macro, meso, and micro-level processes producing institutional change (Béland, 2007; Hacker et al., 2015).

First, historical institutionalism presents explicatory arguments for complex puzzles by analyzing simultaneous unfolding processes and their sequence over time (Pierson & Skocpol, 2002). Historical institutionalism pays attention to temporal sequencing, political agendas, actors’ interests, and broader contextual factors at multiple scales that can lead to institutional changes or continuity (Pierson & Skocpol, 2002). This literature often tapped into external shocks or radical contextual shifts as drivers of institutional transformation (Mahoney & Thelen, 2010). In a way, close to punctuated equilibrium models, historical institutionalism looks at critical junctures as pivotal moments in which actors can alter relatively stable structures (Stefes, 2019).

In this line, Peter Hall (1993) introduced a model of institutional change along three key variables that could yield first, second or third-order policy changes: “(i) the overarching policy goals; (ii) the techniques or policy instruments used to attain those goals, and (iii) the precise settings of these instruments” (p. 278). First-order changes happen when only the instrument settings are altered, while second-order change occurs if a new set of instruments is chosen to achieve the same overarching policy goals (Hall, 1993). However, third-order or paradigm shifts occur when policy goals, instruments, and settings are replaced, producing an institutional displacement, as in the transition from Keynesian economics to fiscal policy under Thatcher (Hall, 1993). Nonetheless, from this original stance to the present, numerous works have shown that

significant institutional change or paradigm shifts also can occur through modest adjustments over extended periods (Hall, 2010; Mahoney & Thelen, 2010).

Linked to historical institutionalism, the power distributional perspective shows how endogenous and more subtle institutional variations yield consequential institutional shifts (Béland, 2007; Hacker et al., 2015). Rather than seeing institutions as stable formations, this approach argues that “dynamic tensions and pressures for change are built into institutions” (Mahoney & Thelen, 2010, p. 14). Therefore, the interaction between the political context, agents of change, and institutional features is crucial for understanding variations. The power distributional lens proposes four mechanisms of institutional change: displacement, layering, drift, and conversion (Mahoney & Thelen, 2010). First, displacement happens when a new set of rules replaces old ones, like Hall’s (1993) third-order change. Second, layering produces incremental modifications by adding new rules, revisions, or amendments to existing ones (Béland, 2007). Thirdly, drift occurs when rules remain stable, but drastic environmental changes alter policy outcomes (Hacker et al., 2015). Last, conversion results from agents reinterpretation or twist to existing norms, creating different outcomes (Hacker et al., 2015). Accordingly, layering, conversion, or drift effects cannot be judged *ex-ante*. The analysis must consider how they unfold over time vis-a-vis other political and contextual factors.

Relatedly, policy instruments literature shares with historical institutionalism that institutions portray the political legacies of historical struggles over power and resources. From a policy sociology approach, instruments are conceived as institutions as they set expectations, involve a fight for resources, create boundaries, and are subject to change (Capano & Howlett, 2020). Instruments are not neutral devices as they reveal “a (fairly explicit) theorization of the relationship between the governing and the governed ”(Le Galès, 2011, p. 14), unveiling

policymakers' political and ideological preferences. Therefore, certain constituents or actors aim to advance their preferred policy instruments at the expense of others, seeking increasing returns to their relative power (Steiner-Khamsi, 2016).

In essence, zooming into which instruments are advanced, for what reasons, and by whom unearths the power dynamics shaping the adoption process and its outcome. Undoubtedly, event sequencing is fundamental to recognize why, at certain moments, specific policies are adopted (Porto de Oliveira & Osorio Gonnet, 2023; Steiner-Khamsi, 2016). Thus, the analysis requires examining changes in the political environment or stakeholders' ability to employ their veto power (Mahoney & Thelen, 2010). Similarly, threshold effects, in which specific processes acquire a critical mass to gain significance, can alter power relationships and open windows for change (Béland, 2007; Porto de Oliveira & Osorio Gonnet, 2023). Therefore, the joint account of historical processes, types of institutional changes, and agents' position vis-a-vis instruments enables a thorough explanation of policy trajectories.

Education in Colombia: A historical battleground for opposing paradigms

During the 1990s, two competing policy paradigms underpinned a heightened decentralization process in Colombia. One paradigm has been coined as 'democratic education', rooted in human rights principles and portraying education as a public good guaranteed by the state for all (FECODE, 1991; Pulido Chaves, 2014; Rodríguez, 2018). This paradigm was supported by the Colombian Federation of Educators (FECODE, in Spanish) –grouping teachers unions nationwide– and left-wing political parties, advocating for strong state involvement in educational financing and provision (Giraldo-Paredes & De La Cruz-Giraldo, 2016). The second, grounded in neoliberal and managerial ideas, aimed to reduce the state's role and transform the public sector

by incorporating private sector practices like outsourcing, standardization, and distance monitoring (Edwards et al., 2023; Miñana Blasco, 2010). This one was sustained by national elites and right-wing parties governing Colombia, in alignment with international organizations driving decentralization reforms across Latin America (Burki & Perry, 1998; Giraldo-Paredes & De La Cruz-Giraldo, 2016). The two agendas would transpire into the 1990s policies and legislation, rendering a complex decentralized governance structure with enduring future consequences.

In 1991, the new Political Constitution proclaimed Colombia as a Social State of Law that guarantees the right to education for all, while also laying the ground for decentralization (Azóca & Giraldo, 1994). The Constitution's social rights approach was forged in a National Assembly formed by diverse stakeholders, including key figures from FECODE, such as Abel Rodríguez (Bocanegra Acosta, 2010). In contrast, President Gaviria's 'Education Liberalization Plan' (1991-1994) advanced a neoliberal reform reducing the state intervention by transferring schools' administration to municipalities and fostering private actors' involvement in educational provision (Edwards et al., 2023). The tension between paradigms grew as the Constitution required a new law to regulate the right to education, inspiring Law Project 120 aligned with Gaviria's neoliberal policies (Azóca & Giraldo, 1994). However, FECODE mobilizations and strikes repealed Law Project 120, gaining a crucial role in subsequent law negotiations (Suárez, 2023).

Sanctioned in 1994, General Education Law 115 consecrated many of teachers' historical demands, foundational to the democratic education paradigm. During the 1980s, the political activism of intellectuals at the heart of pedagogical universities merged with FECODE's resistance to the 'teacher-proof' curriculum reform, giving birth to the 'Pedagogical Movement' (Pulido Chaves, 2014; Rodríguez, 2018). Founded in 1982, the Pedagogical Movement revindicated a professional conception of classroom teachers as educational intellectuals with a transformative

role in society (Cobo, 2022; Suárez, 2023). The 1994 General Education Law moved in this direction, eliminating the national curriculum, establishing school autonomy, and creating the Institutional Educational Project (PEI in Spanish), where schools should articulate their curricula and pedagogical orientation with local demands. In line with the 1991 Constitution, the new law consolidated education as a social service to be provided by the state, ruling out municipalization (Giraldo-Paredes & De La Cruz-Giraldo, 2016; Suárez, 2023).

However, the 1994 General Education Law also deepened education decentralization, fostering a democratic involvement at subnational government levels by reconfiguring the NME functions and responsibilities (Azóca & Giraldo, 1994). Since then, the NME has been responsible for setting policy guidelines, overseeing education quality, financing subnational entities, and organizing teachers' careers (Azóca & Guirado, 1994). Schools are locally administered by 'Certified Territorial Entities' (ETCs in Spanish), a subnational unit whose functions were restructured in 2001 (Edwards et al., 2023). At present, ETCs include Colombia's 32 departments and 65 municipalities above 100,000 inhabitants, responsible for administering schools and their staff, assessing principals' performance, developing school improvement plans, and having schools' information systems, among others. Thus, on paper, ETCs have significant degrees of political and administrative autonomy (Edwards et al., 2023). However, Colombia's decentralization efforts have faced severe economic and capacity constraints as ETCs should use their resources to finance their policies, thus reducing the exercise of autonomy. (Edwards et al., 2023).

In 2001, a severe economic crisis enabled Andrés Pastrana's government to advance a significant 'counter-reform' to the 1990s one (Rodríguez Guarín, 2023; Suárez, 2023). The Legislative Act 01/2001, then consolidated by Law 715 of the same year, modified the Constitution

and existing educational legislation, shifting towards neoliberal principles (Bocanegra Acosta, 2010; Rodriguez Guarín, 2023). Among the main changes, the counter-reform created a new education financing mechanism resulting in severe expenditure cuts in education; expanded decentralization by expediting the certification procedure for territorial entities; established a per-capita financing formula, and created economic incentives for quality improvements (Edwards et al., 2023; Pulido Chaves, 2014). Endowed with special powers during the crisis, Pastrana established a new career pathway for teachers under the 1278 decree, which coexisted with the former one (decree 2277 from 1979). The new 1278 decree organized professional progress under performance appraisal in contrast to the longevity criteria (Bayona-Rodríguez & Urrego-Reyes, 2019). The unconsulted reform became a source of multiple clashes between the government and FECODE in the following years.

Data and methods

This paper follows a case study methodology to explore, from a historical perspective, the trajectory of SAWA policies in Colombia. This approach allows tracing the development of phenomena over time and at different scales –international, national, local– to account for how processes unfold, what the actors’ roles are over time, and in which ways context shaped the process (Bartlett & Vavrus, 2016). I used two data sources to answer the research questions: semi-structured interviews (n=48) with key informants and secondary document analysis. For the interviews, I developed a semi-structured protocol addressing four key areas: (i) the main challenges and changes in Colombia’s education sector in the last decades; (ii) the rationale behind SAWA instruments adoption and the participants’ role; (iii) the different position among most

relevant stakeholders; (iv) the process of instrument adoption, including enabling factors, policy alternatives, and contextual adaptations.

Data collection took place between June and July 2023, where I combined purposeful sampling criteria and snowballing techniques for identifying interviewees (Creswell & Creswell, 2023). Initially, I reviewed policy documents, media outlets, and academic publications, and I developed a list of potential interviewees based on participants' roles during the period under investigation. After each interview, I asked participants for introductions with other informants already identified or mentioned during the interview. The sample includes high-level officials from the NME, former Ministers, vice-ministers, and chiefs of staff (n=17), top-rank officials from the Colombian Institute for Educational Evaluation (ICFES in Spanish) (n=8), and education heads of ETCs in Antioquia, Bogotá and Barranquilla (n=3, both identified as "GOV"). Furthermore, I interviewed leaders and members from FECODE (n=6, "UNI"); World Bank, IDB, and OECD staff or consultants (n=5, "IOs"); leaders of private sector organizations working in education (n=5, "PRIV") and academics (n=4, "ACAD").

Most interviews were in person in Bogotá, with some exceptions via Zoom, spanning 60-100 minutes each. Participants' identities are masked using general tags to classify their roles. Finally, interviewees were recorded with participants' approval and later anonymized and transcribed for the analysis.⁶ During interviews, I asked participants about the main changes in Colombia's education policies in recent decades, the purpose and uses of national large-scale assessments (NLSAs), their views on decentralization, school autonomy, accountability, and other

⁶ The project was approved by the Institutional Review Board (IRB) of Teachers College, Columbia University (IRB #23-237). Participants agreed to the project informed consent, where ethical considerations include potential risk of identity recognition beyond the researcher actions given the limited number of people occupying their roles.

policies, as well as their views on the role of national and transnational stakeholders and key events in shaping the policy process through the years.

Data analysis

The qualitative data analysis involved multiple interrelated steps to understand the sequence of events, stakeholders' roles, and the numerous intervening processes leading to the observed policy outcome. First, during my time in Colombia, I took field notes after the interviews to identify emerging themes. Later, I transcribed the interviews and wrote brief reflective memos of emerging issues related to SAWA instruments. Then, I created an initial codebook, refined through an interactive process while reading and analyzing the transcripts. In addition, I reviewed, collected, and examined three types of secondary documents. First, legislation and policy documents published by Colombia's government. Second, media outlets from the two largest newspapers in the country about key historical events (i.e., union strikes or protests due to peace negotiations) to gain contextual information. Third, unpublished documents, meeting minutes, or reports interviewees shared during fieldwork.

I conducted a theoretical thematic analysis guided by my analytical framework and research questions (Braun & Clarke, 2006). However, when analyzing the data and refining the codes and themes, I followed an inductive approach considering emerging factors in the data. In the analysis, I focused on three interrelated processes. First, the NME agenda vis-a-vis other political or social issues, such as the peace agreement negotiation (2011-2018). Second, the role of relevant stakeholders and veto players, particularly FECODE. Lastly, the constraints to the government's agenda derived from the economic and institutional environment, such as shifts in government fiscal resources, limitations due to schools' autonomy, or complexities stemming from

decentralized governance. In parallel, the examination of secondary documents served to contrast events, enrich contextual data, and understand the implications and repercussions of certain historical events not so extensively discussed in the interviews.

The investigation of these dimensions, key events, and actors' positioning shed light on how SAWA was adopted and re-contextualized over time. To reduce biases in interpreting events and processes, I shared my findings with local experts and colleagues to discuss my results. Therefore, the study offers a thorough analytical narrative of the trajectory of SAWA policies in Colombia and contributes to our understanding of how policy ideas and political dynamics can produce large institutional change.

The ‘Bogotá’ laboratory (1998-2001): The local rehearsal of SAWA policies

None of what happened in Colombia since 2002 can be understood without observing the sweeping changes in Bogotá between 1998 and 2001, led by the central figure of Cecilia María Vélez. Appointed Bogotá’s Secretary of Education in 1998, Cecilia Vélez had little experience in education but was involved in decentralization reforms during her time at Colombia’s National Planning Department. A believer in managerial views on public policy, Cecilia Vélez foresaw her appointment as a unique opportunity to exploit the devolved authority from 1991-1994 decentralization reforms. As stated by an interviewee, *“The problem is that real autonomy to govern only comes when you have financial resources”* (GOV12). Therefore, Bogotá’s fiscal resources granted Vélez a political and administrative autonomous educational agenda.

Undoubtedly, Cecilia Vélez embodied the NPM ideas on decentralized governance vested at the 1998 ‘Summit of Santiago’: distance monitoring, academic standards, learning assessments, and performance incentives for schools and teachers (Burki & Perry, 1998). However, Bogotá

lacked an information system with accurate data on enrollment, teachers' allocations, and system-wide performance indicators, indispensable for central steering:

“The reform started in Bogotá. During this time, we had a massive undertaking in transforming information systems for schools and teacher management, new assessments, and charter schools. (...) We needed to align incentives to expand enrollment, so we started using our power to restrict funds to ETCs for teachers' payroll on accurate and reliable student enrollment information” (Interview, GOV4).

Among many policies deployed in Bogotá, three would be scaled up nationally: (i) low-stakes standardized assessments, (ii) learning standards, and (iii) new information systems. Concerned with showing improvement and giving back schools their performance information, Vélez created a new standardized assessment for 3rd, 5th, 7th, and 9th grades (Peña Borrero, 2005). Unlike the national evaluation SABER 11, used to enter higher education, Bogotá's assessment had no consequences for students –or teachers–. The new test fed the performance machinery to inform government decision-making and deploy distance monitoring. Assessment subjects included math, literacy, and science and took place every year. However, as Colombia eliminated its national curricula in 1994, Bogotá's secretariat used its political autonomy to develop academic standards to be monitored by the assessments.

As every policy impacted teachers, Bogotá's reform success hinged on overcoming FECODE's opposition. Thus, Vélez looked into small pockets to dilute FECODE's veto power. Early on, the Secretary imposed a payday discount in case of a teachers' strike. This decision was challenged in court but endorsed by tribunals arguing it did not affect teachers' right to protest. However, the measure's double penalty of discounting teachers' payday and losing the presentism

bonus seriously weakened teachers' protests. In addition, Vélez gained consensus among numerous principals who benefited from teacher relocations derived from accurate enrollment data provided by information systems. Lastly, Vélez secured strong support from the private sector thanks to her charter school initiative called 'Concession Schools' (see Termes et al., 2017). In contrast, FECODE had little margin to resist changes despite its ideological opposition. First, the Secretariat was entitled to change schools' funding and quality policies. Second, FECODE's dislike of assessments and information systems did not echo public opinion, as changes were portrayed as quality and efficiency-improving policies.

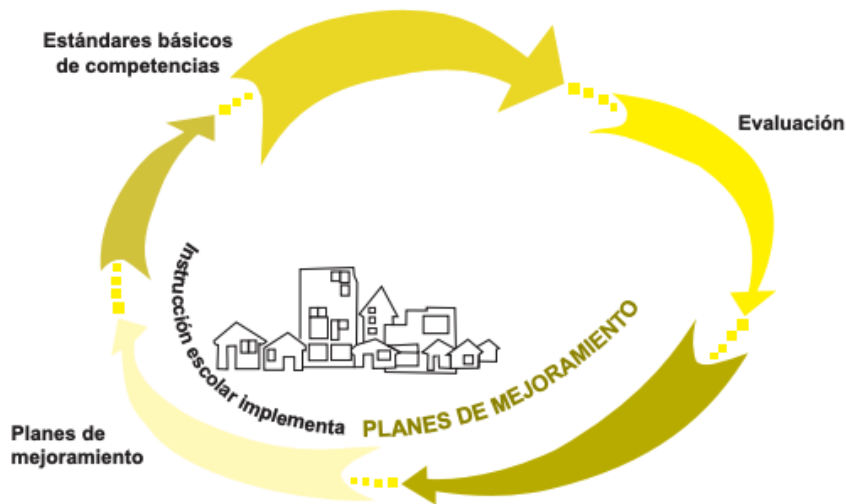
A dominant paradigm: Standards, assessments and performance management (2002-2010)

In 2002, Colombia initiated a new direction in education governance: a progressive pathway towards a vernacular SAWA model under a 'quality assurance' logic. The new orientation pivoted on the government's firm conviction that performance monitoring was indispensable in a decentralized system with extensive teacher autonomy. Appointed head of the NME, Cecilia Vélez proclaimed an "Educational Revolution" in Colombia, focused on three goals: expanding school access, enhancing resource efficiency, and improving education quality (Vélez, 2012). This threefold agenda benefited from the decentralized structure derived from the early 1990s reform but lifted the managerial spirit of the 2001 counter-reform. This new logic would slowly transform school governance.

Under financial distress after the 2001 crisis, the NME faced budget cuts requiring surgical expenditure efficiency to achieve school access and quality goals. However, as occurred in Bogotá, the NME did not have reliable and up-to-date information on school enrollment, teachers'

allocation, or students' performance. According to interviewees, the 1993 Law 60 was sanctioned based on imprecise estimations of enrollment, the number of teachers, or their working conditions. Therefore, in light of the new financing scheme advanced by the 2001 reforms, the NME deployed a new nationwide information system to implement per-capita financing and align ETCs' efforts to expand enrollment and cut unjustified expenses.

The two most relevant governance changes during this period were the new NLSAs and the academic learning standards. Colombia established NLSAs for the 11th grade in 1968, acting since the 1980s as an entry requirement for higher education. Despite its wide coverage, SABER 11th conceded schools or the NME little room for action. Thus, Vélez created a new SABER for grades 5th and 9th with three main goals: (1) producing information about students' outcomes at earlier stages of their school trajectory; (2) allowing the NME and schools to prepare improvement plans, and; (3) to oversee teachers. The government's quality assurance model is represented in Figure 1. Although low stakes, SABER 5 and 9 advanced a new relationship between NME and schools under stricter oversight: *"If teachers have so much autonomy, there's a need for some sort of accountability. We needed to make sure students learn"* (Interview, GOV9).



Note: Top-left corner: Competence-based standards; Bottom-left corner: Improvement plans; Top-right corner: Evaluation; Inside: School instruction implements improvement plans.

Source: Ministerio de Educación (2006, p. 9).

Figure 1: The “Educational Revolution” quality improvement cycle

The new assessments were on core subjects –maths, literacy, and science–, had no direct consequences on students or teachers, and took place every three years. SABER 5 and 9 covered a sample of students in every school –public and private–, although restricting NME’s ability to individualize results at the student or teacher levels. Furthermore, results were not publicly available. The low-stakes and sample-based alternative was argued on the NME’s financial limitations, as monitoring procedures for universal coverage added an overhead cost for a resource-scarce NME. Nonetheless, the exam coverage satisfied the quality assurance rationale, feeding the much-needed performance data for distance monitoring.

In parallel, the new NLSAs required content standards for their design. As posited, the curriculum is controversial in Colombia due to teachers’ autonomy and the decades-long armed conflict (Mantilla-Blanco, 2023). In this context, the NME established high-level competence-based standards, specifying the skills and knowledge that students should demonstrate at the end of each 3-year school cycle. Paradoxically, interviewees mentioned that teachers complained about

standards being too abstract with little instructional guidance from the NME. A second challenge was the notion of competencies, as the Spanish word “competencias” is closely associated with economic competition. Thus, FECODE and left-leaning intellectuals associated this term with the government’s neoliberal agenda rather than with students’ skills. Consequently, FECODE considered the NLSAs and the competence-based standards a government attempt to restrict teachers’ autonomy.

Under Vélez leadership, the NME successfully controlled FECODE’s veto power over the new policies. In the 1990s, FECODE blocked Pastarna’s ‘municipalization’ project, which could have widened regional disparities and atomized teachers’ representativeness (Rodríguez, 2018). However, the late 1990s and early 2000s were a time of internal divisions at FECODE. On the one hand, union leaders were focused on party politics and electoral alliances rather than sector-specific matters (Chambers-Ju, 2017). On the other, FECODE’s unity was wounded after the 2001 teachers’ career reform, dividing ‘new’ teachers into a performance-based pathway, whereas ‘old’ members progress depended on longevity. Ahead of a fragile FECODE, Vélez replicated Bogotá’s strategy of imposing a payday discount for striking teachers (El Tiempo, 2007). Ultimately, the NME succeeded in deploying her policies despite FECODE’s opposition.

Fine-tuning and consolidation: NSLAs at the heart of the new governance (2010-2014)

Under Juan Manuel Santos’ presidency, the educational agenda focused on consolidating the managerial paradigm under a heightened social environment due to peace negotiations with armed groups. Santos’ first minister, Maria Fernanda Campo (2010-2014), shared the pro-managerial style with her predecessor, although with a lower public profile and less ambitious plan. In

continuity with prior policies, Campo also pursued three main goals: reaching universal access to public education, consolidating the information system and assessment culture, and reforming the higher education system.

Early on, the NME attempted a bold reform allowing private universities to become for-profit, producing a power shift among stakeholders in the years to come. The questionable project to reform Law 30 faced fierce opposition from students and unions, fearing the privatization of public universities. Strikes paralyzed the country, spurring discontent in public opinion (El Mundo, 2011). Prioritizing the peace negotiations with the Colombian Revolutionary Armed Forces (FARC in Spanish), President Santos ordered abandoning the reform project: “*There was tremendous social polarization around the peace agreement. Debates within universities were massive. Therefore, once the negotiations with FARC leaked, we had to drop the reform in Congress*” (Interview, GOV7). Campo was weakened after the failed reform, while tensions over the peace agreement would sign Santos’ two-term presidency.

After the crisis, the NME’s agenda pivoted towards improving education quality through strengthening information systems and performance metrics. This shift meant a departure from the prior term –2002 to 2010– when efforts centered on incentivizing ETCs and schools to expand access. Despite enrollment growth in past years, national and international assessments showed low performance and large inequalities across the country (OECD, 2016). Hence, standardized assessments became the guiding policy of the government’s improvement efforts. Campos incorporated SABER in 3rd grade to those in 5th, 9th, and 11th and, from 2012, increased its frequency on an annual basis. Complementary, the NME updated the learning standards in consultation with an expert panel. Certainly, SABER intensification fed the NME vast performance data, although without substantial changes.

In this line, the government implemented a soft quality-improvement strategy resorting on two initiatives: creating reputational incentives for teachers and a national tutoring program called ‘Everybody to Learn’ (PTA as known in Spanish). On the reputational front, the NME aimed to shift attention to test performance through symbolic incentives such as awards for best-performing schools. This idea was epitomized in ‘The Night of the Best,’ a televised event where top-performing ETCs, schools, principals, and teachers received different awards. Then, the so-called ‘PTA’ was an in-service teacher training program through peer tutoring that targeted the worst-performing schools in the country. Designed by McKinsey & Co. and modeled after Mina Gerais’ experience, the PTA used a ‘non-official’ curriculum based on SABER content areas and provided schools with additional pedagogical resources. After initial resistance from FECODE, criticizing PTA interference with school autonomy, the program was well received by schools that benefited from tutoring and the resources provided. Although PTA’s impact on performance is mixed, the program continues until today (Ministerio de Educación Nacional, 2022). In sum, NLSA’s centrality spilled over NME’s timid improvement efforts, which, nonetheless, did not yield significant governance or performance alterations.

“The harder you push, the harder it resists”: A reform pendulum (2014-2018)

During the second term of President Santos, Colombia deepened its managerial approach by incorporating new competition mechanisms typical of market-oriented reforms. Claiming ‘education quality’ as its flagship, Minister Gina Parody’s (2014-2016) agenda centered on standardization, assessments, economic incentives, and performance monitoring. After three presidential terms, schooling access showed major improvements and information systems were robust and reliable (OECD, 2016). However, education quality remained the country’s biggest

challenge: “*Our view was that the system had too many incentives to expand access, yet almost none on quality*” (Interview, GOV18).

The quality agenda was at the heart of Colombia’s national government, not only aspiring to make Colombia “the most educated in Latin America by 2025” but also including it among its development pillars towards 2018 (Ministerio de Educación Nacional, 2016). At the time, Colombia was undergoing an OECD membership assessment, the leading proponent of SAWA policies. Remarkably, NME officials thinking was closely aligned with OECD’s educational views: “*We used OECD reports from 2012 and 2013, as well as drafts version of their 2016 Review of Education Policies as a Bible... For us, OECD’s recommendations were critical*” (Interview, GOV15). Hence, incorporating new instruments and adjusting the settings of existing ones, the NME advanced standardization and competition as reform drivers:

“We aimed to create a system with clear standards grounded in competition among teachers, schools, and ETCs as a driver of improvement while making information and rankings openly available. We truly believed in creating and aligning incentives for improvement” (Interview, GOV18).

Under the new logic, SABER raised as the government’s yardstick for measuring success or failure. The NME deployed a set of financial rewards to incentivize improvement, including bonus payments for teachers or additional resources for top-performing schools (Baxter & León Cadavid, 2021). The new scheme was epitomized in the ‘Synthetic Index of Educational Quality’ (ISCE in Spanish), borrowed from Mina Gerais’s successful experience in Brazil (De Zubiría, 2016). The ISCE combined SABER results with other performance metrics, such as students’ absenteeism and survival rates, producing a standardized and comparable measure across all

Colombian schools. Seeking more fine-grained performance indicators, the government worked towards individualizing SABER results, reaching every Colombian student for the first –and only– time in 2017. Despite the government not creating school rankings to avoid ‘naming and shaming’ practices, results were publicly available and used by different media outlets to judge and compare schools.

To leverage ISCE’s impact, the NME established ‘the Day for Educational Excellence to raise public awareness of the importance of education quality and government efforts. During this day, schools received a individualized performance report, including an annual minimum improvement target (Ministerio de Educación Nacional, 2016). Known as ‘Day E’, the Minister aimed to incentivize results ownership by schools’ principals and teachers, dedicating time for exchange and debates on how to use them to develop an improvement plan for reaching next year’s target.

Learning standards are one of the main ingredients of SAWA policies, central to performance measuring practices and to regulate school autonomy. In its educational standardization crusade, the NME unfolded a two-step plan. First, in 2015 the government published a simplified version of learning standards called ‘Basic Learning Rights’ (DBA in Spanish). If learning standards were criticized for their high-level abstraction for teaching practices, DBA’s were attacked for portraying a narrower expression of school curricula. The second and more ambitious step was creating a national curriculum, seen as a missing piece for improving quality within an unequally resourced school system (see OECD, 2016, p. 153). Interestingly, NME officials, academics, and union leaders shared this diagnosis. A FECODE leader argued on school autonomy: *“I believe that we should not go to extremes in this regard. Autonomy, undoubtedly, requires a better-trained teacher workforce”* (Interview, UNI3).

Similarly, a top-ranked NME official claimed that “*school autonomy has been harmful to our schools... If you have Singapore’s professors, they can be autonomous, but Colombia’s professors do not have that level. They need guidance*” (Interview, GOV22). The NME worked on this project with different universities for almost three years (Ministerio de Educación Nacional, 2016). Still, with little support from other stakeholders, the initiative did not prosper.

In turbulent times due to the peace agreement negotiations, Minister Parody’s confrontational style ahead of veto players eroded stakeholders’ support for their reforms. On the one hand, SABER alterations happened without a broader consensus among the actors involved, such as FECODE or the Minister of Finance. The increased periodicity and coverage required further financing in an indebted sector, yielding tensions with the Finance Minister. On the other hand, the NME team debated whether further to increase exam stakes, such as school sanctions: “*We had tremendous disputes about it, yet those against higher stakes prevailed. We wanted to avoid their negative consequences, such as cheating*” (Interview, GOV21). As in the past, FECODE was not involved in SABER reforms. However, as reforms sidestep direct negative consequences, the union did not show significant opposition to the new agenda.

Table 1: Evolution of ‘SABER’ assessment over the years

Period	Assessment Years	Grades	Subjects	Frequency	Coverage
Uribe (I & II) 2002-2010	2002/03; 2006; 2009	5th & 9th	Math, Literacy & Science	Tri-annual	Every school/ sample of students
Santos (I) 2010-2014	2012; 2013; 2014	3rd; 5th & 9th	Math, Literacy, Science & Civic	Bi-annual / Annual	Every school/ sample of students
Santos (II) 2014-2018	2015; 2016; 2017;	3rd; 5th & 9th	Math, Literacy, Science & Civic	Annual	Every student
Iván Duque 2018-2022	2019; 2021; 2022	3rd; 5th; 7th & 9th	Math, Literacy, Natural Sciences; Environmental Education & Civics	Annual**	Sample-based

** Note: SABER 11th is not included as this assessment has maintained its characteristics since its creation in 1968: it is an annual examination for every student at the end of secondary schooling, required for entering public or private universities.*

*** In 2018, SABER was suspended, and tests underwent substantive changes in design. During the Pandemic, the test was also postponed and reinstalled in 2021.*

In 2015, FECODE led a massive and long-lasting national strike, protesting for salary raises, reopening promotion reviews –frozen since 2011– and a new teacher appraisal system. Parody was unwilling to negotiate, but she underestimated FECODE’s strength. After 12 years under the dual career regime, most of FECODE’s members were affected by the 1278 statute regulations (Bayona-Rodríguez & Urrego-Reyes, 2019). Hence, time produced an institutional drift, where the public opinion echoed FECODE’s complaints. Once again, fearful of losing FECODE’s support ahead of the peace agreement referendum, President Santos settled the conflict by conceding to teachers’ demands. In 2016, during the peak of social turmoil around the referendum, the NME ‘sex education’ protocol was ferociously attacked by former President Uribe and its right-wing coalition “Unidad Democrática.” Uribe, opposed to Santos’ peace agreement, mobilized protests

led by catholic families accusing the NME of sexual indoctrination (Semana, 2016). Unfairly accused but with few allies due to past confrontations, Parody resigned in 2016. After her departure, most of the pro-competition policies were dismantled.

Retrospectively, the lack of buy-in from government agents and FECODE condemned the sustainability of the proposed policy changes. In their own words: “*We should have been aware that wider consensus was necessary to make reforms last over time*” (Interview, GOV17). More so, Parody's team overlooked the timing and sequencing of changes. In contrast to Cecilia Vélez success in deploying systemic reforms, they failed: “*When you push hard, the system push backs harder (...). Cecilia Vélez reform was slower, it was an 8-year long effort... whereas we tried to do it all in two*” (Interview, GOV15). In the following years under Iván Duque (2018-2022), most market-based adjustments were abandoned without altering the already institutionalized assessment practices. For instance, economic rewards for top-performing schools or SABER universal coverage were withdrawn. Nonetheless, the ISCEs and the annual frequency continued,⁷ and SABER expanded to include 7th grade. The endurance of these policies seems to confirm the institutionalization of Colombia’s SAWA model under a quality assurance logic.

Discussion: critical junctures and institutional displacement

Since the 2000s, the decades-long ideological dispute over Colombia’s education has been dominated by proponents of managerialism. Undoubtedly, the early 1990s reforms were a defining moment in this historical dispute. The short but intense cycle between 1991 and 1994 sedimented

⁷ To be precise, SABER 3, 5 & 9 was cancelled in 2018 due to fiscal cuts and in 2020 because of the pandemic. Since 2021, the exam resumed its annual occurrence with a pilot test for grade 7th, fully incorporated into the exam calendar from 2022 until today.

the ideas of two competing policy paradigms embodied by two historical processes: (i) the Pedagogical Movement's strive for school autonomy and; (ii) a decentralized government endorsed simultaneously by a neoliberal strive to reduce the state role and those in favor of reducing Bogotá's authority (Azóca & Giraldo, 1994; Suárez, 2023). Overall, the consolidation of a Social State of Law in the context of decentralization reforms was seen as a triumph of democratic values over managerial or neoliberal views (Miñana Blasco, 2010; Rodríguez, 2018; Rodríguez Guarín, 2023)

Looking back, the 1991-1994 reforms entailed a critical juncture in Colombia's education system that shaped the years to come (Pierson & Skocpol, 2002; Stefes, 2019). The adopted normative framework consolidated governance decentralization and school autonomy, dismissing alternative pathways like further municipalization or the national curriculum (Edwards et al., 2023). In this context, the NME ended up assuming a financing and quality monitoring role, as schools' administration fell under ETCs and curricula on principals and teachers. Consequently, school autonomy and decentralization rendered lock-in within Colombia's education system, proving hard to reverse while restricting future policies (Mahoney & Thelen, 2010).

As argued, the 2000s saw a paradigm shift from social and democratic ideals of the 1990s toward managerial and outcomes-based governance. Coincident with Hall's (1993) conceptualization, the managerial paradigm involved new policy goals of educational standardization, distance monitoring, and outcomes-based governance, as well as new instruments, namely learning standards, NLSAs, and performance targets (Gunter et al., 2016; Wilkins et al., 2024). However, the study demonstrated that the paradigm shift occurred through subtle but significant institutional changes (Mahoney & Thelen, 2010). In other words, the 40-year continuity of autonomy and decentralization policies was not static. Fraught in dispute, institutional

arrangements are sustained by the ongoing power-distributional tensions among actors who constantly update their struggles and alliances (Hacker et al., 2015; Mahoney & Thelen, 2010). In 2001, a change in the power equilibrium between political elites and FECODE enabled a ‘neoliberal’ counter-reform (Rodríguez Guarín, 2023). Although 2001 changes appeared less radical than those of the 1990s, the counter-reform layered approach produced significant transformations in education financing and teachers’ careers against a weakened opposition. The counter-reform updated the historical dispute among policy paradigms in Colombia and set the ground for upcoming shifts.

Unquestionably, the complexities behind ETCs and schools’ autonomy have signed education governance in Colombia. Despite ETCs’ broad political and administrative autonomy to manage schools, their capacity is diluted by their scarce financial independence and limited human resources (Edwards et al., 2023). Hence, paradoxically, in a decentralized system, most ETCs follow the educational policies dictated at the national level. Similarly, while school autonomy was aimed at professionalizing teachers, policymakers, academics, and union members have been critical of teachers’ ability and training for exercising such devolved responsibility (Rodríguez, 2018). Illustrative of this has been the school Institutional Education Projects adopted in 1994, which became a bureaucratic requirement rather than a vivid instrument of professional determination (Cobo, 2022; Wasserman, 2021) Hence, the learning standards and NLSAs have gradually limited school autonomy.

In this context, Colombia’s adoption of SAWA policies lies in political and educational elites’ belief in managerial practices as the best mechanism to regain political, economic, and administrative control over a decentralized school system. This re-centralizing paradox has precedents in countries with highly autonomous environments like England or Australia (Wilkins

et al., 2024). Nonetheless, Colombia's mix of learning standards, low-stakes assessments, and performance management seems more closely aligned with decentralized northern European or Latin American countries following the quality assurance rationale (Parcerisa et al., 2022). In Colombia, the low-stakes SAWA model allowed for crafting a coherent quality improvement narrative, diminishing veto players' ability to oppose them. However, as discussed next, the low-stakes SAWA model has not been static.

Colombia's SAWA: a slowly cooked educational reform

Examining the puzzling trajectory of SAWA policies in Colombia brought the vivid discussion of institutional change to the center stage. In this debate, we argued and showed that incremental variations can result in significant transformations over time and stressed power as a constitutive institutional dimension critical to understanding continuity and change (Béland, 2007; Hacker et al., 2015). In between agency-oriented and structural determinants, the Colombian case has highlighted the importance of key agents, often personifying policy ideas and defying political resistance to drive change (Mahoney & Thelen, 2010). These nuances are exemplified in Cecilia Vélez, a central personage in Colombia's reforms. Starting at Bogotá's Secretariat, Vélez exploited the political autonomy granted during the 1990s reform to unfold its threefold agenda grounded in standardization, assessments, and performance monitoring. As head of NME, Vélez benefited from a new political environment with a strong government and a weak union to enact Law 115 quality oversight mandate, and the per-capita financing requirement in Law 715 (Miñana Blasco, 2010; Vélez, 2012). Therefore, combining a layered adoption of new policies and contextual changes (Hacker et al., 2015), Vélez scaled up Bogotá's policies – including the new NLSAs for grades 5 and 9, academic standards, and information systems.

To impose a managerialist agenda, Vélez combined a sound improvement narrative with a firm stance against a weak union and resourceless ETCs (Baxter & León Cadavid, 2021). Criticizing system-wide inequities, the NME argued that learning standards and information systems would result in an ‘educational revolution’ for improving resource efficiency and educational equity in the country. At the time, FECODE faced a significant divide derived from the new teachers’ statute: its members were separate into two different professional pathways and organized around contradictory incentives, longevity versus performance (Bayona-Rodríguez & Urrego-Reyes, 2019). Ahead of this, applying a controversial interpretation of teachers’ rights to protest, Vélez discounted striking teachers’ payday, diminishing demonstrations and, consequently, FECODE’s veto power (Chambers-Ju, 2017). Against these strategies, FECODE nor ETCs were capable of effectively opposing NME’s initiatives.

Under the same ideological frame, the 2010-2014 term marked the consolidation of the managerial paradigm following minor policy alterations (Hall, 1993). NLSAs gained further centrality as an accountability instrument for the NME, evidenced by its increased frequency – annual since 2012–, incorporation of 3rd grade, and subordination of other policies, such as the PTA program (Baxter & León Cadavid, 2021). During Parody’s tenure (2014-2016), the ‘quality’ agenda attempted a marketized version of managerial practices by incorporating further competitive dynamics and higher stakes within the accountability system (Díaz Ríos, 2020). The NME created the ISCE, published performance results, relaunched a simplified version of learning standards (DBA), and created economic rewards for best-performing schools. These changes strived for a performance-based accountability logic aligned with the global tenets of NPM (Gunter et al., 2016). However, several disputes from the NME with veto players hindered the sustainability of most performance-based reforms. On the material front, individualized

assessments and economic incentives represented an unsustainable financial burden for a sector already indebted and facing unions' salary demands, abandoned after 2017. On the political side, Parody's defeat against FECODE and lack of broader buy-in among critical stakeholders erode the sustainability of the competitive logic.

Table 2: SAWA instruments trajectory and policy changes in Colombia (2002-2018)

Period	Policy goals (ordered)	Instruments	Settings	Veto players	Approach to Institutional change	Policy change
Uribe (I & II) 2002-2010	Access (1), efficiency (2), and quality (3)	NLSAs, Information systems, learning standards*	Unchanged	Weak opposition	Layered and conversion	Third order
Santos (I) 2010-2014	Access (1), quality (2), and efficiency (3)	Teacher training in lowest performing schools (PTA)	Altered	Weak opposition	Layered	First order
Santos (II) 2014-2018	Quality (1), efficiency (2), and access (3)	Synthetic Index of Education Quality, publication of ISCE's results, economic rewards, revised learning standards	Altered	Strong opposition	Layered and drift	Second order

* Sustained in all three periods

Concluding thoughts: A vernacular expression of A global reform

As with every reform, the gap between its deal version and its local expression rises from the encounter of global forces and domestic traits (Steiner-Khamsi et al., 2024). More importantly, the timing at which reforms can be introduced is crucial in determining what policy alternatives are pursued (Porto de Oliveira & Osorio Gonnet, 2023). In Colombia, the installment of SAWA policies and its shift towards a managerial education paradigm was possible thanks to the 2001

counter-reform. At the time, the SAWA model resonated with the governments' will to regain control after decentralization reforms and echoed a regional trend where NLSAs and accountability policies were on the rise (Burki & Perry, 1998; Moreno Salto, 2023; Rivas, 2021). Once in place, the SAWA model became increasingly institutionalized, consolidating an orientation based on standards, assessments, and performance monitoring as governance pillars of Colombia's school system.

Interestingly, as pointed out by Díaz Ríos (2020), SAWA evolution did not render a linear trajectory from a low-stakes stage towards a test-based accountability model with increased choice and higher stakes like in Chile. For instance, the early 2000s reforms on per-capita financing were driven by funding concerns and not to create competition, whereas school choice in public schools remains highly restricted (Díaz Ríos, 2020). Similarly, Consecion Schools –Colombia's charter schools equivalent– are mostly circumscribed to Bogotá despite the government's attempt to expand nationally (Termes et al., 2017).

However, two aspects require further scrutiny. First, building on Díaz Ríos (2020) analysis, this paper makes a step further and contends that Colombia's SAWA reform –grounded in educational standardization, an intensifying assessment culture, and their associated programs– produced significant governance changes for the NME, ETCs, and schools under a managerial logic. Despite market-like mechanisms were abandoned, SAWA tenets and instruments remain central these days. For instance, learning standards overcome criticisms to become an essential guide for schools' daily work (OECD, 2016). Similarly, after an impasse during the pandemic, the ISCE is published regularly, and SABER recovered its annual frequency since 2021, covering grades 3rd, 5th, 7th, 9th, and 11th (ICFES, 2024). These policies not only constrain school

autonomy due to increased standardization but also create reputational consequences for schools, which should not be overestimated (Bunar & Ambrose, 2016; Parcerisa et al., 2022).

Second, the challenges in consolidating a test-based and competitive-driven accountability system were not because of ideational discrepancies among policymakers but material and political constraints at the time of reforms. Colombia's education Ministers endorsed schools' managerial autonomy in hiring practices, more competitive incentives, and stricter consequences for low-performing schools. However, these contentious reforms required larger institutional changes, likely to face significant opposition from veto players, and increased government expenditure in a deficitary system (Díaz Ríos, 2020). Nonetheless, SAWA policies, particularly accountability ones, have proven resilient to political and economic turmoil, offering valuable payoffs to government elites ahead of a decentralized and atomized system.

All in all, these nuances signal the vivid ideological and political forces in constant dispute behind policy reforms underlying the observed continuity and their subtle changes. The Colombian case has shown how incremental alterations can yield significant governance transformations. Policy paradigms are crucial templates guiding policy change, although we often find them merged in complex ideational and policy mixes with competing views (Capano & Howlett, 2020). After 2018, NLSAs and performance monitoring remained central to educational policies in Colombia, including under the most recent left-wing government of Gustavo Petro (2022-2026). The new government brought new winds of change to the country, which has become an OECD member since 2021. In what ways will the OECD membership affect SAWA policies in Colombia? What changes will Petro's government advance into SAWA policies? The decades-long ideological dispute over Colombia's education is far from over, rendering SAWA policies' future an intriguing case.

Funding acknowledgment: This project received funding from the Institute of Latin American Studies at Columbia University and the Thinker Foundation.

References

- Azóca, J. C. T., & Giraldo, H. D. (1994). El proceso de descentralización educativa en Colombia. *Revista Colombiana de Educación*, 29, Article 29.
<https://doi.org/10.17227/01203916.5364>
- Bartlett, L., & Vavrus, F. (2016). *Rethinking Case Study Research*. Routledge.
- Baxter, J. G., & León Cadavid, M. C. (2021). From Political Intentions to Structural Interventions A Review of Two Decades of Education Policy Reforms in Colombia. In A. Rivas (Ed.), *Examining Educational Policy in Latin America* (pp. 96–108). Routledge.
- Bayona-Rodríguez, H., & Urrego-Reyes, L. A. (2019). 240 años de profesión docente en Colombia. *Educación y Ciudad*, 2(37), Article 37.
<https://doi.org/10.36737/01230425.v2.n37.2019.2143>
- Béland, D. (2007). Ideas and Institutional Change in Social Security: Conversion, Layering, and Policy Drift. *Social Science Quarterly*, 88(1), 20–38.
- Bevir, M. (2009). *Key Concepts in Governance*. SAGE Publications Ltd.
<https://doi.org/10.4135/9781446214817>
- Bocanegra Acosta, H. (2010). Las políticas educativas y el magisterio colombiano en la década de los 80. *Diálogos de saberes*, 32, Article 32.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Bunar, N., & Ambrose, A. (2016). Schools, choice and reputation: Local school markets and the distribution of symbolic capital in segregated cities. *Research in Comparative and International Education*, 11(1), 34–51. <https://doi.org/10.1177/1745499916631064>
- Burki, S. J., & Perry, G. E. (1998). *Más Allá del Consenso de Washington: La Hora de la*

- Reforma Institucional*. The World Bank. <https://doi.org/10.1596/0-8213-4282-7>
- Capano, G., & Howlett, M. (2020). The Knowns and Unknowns of Policy Instrument Analysis: Policy Tools and the Current Research Agenda on Policy Mixes. *SAGE Open*, *10*(1), 215824401990056. <https://doi.org/10.1177/2158244019900568>
- Cardenas, S. (2009). *School Based Management in Latin America*.
<https://doi.org/10.13140/2.1.1985.3443>
- Chambers-Ju, C. (2017). *Protest or Politics? Varieties of Teacher Representation in Latin America* [UC Berkeley]. <https://escholarship.org/uc/item/4rc371rp>
- Cobo, O. M. (2022). El currículo oficial en las dos últimas reformas educativas en Colombia. *Revista Educación, Política y Sociedad*, *7*(1), Article 1.
<https://doi.org/10.15366/rep2022.7.1.001>
- Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (Sixth edition.). SAGE.
- Cruz, A. L. A. (2012). La influencia del Banco Mundial en las políticas educativas de Colombia. *Pedagogía y Saberes*, *36*, Article 36. <https://doi.org/10.17227/01212494.36pys69.79>
- De Zubiría, J. (2016, April 4). El Día E y la calidad de la educación. *Semana*.
<https://www.semana.com/educacion/articulo/la-calidad-en-la-educacion/468070/>
- Díaz Ríos, C. M. (2020). The Role of Policy Legacies in the Alternative Trajectories of Test-Based Accountability. *Comparative Education Review*, *64*(4), 619–641.
<https://doi.org/10.1086/710767>
- Edwards, D. B., Moschetti, M. C., Grillet, P., Quilabert, E., & Caravaca, A. (2023). *Descentralización y privatización educativa en Colombia: Facilitadores y tensiones*. Education International.

- El Mundo. (2011, November 17). Los universitarios colombianos suspenden la huelga pero seguirán con las protestas. *El Mundo*.
<https://www.elmundo.es/america/2011/11/17/colombia/1321535739.html>
- El Tiempo. (2007, May 24). Fecode dice que seguirá el paro. *El Tiempo*.
<https://www.eltiempo.com/archivo/documento/MAM-2504861>
- FECODE. (1991, December). Propuestas de FECODE sobre Ley General de Educación. *Educación y Cultura*, 25, 5–8.
- Frostenson, M. (2015). Three forms of professional autonomy: De-professionalisation of teachers in a new light. *Nordic Journal of Studies in Educational Policy*, 2015(2), 28464.
<https://doi.org/10.3402/nstep.v1.28464>
- Ganimian, A. J. (2016). Why do some school-based management reforms survive while others are reversed? The cases of Honduras and Guatemala. *International Journal of Educational Development*, 47, 33–46. <https://doi.org/10.1016/j.ijedudev.2015.12.001>
- Giraldo-Paredes, H., & De La Cruz-Giraldo, G. N. (2016). La influencia neoliberal en las políticas educativas en Colombia. *Criterio Libre Jurídico*, 13(2), Article 2.
<https://doi.org/10.18041/crilibjur.2016.v13n2.26208>
- Gunter, H. M., Grimaldi, E., Hall, D., & Serpieri, R. (Eds.). (2016). *New Public Management and the Reform of Education: European lessons for policy and practice*. Routledge.
<https://doi.org/10.4324/9781315735245>
- Hacker, J. S., Pierson, P., & Thelen, K. (2015). Drift and conversion: Hidden faces of institutional change. In J. Mahoney & K. Thelen (Eds.), *Advances in Comparative-Historical Analysis* (pp. 180–208). Cambridge University Press.
<https://doi.org/10.1017/CBO9781316273104.008>

- Hall, P. A. (1993). Policy Paradigms, Social Learning, and the State: The Case of Economic Policymaking in Britain. *Comparative Politics*, 25(3), 275–296.
<https://doi.org/10.2307/422246>
- Hogan, J., & Howlett, M. (Eds.). (2015). *Policy Paradigms in Theory and Practice*. Palgrave Macmillan UK. <https://doi.org/10.1057/9781137434043>
- Högberg, B., & Lindgren, J. (2021). Outcome-based accountability regimes in OECD countries: A global policy model? *Comparative Education*, 57(3), 301–321.
<https://doi.org/10.1080/03050068.2020.1849614>
- ICFES. (2024). *Vuelven las Pruebas Saber 3°, 5°, 7° y 9°—Edición 2023*.
https://www.icfes.gov.co/inicio/-/asset_publisher/qcrs/content/vuelven-las-pruebas-saber-3%C2%B0-5%C2%B0-7%C2%B0-y-9%C2%B0-edici%C3%B3n-2023-
- Klijn, E.-H. (2012). New Public Management and Governance: A Comparison. In D. Levi-Faur (Ed.), *The Oxford Handbook of Governance* (p. 0). Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199560530.013.0014>
- Le Galès, P. (2011). Policy Instruments and Governance. In M. Bevir (Ed.), *The SAGE Handbook of Governance* (pp. 142–159). SAGE Publications Ltd.
<https://doi.org/10.4135/9781446200964>
- Le Galès, P. (2016). Performance measurement as a policy instrument. *Policy Studies*, 37(6), 508–520. <https://doi.org/10.1080/01442872.2016.1213803>
- Lennert da Silva, A. L. (2022). Comparing teacher autonomy in different models of educational governance. *Nordic Journal of Studies in Educational Policy*, 8(2), 103–118.
<https://doi.org/10.1080/20020317.2021.1965372>
- Lundström, U. (2015). Teacher autonomy in the era of New Public Management. *Nordic Journal*

- of Studies in Educational Policy*, 2015(2), 28144. <https://doi.org/10.3402/nstep.v1.28144>
- Mahoney, J., & Thelen, K. (2010). A Theory of Gradual Institutional Change. In J. Mahoney & K. Thelen (Eds.), *Explaining Institutional Change: Ambiguity, Agency, and Power* (pp. 1–37). Cambridge University Press. <https://doi.org/10.1017/CBO9780511806414.003>
- Mantilla-Blanco, P. L. (2023). “We Think We’re Far from Conflict, but That’s Not True”: Peace Building and Remembrance through Memory Sites in Colombia. *Comparative Education Review*, 67(1), 78–99. <https://doi.org/10.1086/722801>
- Mentini, L., & Levatino, A. (2023). A “three-legged model”: (De)constructing school autonomy, accountability, and innovation in the Italian National Evaluation System. *European Educational Research Journal*, 14749041221148280. <https://doi.org/10.1177/14749041221148280>
- Miñana Blasco, C. (2010). Políticas neoliberales y neoinstitucionales en un marco constitucional adverso. Reformas educativas en Colombia 1991-2010. *Propuesta Educativa*, 34, 37–52.
- Ministerio de Educación Nacional. (2016). *Colombia, la más educada en 2025. Informe de gestión 2014-2016*. Ministerio de Educación Nacional. https://redes.colombiaaprende.edu.co/ntg/men/archivos/Informe_de_gestion_2014_2016.pdf
- Ministerio de Educación (2006). *Estándares Básicos de Competencias en Lenguaje, Matemáticas, Ciencias y Ciudadanas. Guía sobre lo que los estudiantes deben saber y saber hacer con lo que aprenden*. Bogotá: Ministerio de Educación Nacional. Available at: https://www.mineducacion.gov.co/1621/articles-340021_recurso_1.pdf
- Moreno Salto, I. (2023). *Harmony and Cacophony in Large-scale Assessments in Education: A Mexican Account*. Brill. <https://brill.com/display/title/69042>

- OECD. (2016). *Education in Colombia*. OECD Publishing. <https://www.oecd-ilibrary.org/docserver/9789264250604-en.pdf?expires=1663001166&id=id&accname=ocid177456&checksum=199C36D991FF309DE8425BE098B830B9>
- Pagès, M. (2021). Enacting performance-based accountability in a Southern European school system: Between administrative and market logics. *Educational Assessment, Evaluation and Accountability*, 33(3), 535–561. <https://doi.org/10.1007/s11092-021-09359-7>
- Parcerisa, L., Verger, A., Martín, M. P., & Browes, N. (2022). Teacher autonomy in the age of performance-based accountability: A review based on teaching profession regulatory models (2017-2020). *Education Policy Analysis Archives*, 30, (100)-(100). <https://doi.org/10.14507/epaa.30.6204>
- Peña Borrero, M. (2005). *Aprendizajes sobre gestión educativa. La experiencia de Bogotá 1998—2003*. Empresarios por la Educación.
- Pierson, P., & Skocpol, T. (2002). Historical Institutionalism in Contemporary Political Science. In I. Katznelson & H. V. Milner (Eds.), *Political science: The state of the discipline* (pp. 693–721). W. W. Norton & Co. <http://www.gbv.de/dms/sub-hamburg/347252621.pdf>
- Porto de Oliveira, O., & Osorio Gonnet, C. (2023). Comparative Transfer Analysis: Policy Instruments, Space and Time. *Journal of Comparative Policy Analysis: Research and Practice*, 25(4), 367–384. <https://doi.org/10.1080/13876988.2023.2235308>
- Pulido Chaves, O. O. (2014). Veinte años de la Ley 115 de 1994. *Educación y Ciudad*, 27, 15–26.
- Rivas, A. (Ed.). (2021). *Examining Educational Policy in Latin America: Comprehensive Insights into Contemporary Reform*. Routledge.

- Rodríguez, A. (2018). Fecode, Movimiento Pedagógico, Seretaría de Educación de Bogotá, educación rural, profesionalización docente, Ley General de Educación, Plan Decenal de Educación, Expedición Pedagógica Nacional. In J. D. Herrera & H. Bayona-Rodríguez (Eds.), *21 Voces: Historias de vida sobre 40 anos de educacion en Colombia* (pp. 21–46). Ediciones Uniandes.
- Rodriguez Guarín, L. A. (2023). *La contrarreforma educativa neoliberal en Colombia*. LETRAME.
- Semana. (2016, October 4). *Gina Parody renuncia como ministra de Educación*.
<https://www.semana.com/educacion/articulo/renuncia-de-gina-parody/497734/>
- Stefes, C. H. (2019). Historical Institutionalism and Societal Transformations. In W. Merkel, R. Kollmorgen, & H.-J. Wagener (Eds.), *The Handbook of Political, Social, and Economic Transformation* (pp. 95–105). Oxford University Press.
<https://doi.org/10.1093/oso/9780198829911.003.0009>
- Steiner-Khamsi, G. (2016). Comparing the Receptions and Translations of Global Education Policy, Understanding the Logic of Educational Systems. In T. D. Jules (Ed.), *The Global Educational Policy Environment in the Fouth Industrial Revolution* (Vol. 26, pp. 35–57). Emerald Group Publishing Limited.
- Steiner-Khamsi, G., Jóhannesdóttir, K., & Magnúsdóttir, B. R. (2024). The school-autonomy-with-accountability reform in Iceland: Looking back and making sense. *Nordic Journal of Studies in Educational Policy*, 0(0), 1–19.
<https://doi.org/10.1080/20020317.2024.2375087>
- Suárez (Ed.). (2023). Este libro y sus razones. In *40 años del Movimiento Pedagógico 1982-2022. Los nuevos desafíos de la educación en Colombia hoy* (pp. 9–16). Magisterio

Editorial.

Termes, A., Verger, A., & Bonal, X. (2017). Mitos y asunciones de las escuelas chárter: Un análisis de los colegios en concesión de Bogotá. *Educação & Sociedade*, 38(141), 911–934. <https://doi.org/10.1590/ES0101-73302017173097>

Vélez, C. M. (2012). *La gestión de la educación en Colombia 2002—2010*.

Verger, A., Fontdevila, C., & Parcerisa, L. (2019a). Constructing School Autonomy with Accountability as a Global Policy Model: A Focus on OECD's Governance Mechanisms. In C. Ydesen (Ed.), *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex* (pp. 219–243). Springer International Publishing. https://doi.org/10.1007/978-3-030-33799-5_11

Verger, A., Fontdevila, C., & Parcerisa, L. (2019b). Reforming governance through policy instruments: How and to what extent standards, tests and accountability in education spread worldwide. *Discourse: Studies in the Cultural Politics of Education*, 40(2), 1–23. <https://doi.org/10.1080/01596306.2019.1569882>

Voisin, A., & Dumay, X. (2020). How do educational systems regulate the teaching profession and teachers' work? A typological approach to institutional foundations and models of regulation. *Teaching and Teacher Education*, 96, 103144. <https://doi.org/10.1016/j.tate.2020.103144>

Wasserman, M. (2021). *La educación en Colombia*. Debate.

Wilkins, A., Collet, J., Esper, T., Grimaldi, E., & Gobby, B. (2024). Assembling New Public Management: Actors, networks and projects. In D. B. E. Jr, A. Verger, K. Takayama, & M. McKenzie (Eds.), *Researching Global Education Policy: Diverse Approaches to Policy Movement*.

https://www.academia.edu/78197910/Reassembling_New_Public_Management_Actors_networks_and_project

4.2. Article 2: Not everything learned is transferred: Unravelling knowledge use and lesson-drawing behind school autonomy with accountability reforms in Colombia

Summary

Journal: Globalisation, Societies and Education

Article Status: Accepted with minor revisions, on November 8th, 2024. Open Access.

Article's Table of Content

Introduction

Lesson-drawing and evidence use in global education governance

Colombia: case and context

Education reforms and the OECD membership

Research design

Network analysis of policy documents

Qualitative analysis: documents and interviews

Unpacking evidence use in policy planning

The emergence of IOs in knowledge networks

International references as a legitimising strategy

The politics of evidence: What matters for lesson-drawing

OECD's rise in Colombia's education

Brokers wisdom behind policy transfer: A tailored advice

Context matters: The limits of lesson-drawing

Discussion and conclusion

References

Not everything learned is transferred: Unravelling knowledge use and lesson-drawing behind school autonomy with accountability reforms in Colombia

Abstract

In times of evidence-based regulation, policymakers are expected to make decisions based on existing knowledge to address policy problems. In search of policy alternatives, governments often utilise international evidence to borrow lessons, often resulting in the transfer of foreign policies. However, the linkages between knowledge use, lesson drawing, and policy transfer are messier than assumed, shaped by knowledge brokers' interests, politicians' ideological preferences, and institutional constraints. This article unpacks how and when evidence use and lesson-drawing leads to policy transfer by exploring Colombia's adoption of school autonomy with accountability reforms. Following a sequential mixed-methods design combining network and content analysis of policy documents, and interviews with key informants, the study yields three important findings. First, Colombia policy documents represent a 'staging' use of knowledge to comply with international standards of evidence-based policymaking, not reflecting adopted solutions. Second, that the OECD has emerged as a key policy broker in Colombia since the country's accession process since 2011. This is important because of the OECD's global influence through knowledge brokerage and main promoter of autonomy and accountability policies in education. Finally, policy transfer was limited to knowledge and lessons compatible with the local political and institutional landscapes in the country.

Keywords: lesson-drawing; education reform; OECD; policy brokerage; evidence use

Introduction

In an ideal world, when governments are confronted with policy problems, solutions would emerge from scientifically-driven scrutiny of existing evidence on best –and worse– practices, portrayed as relevant lessons to learn from. In the era of evidence-based regulation, this mythical view of the problem identification and alternative-seeking process assumes that policymakers ground their choices on already-tested solutions (Lingard, 2013; Parkhurst, 2017). Nevertheless, this romantic conception of policymaking overlooks that evidence use and lesson-drawing are more nuanced phenomena influenced by decision-makers ideological preferences, interest groups' pressure, or institutional boundaries upon policymakers' choices (Nutley et al., 2007; Pawson, 2006; Weyland, 2007). Frequently, decision-makers cherry-pick pieces of evidence aligned with their preferences (Dunlop, 2014), or simply refer to international examples as a facade to advance their predilected choice (Steiner-Khamsi, 2021).

It is well established that lesson-drawing or policy learning drives policy transfer of education reforms, yet the linkages between these phenomena are not always intelligible (Porto de Oliveira & Osorio Gonnet, 2023). Researchers, think tanks, and international organisations (IOs) produce evidence from policy experiences, becoming lessons decision-makers can learn to solve a given problem. However, for learning to happen, evidence on a given matter requires interpretation, complexity reduction, and translation into policymakers' language – a role uptake by knowledge brokers (Bali et al., 2019; Sriprakash & Mukhopadhyay, 2015). Therefore, scrutinising knowledge brokers' agendas and strategies is crucial for understanding the multifaceted relationship between evidence use and lesson-drawing.

In times of evidence-based governance, the OECD and the World Bank have become crucial knowledge brokers in education (Niemann & Martens, 2018; Zapp, 2017). Due to the

OECD and World Bank's dominance in education, scholars have sought to unpack their brokerage strategies used to influence national policies among countries (Baek, 2022; Cardoso, 2020; Verger et al., 2019b; Ydesen et al., 2023). Certainly, knowledge mobilisation and policy brokerage lie at the core of IOs' global governance mechanisms, fostering cross-national comparisons and policy advice to drive education reforms (Seitzer et al., 2023; Sorensen et al., 2021). However, rather than passively following IOs' recommendations, local decision-makers may use global influences and discourses –evidence, standards, policies– to advance their interest within their domestic boundaries (Steiner-Khamsi, 2021). Hence, the way in which evidence use may lead to policy reforms is far from a linear process.

In particular, studies have explored the relationship between evidence use and policy reforms in contexts where scientific commissions or expert panels are pivotal in advising governments (McDonnell & Weatherford, 2013; Schrefler, 2010). Through bibliometric analysis of education reform documents in Norway, Steiner-Khamsi and colleagues (2019) concluded that most expert knowledge in background papers is lost in political translation into policy documents. Similarly, Baek (2022) found that the United States, Korean, and Norwegian policymakers selectively refer to or borrow international knowledge depending on available local expertise, how contentious a given reform is, or the legitimacy of cited authors. However, less is known about knowledge use in education policymaking in countries under significant influence of IOs or where domestic politics outweigh scientific clout in the policy process. Given that dozens of countries follow these patterns, we ought to account for whether similar trends occur or, if not, how and to what extent knowledge shapes policy in such cases.

To address this gap, Colombia is a suitable case to unravel the linkages between evidence use, policy brokerage, and lesson-drawing behind education reform. Two features make the

Colombian case particularly appealing. First, Colombia has a 60-year-long tradition of policy planning led by the National Planning Department, defined as an evidence-based organisation (OECD, 2014). The Planning Department leads the development and publication of multiple policy documents, where the National Development Plan stands out for outlining the national government's policies for its four-year period (Barón, 2013; Zapata-Cortés, 2020). Second, after a 7-year evaluation process (2011-2018), Colombia became an OECD member in 2020 (OECD, 2021). While IOs' membership is a sensitive moment for policy transfer due to signing into transnational agreements (Carroll, 2014), the OECD does not impose policies in the education sector but rather seeks to influence countries' decisions through knowledge mobilisation and policy advice (Centeno, 2021; Sorensen et al., 2021). Therefore, as Colombia was the first OECD country to undergo an education sector review as part of the membership process, it makes it a relevant case to explore how the process influenced lesson-drawing and policy transfer.

In this context, this study focuses on knowledge use, policy brokers and lesson-drawing, shaping Colombia's adoption of school autonomy with accountability (SAWA) policies. Since the 2000s, SAWA policies have become the OECD's model for governance reform worldwide (Högberg & Lindgren, 2021; Verger & Fontdevila, 2025). SAWA reforms advance a shift towards managerial practices and outcomes-based governance in education through standardised testing and competitive improvement incentives, often causing resistance among relevant stakeholders. Interestingly, policymakers often rely on international best practices as a legitimising strategy when pursuing controversial reforms like SAWA, even when pursued policies differ from those referenced (Takayama, 2010). Therefore, the study uses SAWA policies as a case for exploring three research questions: (i) How and whose evidence was used in shaping Colombia's education policy between 2002 and 2022? (ii) To what extent has lesson-drawing shaped policy transfer in

Colombia's SAWA policies? (iii) In what way does evidence use lead to policy learning as a driver of policy transfer?

To answer these questions, I adopted a sequential mixed-methods design combining network analysis of references in policy documents (n=25), followed by thematic analysis of the documents and interviews (n=29) with decision-makers and key informants in Colombia. First, analysing documents' citations and content helps understand the relationship between evidence use and lesson-drawing behind SAWA policies. Then, interviews unpack the network and document findings by laying out the role of policy brokers in lesson drawing.

The analysis reveals two contrasting uses of knowledge in policy reforms. At first glance, the increasing number of citations over the years confirms Colombia's alignment with the global turn towards evidence-based policymaking. However, after a deeper dive, Colombia's policymakers seem to employ references to conform with international standards, having little engagement with the cited knowledge. I call this a 'staging' performance, as there is a significant decoupling between the evidence used and the lessons drawn by policymakers when adopting policy alternatives. In contrast, the OECD's centrality found in policy documents was corroborated by decision-makers' interviews. Colombia not only relied heavily on OECD-produced knowledge but also the OECD gave concrete and tailored advice to Colombia officials on SAWA policies, playing a pivotal role in its adjustments and calibrations over time. Illustrated by Colombia's national assessments and school curricula reforms, this study unpacks the importance of lessons' contextual similarities for policy learning to drive global policy transfer.

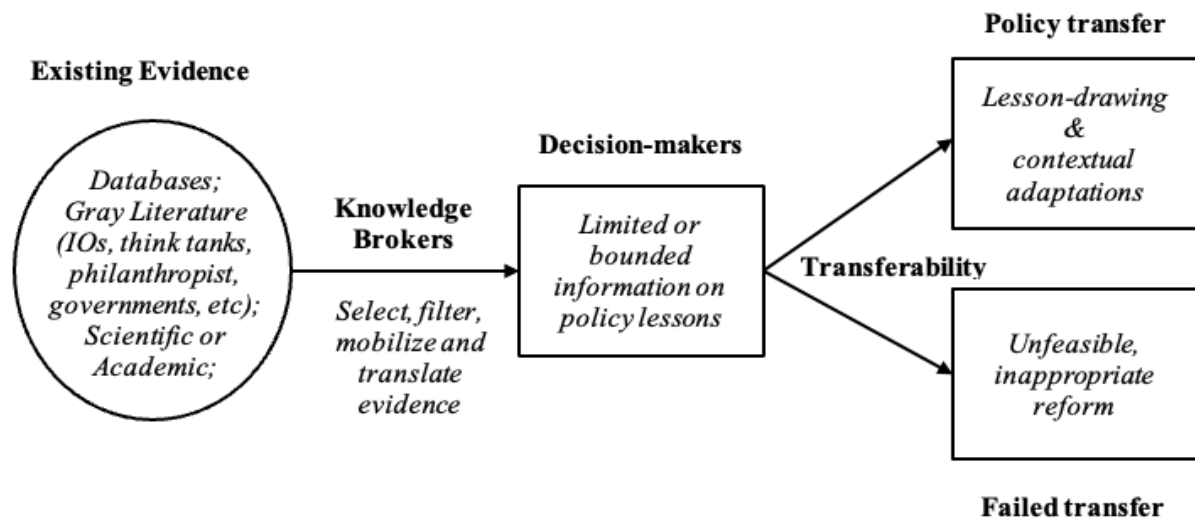
Lesson-drawing and evidence use in global education governance

In policy transfer and diffusion literature, learning is a crucial mechanism underpinning policy adoption in national or cross-national settings (Dolowitz & Marsh, 2000). However, learning challenges its operationalisation and demarcation from other mechanisms like emulation (Meseguer, 2004; Volden et al., 2008). Broadly speaking, learning is understood as “an improved understanding of causal relationships in the light of experience” (Meseguer, 2004, p. 71). From a rational-choice perspective, scholars interpreted learning as a cognitive update from decision-makers searching for policy alternatives (Meseguer, 2004; Shipan & Volden, 2008). This view is contested by the notion of bounded learning, which stresses that agents’ options are limited by cognitive shortcuts and selective pieces of information, often mobilised by knowledge brokers or epistemic communities (Haas, 1992; Seitzer et al., 2023; Weyland, 2007).

In this study, I use learning as a synonym for lesson-drawing, as “lessons constitute what is learned” (Rose, 1991, p. 7). In other words, lesson-drawing results from agents’ observations, assessments, and causal inferences about different experiences on similar problems and judgment about lessons’ transferability into their contexts (Rose, 1991; Steiner-Khamsi, 2021). Accordingly, lesson-drawing does not always lead to policy transfer, nor does it solely enter the scene during the alternative-seeking stage. For instance, when framing a policy problem, politicians reference ‘best practices’ or international examples to gain legitimacy vis-a-vis domestic stakeholders, a practice known as externalisation (Baek, 2022; Santos & Centeno, 2023; Takayama, 2010). More often than not, actors first choose their preferred solution and mobilise evidence, discourses, and resources to create a policy problem (Béland & Howlett, 2016). Nonetheless, policy transfer or borrowing involves an additional step to lesson-drawing. Adopting an idea, a policy instrument, or a whole reform package requires adequate domestic conditions, namely enough support and an

opportunity window for adoption, combined with its recalibrations to accommodate the country’s political and institutional context (Cowen, 2009; Díaz Ríos, 2024).

As argued, knowledge brokerage connects the available information, data, and evidence into consumable ‘best practices,’ ‘what works,’ or even lessons to avoid (Bali et al., 2019; Verger et al., 2019b). Scientists and decision-makers belong to different communities, which largely explains the gap between evidence production and its policy use (Steiner-Khamsi et al., 2024). However, in the era of evidence-based regulation, knowledge brokerage surged as a central governance strategy for the OECD and the World Bank, although they adopt divergent approaches to mobilise policy reforms (Centeno, 2021; Zapp, 2017). Whereas the World Bank mostly recommends on-size-fits-all policies (Zapp, 2017), the OECD fosters horizontal learning, advancing context-sensitive advice for their member countries (Seitzer et al., 2023). Moreover, the World Bank is known for imposing conditionalities as a central mechanism for policy transfer (Hossain, 2022). However, lesson-drawing is a voluntary transfer mechanism (Dolowitz & Marsh, 2000; Meseguer, 2005), which requires more subtle forms of influence.



Source: self-elaboration

Figure 1. A streamlined model of evidence use and policy transfer

In this realm, the OECD exercises ‘soft’ governance by combining ideas generation, policy evaluation, and data production as influencing mechanisms (Martens & Jakobi, 2010). Its legitimacy in education derives from knowledge generation and benchmarking, particularly via its notable Program for International Student Assessment (PISA) (Grek, 2009; Sorensen et al., 2021). Hence, beyond data gathering and dissemination, the OECD acts as a policy broker for translating evidence into policy recommendations and guidelines for their member states through diverse instruments, including PISA, the Review of National Education Policy, and the Education Policy Outlook, among others (Seitzer et al., 2023).

In recent years, the OECD has advanced SAWA policies as its preferred education governance model. Rooted in new public management principles, SAWA fosters devolved school authority over curricula, budgeting, teaching practices, and staffing, mixed with centralised performance-based accountability instruments like standardised assessments, inspections, or learning standards (Verger et al., 2019a). This policy mix succeeded globally largely thanks to its flexibility to fulfil different purposes. In England and Chile, SAWA policies advanced market-based reforms with high-stakes accountability systems (Parcerisa & Falabella, 2017). In contrast, in Nordic European or Latin American countries, assessments, monitoring, and performance metrics respond to a quality assurance rationale (Díaz Ríos, 2024; Högberg & Lindgren, 2021). Under the surface, SAWA configures diverse policy arrangements with contextual specificities across the most diverse countries.

Colombia: Case and Context

Colombia’s policy-planning tradition of consolidation stems from 1968 when the Constitutional Reform established the legal and administrative bases for its first National Development Plan

(NDP). However, it was not until the 1990s decentralisation reforms that the NDP became a central instrument in Colombia's government policy. At that time, decentralisation responded to international pressures and power redistribution efforts as decision-making was concentrated in Bogota (Edwards et al., 2023). The 1991 Constitution and Law #152 from 1994 consolidated the administrative decentralised organisation, where the NDP became a mandatory instrument for each administration requiring Congressional approval as a national law with its corresponding financial resources and implementation monitoring mechanisms.

In theory, decentralisation reforms promoted a bottom-up planning process through different local planning bodies, but, in reality, the new scheme consolidated the prior centralised top-down logic (Zapata-Cortés, 2020). The elaboration of the NDP is a complex process merging the government's political agenda, expert bodies' knowledge, input from community-based representatives, and certain policy guidelines and targets for their adoption at sub-national levels (Zapata-Cortés, 2020). In addition, the NDP includes requirements from international agreements, such as the SDGs, and prioritised areas highlighted by the other government planning bodies, such as the Council for Social and Economic Policies (CONPES in Spanish). Nonetheless, the NDP materialises most of the government's campaign promises, resulting in a larger weight of political priorities at the expense of experts' views or community members' input (Barón, 2013).

In education, the Ministry leads the creation of two relevant planning documents: the Education Sector Plan (ESP) and the 10-year Education Development Plan. The ESP follows and expands on the policies established by the NDP for each government period, following the political agenda. In contrast, the 10-year plans, created by the 1994 General Education Law (#115), undergo a year-long stakeholder consultation process, including input from expert panels ("Mision de Sabios"), providing general guidelines and strategic priorities for the future. The expectation is

that both NDP and ESPs should be aligned with 10-year education plans, although coordination mechanisms are not clearly outlined.

In addition, the CONPES and the Minister of Education publish different documents that inform policy decisions or refer to specific guidelines outlined in national education plans (see Annex I for a complete list of documents). In sum, despite the blurred linkages between document types, planning documents set the government's policies, while other publications aim to inform or guide sectoral development. Although both types of documents contribute to policy discourse formation in education, not all are given the same consideration in the policy process.

Education reforms and the OECD membership

In the last two decades, Colombia has embraced managerial and accountability policies in education to regain control over its highly decentralised school system, configuring a quality assurance model (Díaz Ríos, 2024; Author, 2024). In Colombia, administrative functions over schools fall under sub-national units called Territorial Certified Entities; principals and teachers have significant pedagogical autonomy, but national policy guidelines, quality evaluation, teachers' salaries and certification, and funding flows remain centrally controlled (Edwards et al., 2023; Author, 2024). In this context, three consecutive national administrations followed managerial principles implementing different accountability instruments to monitor schools at a distance: Álvaro Uribe (2002-2010), Juan Manuel Santos (2010-2018), and Iván Duque (2018-2022), shown in Table 1.

Concomitantly, this 20-year period witnessed Colombia's strengthening ties with the OECD. Having participated in PISA since 2006, the OECD membership signalled Colombia's aspiration to become a relevant player among developed countries. After two years of negotiations,

Colombia started its formal accession process in 2013 (Vargas-Alzate, 2021), being the first to face an education sector review to become a new member (OECD, 2013, 2021). During this time, an OECD commission visited Colombia multiple times and met with numerous stakeholders. The result was the committee’s favourable recommendation on Colombia’s admissions and the publishing of the Review of National Education Policies (OECD, 2016), containing policy recommendations for the education sector. Hence, in light of these strengthening links, a close examination of OECD’s influence mechanisms can illuminate recent education reforms in Colombia.

Table 1: SAWA instruments trajectory and policy changes in Colombia (2002-2018)

Reform Period	SAWA elements	Main Events	Main policy documents
2002-2010: The education revolution	National Large-Scale Assessments (SABER) on maths, literacy (2006) and civics (2009) for grades 5th and 9th, every 3-years. Learning standards linked to ‘SABER’ test.	2006: Colombia’s first participation in PISA 2009: First peace agreement with guerrilla groups	National Development Plans (2) 2002-2006 & 2006-2010 Education Sector Plans (2) 2002-2006 & 2006-2010
<i>President: Alvaro Uribe</i>	Education information systems (EMIS) for per-pupil funding and performance monitoring Development of school improvement plans based on SABER results New teacher assessment system including annual performance reviews and for promotion;		Ten-year Education Plan 2006-2016
2010-2018: Colombia, the most educated by 2025	Increased frequency and coverage of LSLAs program SABER (2012; 2014; 2015; 2016; 2017; 2018) Centralized monitoring with performance incentives: Synthetic Education Quality Index (2015-2018)	2013: Start of formal OECD’s accession process 2014: Large teachers’ unions strike 2016: Education Ministry	National Development Plans (2) 2010-2014 & 2014-2018 Education Sector Plans (2) 2002-2006 & 2006-2010
<i>President:</i>			

<i>Juan Manuel Santos</i>	Revised Learning Standards (“DBA”)	steps down after ‘sex-education’ scandal	Ten-year Education Plan 2016-2026
	Consolidation of the teacher’s evaluation system	2018: Signing of peace agreement with armed groups	
2018-2022: A pact for Colombia, a pact for equity	Changes to assessment policy during COVID-19 pandemic: “Assess to Progress”	2020: Accession to OECD 2020-2021: Covid-19 crisis	National Development Plans 2018-2022
<i>President: Iván Duque</i>			Education Sector Plans 2018-2022

Source: self-elaborated

Research design

In this study, I used a sequential mixed-methods design (Creswell & Clark, 2018), combining (i) network analysis of policy documents’ citations, (ii) documents’ content analysis, and (iii) networked-cued interviews with key informants (Baek, 2022). The rationale for combining quantitative and qualitative methods lies in the study’s twofold objectives: to unpack whose and what evidence was utilised in Colombia’s education sector and to examine whether evidence use and lesson-drawing led to policy transfer.

The bibliometric analysis of references sheds light on descriptive features such as the number, frequency, type of citations, and other attributes of references used in policy documents. Complementarily, the network lens incorporates tools to analyse the resulting knowledge grids formed by these references to explore the meaning of actors’ centrality, their ties, or changes over time (Pizmony-Levy & Baek, 2022). However, to account for lesson-drawing, simply analysing references would find two important limitations. First, citations alone say little about what policies, experiences or how knowledge is used. Second, in a country like Colombia, where policy planning is highly dependent on political decisions, the content of documents would not suffice to know

what lessons were drawn. Therefore, interviews with decision-makers delve deeper into network findings, providing a more comprehensive picture of the implications of knowledge use for policy reforms.

Network analysis of policy documents

To understand what and whose knowledge is used in education reforms, I retrieved and analysed 25 policy documents published by Colombia's government between 2002 and 2020. As explained, Colombia follows a hierarchical and centralised policy planning cycle. Therefore, to create my sample, I started by retrieving every national policy-planning document for the 2002-2022 cycle, including NDPs (n=5) and ESPs (n=5). Second, I included publications from the Ministry of Education, such as administration's *memoirs* (n=7) or CONPES education-related publications (n=5). Finally, I included every long-term planning document (n=3) published between 2002 and 2020. Examples of these are "*Vision Colombia 2019: A proposal*" (document #3) or the 10-year education sector plans (see Annex I).

Following that, I coded each document's references into a database totalling 1233 citations, as detailed in Table 2. I analysed the frequency of citations and coded different attributes for every reference, including (i) year of publication, (ii) publisher or institutional affiliation of the author, (iii) location of publication, or if authors or publishers were IOs. Subsequently, I used UCINET 6.289 to analyse the database and NetDraw 2.097 (Borgatti et al., 2009) to create the 2-mode networks of (a) policy documents and references and (b) documents and authors. Then, I converted each two-mode network into a one-mode network, forming a balanced matrix with equal rows and columns containing the identification numbers of all references and authors, respectively. Citation relationship – when policy document X cites reference Y – was coded 1, and no citation

relationship was coded 0. Then, I used these two matrixes to analyse the structure of knowledge networks in Colombia (Karseth et al., 2025).

While most bibliometric analyses of policy documents focus on cited publications (Pizmony-Levy & Baek, 2022; Verger et al., 2019b), I also delve into cited authors. Countries' policy documents often cite multiple publications with scarce overlap (Steiner-Khamsi et al., 2019). In contrast, if an author is cited across multiple documents, this repetition signals it as a legitimate voice for constructing an argument. Then, to analyse authors' relevance in knowledge networks, I examined the number of citations an author has received (in Table 3) and calculated their normalised degree centrality measure in the one-mode networks. Degree centrality captures how many times an author is co-cited with other authors. Then, the normalised measure divides the degree centrality of each author by the maximum possible degree in the network, allowing for comparisons across networks of different sizes or from different periods (Prell, 2012). Finally, distances between nodes and their relative size in the network can be interpreted to analyse actors' relevance (Prell, 2012).

Qualitative analysis: documents and interviews

The study's second phase began with a content analysis of policy documents, exploring what and how evidence was used when discussing SAWA policies. Given the controversies around SAWA reforms, I aimed to understand how evidence was used, who was cited as an authoritative source, and what lessons were drawn. I coded every policy document, identifying each mention of SAWA-related instruments with a unique ID, such as learning standards, national or international assessments, performance incentives, etc. Then I coded the policy name, why it was mentioned –

i.e., problem framing; alternative seeking; lesson-drawing–, its origin, meaning to what country it referred to, and tone used – positive, neutral, negative –.

Finally, results from the network and content analysis of policy documents informed my in-depth interviews (n=29) with key informants involved in adopting or adjusting SAWA policies during the analysed period (2002-2022). Interviews took place between June and July 2023. I used purposeful sampling criteria based on interviewees’ roles to identify decision-makers and other key informants.⁸ The final sample comprised top-ranked officials from the Ministry of Education (n=19) labelled as “MOE”, local education secretariats (n=3), the National Planning Department (n=3) labelled as “NPD”, and project leaders or consultants from IOs working in the country (n=4) labelled as “IO”.⁹ Following Baek (2022), I developed a network-cued protocol focusing on three key areas: (i) the rationale behind SAWA instruments adoption and the participants’ role; (ii) the role of domestic or international evidence in the policy process; (iii) other experiences or best practices from countries or IOs used as a reference for policy decisions, and; (iv) the influence of OECD membership in SAWA policies.

During the interviews, I shared the main network findings on centrality, types of publications, and experiences mentioned in policy documents while asking informants to share their thoughts and reactions. Furthermore, I asked specifically what experiences they considered when deciding on policy instruments, how they used existing evidence, from whom they received advice, and what role IOs played in each case. This way, informants could make sense of network findings and expand on the role of evidence, lesson-drawing, and knowledge brokers in

⁸ The project was approved by the Institutional Review Board (IRB) of Teachers College, Columbia University (IRB #23-327).

⁹ Participants’ identity is masked using general tags to classify their roles. However, during the initial contact and before starting each interview, I explained the risks of participation in this study due to the reduced number of people occupying their positions. This meant that their identity could be recognized beyond my actions to prevent their disclosure, as discussed in my IRB protocol.

Colombia's reform. Following a thematic content analysis, I analysed documents and interviews, centred on how and what evidence was used, which lessons were borrowed, and what factors shaped policy transfer (Braun & Clarke, 2006).

The following section introduces the general trends regarding the evidence used in Colombia's policy documents. Then, I present the documents' content analysis to unpack how and for what purposes evidence was used in the context of SAWA policies. Finally, I examine and expand these findings with interview data on the linkages between evidence use and lesson drawing behind Colombia's SAWA reform.

Unpacking evidence use in policy documents

In line with the global shift towards evidence-based regulation, citations in Colombia's policy documents have grown considerably over the years. This trend, shown in Table 2, is particularly noticeable in Colombia's main planning documents, the NDPs and ESPs. For instance, the 2002-2006 NDP uses only six references, the NDP 2010-2014 has 26, and the latest NDP from 2018-2022 has 114.¹⁰ The relative weight of citations also changed: neither of the four NDPs or ESPs published between 2002-2010 included a reference list, while both from the latest period (2018-2022) had a substantive list, including several footnotes.

Upon further investigation, Colombia used evidence from considerably different sources in each document while drawing mostly on grey literature and domestic publications. The analysis reveals a wide range of citations, with just 54 publications —4.37% of all references— cited more than once. Concerning the type of evidence used, only 6.3 % (n=78) of the references are peer-

¹⁰ In the case of NDPs, I only retrieved citations from the education sub-sections.

reviewed publications, and almost half (43,7%) are government-published documents, including statistical reports, databases, or prior planning documents.

Table 2. Reform waves and the number of analysed documents and references

Reform period	Analysed documents (total)	Number of references	Average Citations per-Page
2002-2010: The education revolution	11	552	0.36
2010-2018: Colombia, the most educated by 2025	10	177	0.43
2018-2022: A pact for Colombia, a pact for equity	4	504	1.2
Total	25	1233	0.52

Pivoting to cited authors in Colombia’s documents, I observe a similar high dispersion pattern, except for a reduced group repeated throughout most publications. These authors are government bodies –the Ministry of Education or National Planning Department–, and IOs. As shown in Table 3 and the knowledge networks below, the OECD and the World Bank rank at the top of citations, where their centrality in policy documents also grew over time. Since 2010, IOs have replaced government branches as the most central authors since 2010. In other words, line 4 in Table 3 shows that in the first period (2002-2010), the OECD had quite a low centrality being co-cited with 12% of all possible authors, whereas in the last period (2018-2022), when Colombia has gained its membership, it was at the top being co-cited with 27.6% of all possible authors. This growing trend will be echoed during interviews, signalling the increasing relevance of OECD in Colombia’s education sector.

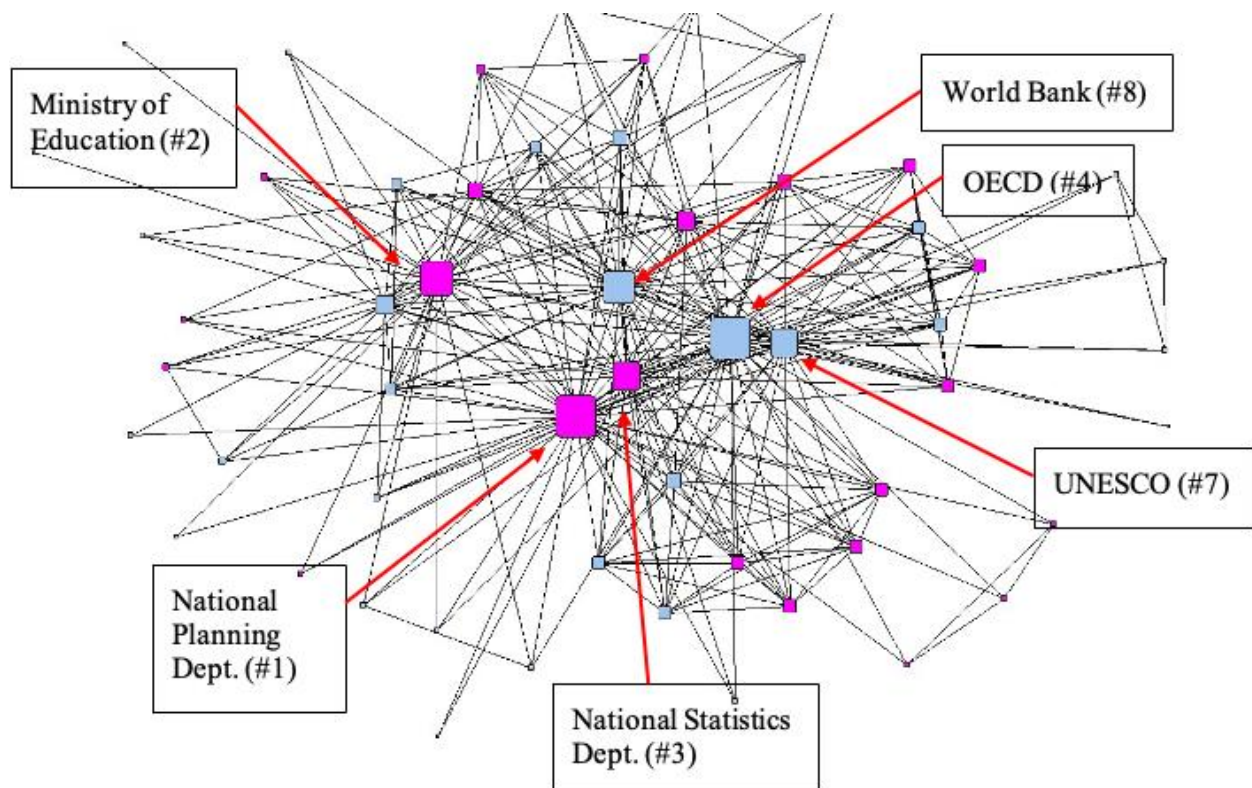
Table 3. Most cited authors in Colombia’s policy documents, order by document appearances

#	Author	Author ID	Reform periods	Appearances in policy documents	Total number of citations	Normalised Degree Centrality Period 1 (2002-2010)	Normalised Degree Centrality Period 2 (2010-2018)	Normalised Degree Centrality Period 3 (2018-2022)
1	National Ministry of Education	320	1, 2 & 3	20	149	0.161	0.181	0.212
2	National Planning Department	129	1, 2 & 3	19	117	0.152	0.2	0.263
3	National Statistics Department	128	1, 2 & 3	15	69	0.139	0.096	0.256
4	OECD	373	1, 2 & 3	14	38	0.12	0.21	0.276
5	Colombian Institute for Evaluating Education Quality (ICFES)	229	1, 2 & 3	11	17	0.136	0.125	0.2
6	UNESCO	485	1, 2 & 3	9	13	0.145	0.054	0.192
7	World Bank	528	1, 2 & 3	8	34	0.138	0.033	0.276
8	ECLAC	76	1, 2 & 3	6	18	0.116	0.085	0.244
9	UNDP	484	1, 2 & 3	5	9	0.123	0.025	0.18
10	United Nations	490	1, 2 & 3	5	20	0.116	0.05	0.276

The emergence of IOs in knowledge networks

The network analysis of Colombia’s documents confirms the bibliometric findings: a reduced group of authors, led by OECD and the World Bank, lie at the knowledge network core. Figure 2 displays the resulting network, where node size represents authors’ centrality and ties indicate co-citations across policy documents. As explained, while government branches rank at the top of citations, their cited evidence is mainly statistical data, planning documents, or self-published reports. These sources are at the bottom of the evidence hierarchy, denoting low engagement with knowledge production (Nutley et al., 2007).

In contrast, references to IOs' publications show a more profound engagement with different authors and reports when dealing with different policy problems. To illustrate this, PISA is recurrently cited when discussing low-quality education, and UNESCO emerges as an authoritative voice when referring to school access problems. Arguably, the fine-tuned use of international references relates to the different agendas pushed by these IOs in the global education space (Cardoso, 2020; Ydesen et al., 2023).

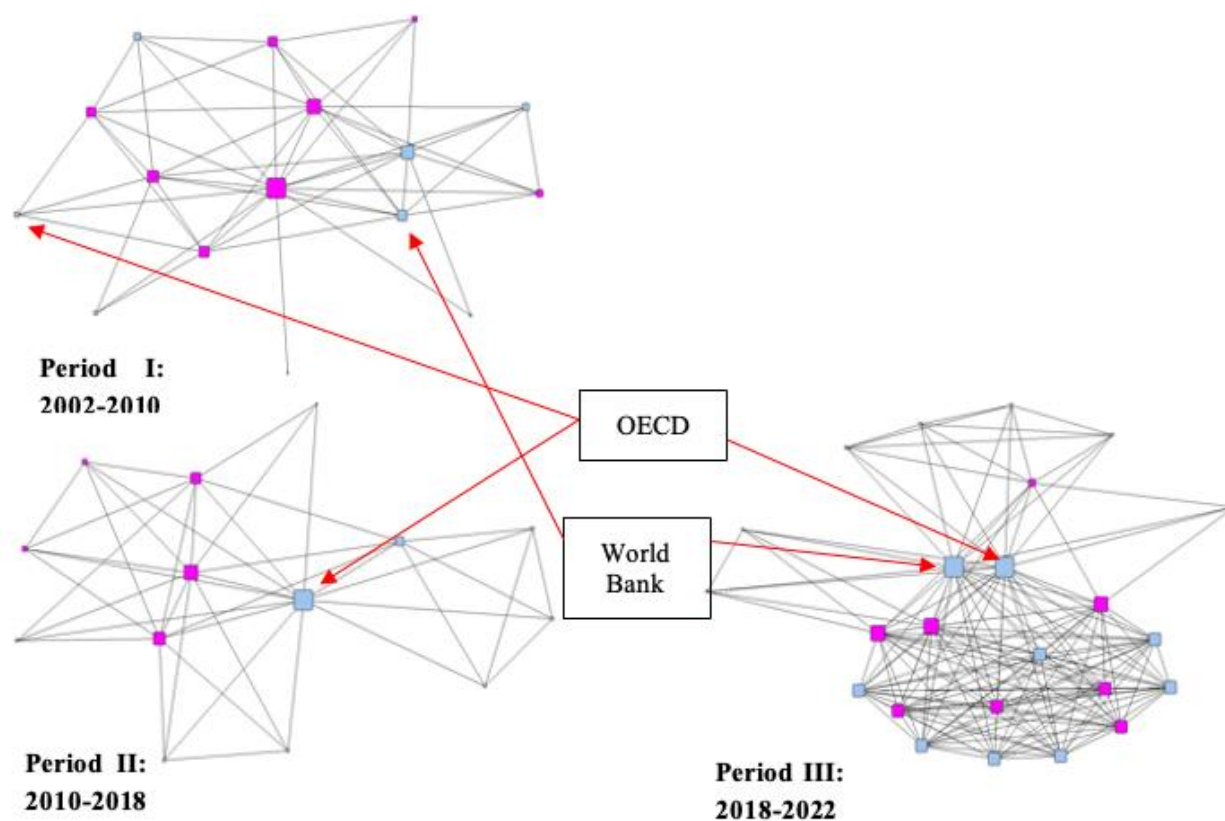


Note: Each node represents an author cited in policy documents. Node size represents authors' centrality. Ties are undirected, with a threshold of 3 citations. Pink nodes are domestic authors, and light blue nodes are international

Figure 2. Network of authors cited in Colombia's education policy documents (2002-2022).

Then, I disaggregated the different authors' networks formed over the three reform periods. On the one hand, we observe an increasing number of cited authors. On the other hand, network changes showcase the OECD's and World Bank's larger centrality. Figure 3 illustrates

how the OECD ‘moves’ from a peripheral network location in the first period (2002-2010) to the network’s core in the second period (2010-2018). In this regard, two considerations are needed. First, the World Bank’s core location in knowledge networks confirms other studies’ findings that show its historical influence in Colombia’s education sector (see Díaz Ríos & Urbano-Canal, 2023). Second, OECD’s centrality confirms the expectation of increasing relevance due to Colombia’s accession negotiations since 2011. Considering this novel finding, the next step involves unpacking how evidence is used, what and whose lessons are drawn, and what role the OECD played in SAWA reforms.



Note: Each node represents an author cited in policy documents. Node size represents authors’ centrality. Ties are undirected, with a threshold of 3 co-citations. Pink nodes are domestic authors and light blue nodes are international authors.

Figure 3: Network of authors cited in Colombia’s policy documents by reform period.

International references as a legitimising strategy

The content analysis of Colombia's policy documents revealed that reference use regarding SAWA policies served almost exclusively for problem-framing purposes rather than alternative seeking or lesson-drawing. This use coincides with prior studies showing that, when framing problems ahead of policy reform, policymakers reference international examples to legitimise their proposed solution (Santos & Centeno, 2023; Steiner-Khamsi, 2021; Takayama, 2010). Similarly, when discussing issues like school autonomy or standardised assessments, documents drawn from international data articulate three problems: (i) students' underperformance, (ii) system inequities, and (iii) poor teaching quality. Then, different SAWA policies are presented as solutions to each problem.

Since its first appearance in 2006, PISA performance has been used to signal Colombia's low education quality and systemic disparities. Arguably, Colombia preferred OECD data to discuss quality issues as it barely referenced other international assessments the country has participated in since the 1990s.¹¹ Interestingly, documents presented a contextualised explanation of Colombia's underperformance by comparing it with other Latin American countries:

“In 2006, 57 countries participated, including six Latin American countries (Argentina, Brazil, Chile, Colombia, Mexico, and Uruguay). The results of these countries were below the international average, and Colombia was at the same level as Argentina and Brazil and below Uruguay and Chile” (Policy document #11, 2010, p. 148).¹²

¹¹ Colombia participates in UNESCO-LLECE Regional and Comparative Study since 1997, the IEA Trends in International Mathematics and Science Study (TIMSS) since 1995 and Trends in International Mathematics and Science Study (PIRLS) since 2001.

¹² Every quote from policy documents was translated from Spanish by the author.

Whereas OECD data served problem-framing purposes on education quality issues, this is not the case when discussing other policies like higher education financing or school access challenges. For instance, citations to academic literature were used to promote the “*Jornada Unica*” initiative, which promotes full-time schooling: “*The literature (Hincapié, 2014; García, Fernández y Weiss, 2013; Cerdan-Infantes y Vermeersch, 2007) indicates that a longer duration of students in educational institutions contributes to quality improvement (...)*”, further supported by its impact in improving PISA performance: “*In fact, a greater number of hours of instruction in mathematics favours student performance according to the PISA 2012 (Avendaño et al., 2015a)*” (Policy Document 17#, 2015, p. 80).

In contrast, Colombia references the OECD to legitimise already adopted policies: “*An example is “Everyone Learns,” a program aimed at transforming education quality, which the OECD has recognised as a world-class pedagogical experience at the Summit of Ministers of Education held in Tokyo 2014*” (Policy document #19, 2014, p. 17). Similarly, regulatory changes in education are backed by the OECD’s advice: “*Following the recommendations of the OECD, within the administrative provisions of the sector, the Single Regulatory Decree 1075 of 2015 was issued, which compiles and orders all the regulations of the educational sector (...)*” (Policy document #24, 2018, p. 266). In sum, these examples demonstrate the emergence of the OECD as an authoritative voice in Colombia’s education sector. While PISA acts as Colombia’s yardstick to measure its efforts in education, the OECD knowledge and recommendations endow policies with external legitimacy in the face of domestic audiences. Therefore, next, I address how PISA and the OECD’s accession process influenced what SAWA lessons had led to policy transfer in Colombia.

The politics of evidence: What matters for lesson-drawing

Turning to the interviews, I find a disconnection between official knowledge sources and the lessons drawn by government officials. Coincidentally with documents and references analysis, the OECD and its knowledge products emerged as an authoritative source for policy decisions. However, three aspects stand out. First, the lessons and examples mentioned by policymakers did not appear in policy documents, signalling a decoupling with official knowledge published Colombia's government. Second, lessons and examples drawn by decision-makers were driven by political relevance and contextual similarity, rather technical motivations. Third, OECD's importance for policymakers changed upon Colombia's formal accession bid in 2011.

Chile is a clear example of a political and contextually relevant reference, whose policies are legitimised by the OECD. Chile sits historically at the top of Latin American countries in PISA rankings and emerged during interviews as Colombia's poster child to compare against and learn from. Every Colombian decision-maker looked at Chile after PISA's results release: *"I always stared at Chile. Then, Mexico and Argentina... but Mexico spent double as much as we did in education, and we obtained the same results!"* (Interview, MOE 9). The 'Chilean obsession' peaked in 2014, when Colombia declared its ambition to become "the most educated country in Latin America by 2025" (Policy Document 19#, 2015). This had a clear meaning for decision-makers: *"It was simple: We wanted to outperform Chile in PISA 2024"* (Interview, MOE 17).

In 2002, Colombia inaugurated the SAWA reform cycle, placing its national large-scale learning assessment (LSLA) program "*SABER*" at the centre of its governance strategy. The revamped LSLA was paired with the newly introduced competency-based learning standards, both aimed at regaining control and monitoring education quality in a highly decentralised and autonomous school system (see Author, 2024). Over the years, *SABER* would have multiple

readjustments increasing its frequency, coverage, and associated consequences. Thus, when searching for lessons on similar policies, Colombia officials turned to their obsession: Chile.

Chile had a long trajectory in national standardised testing and international recognition for its educational improvement strategies targeting the lowest-performing schools (Parcerisa & Falabella, 2017). Therefore, its world-known National System of Educational Assessment, “SIMCE”, was considered the best case to learn from: “*Chile was essential for us... Technicians from SIMCE helped us develop the new assessments, including our sampling strategy and the items’ construction*” (Interview, GOV 4). In 2005, Colombia launched SABER 5 and 9 –for grades 5th and 9th– in addition to the historical SABER 11, used for university accession. The new SABER 5 & 9, with a three-annual occurrence, assessed a sample of students in every school in literacy and mathematics.

However, from all the Chilean lessons, Colombia only borrowed those deemed compatible with their goals and context. To illustrate this, the Chilean SIMCE is the pivotal piece of a marketised education system, where assessment results have high stakes for schools, including access to government funds, reputational effects due to school rankings, and government interventions (Díaz Ríos, 2024; Parcerisa & Falabella, 2017). Although Colombian decision-makers expressed sympathy with Chile’s high-stakes assessment, they discarded them due to foreseeable teacher unions’ resistance and institutional incompatibility: “*Those policies make sense when you have high degrees of school choice like in Chile... but in Colombia, when it comes to state schools, you go to your neighbouring school*” (Interview, MOE 4).

In other words, while Chile’s reputation in PISA was enough to make it a role model, policy transfer was limited to its technical know-how. Colombia officials borrowed SIMCE’s model by transforming the national evaluation institute, “ICFES”, into an independent education quality

agency. However, lesson-drawing rendered selective borrowing by discarding problematic or unaligned policy features with Colombia's institutional legacies leading to a low-stakes SAWA model under quality assurance goals.

OECD's rise in Colombia's education

For Colombia, becoming an OECD member meant global recognition as a liberal-developing country and closing more than 50 years of political violence and recurring economic crises (Padgett & Otis, 2012). The OECD membership was a flagship under Santos' administration (2010-2018), a period signed by the tumultuous and heightened peace agreement with armed groups.¹³ These twofold processes –the OECD accession (2013-2018) and political turmoils derived from the peace negotiations– meaningfully shaped Colombia's education sector and SAWA policies' trajectory.

Colombia's ties with the OECD depart from an early involvement in PISA in 2006 to its membership in 2020. For non-members, joining PISA is voluntary, although it is often perceived as a requirement for belonging to the global community. When asked whether they felt pressure for participating, a top-rank official was clear: *“On the contrary, it was our will to join, not only PISA but also PIRLS and TIMSS”* (Interview, MOE 2).¹⁴ As shown, Colombia drew heavily on PISA data for problem-framing purposes, progressively becoming the yardstick against which it measures its progress. Remarkably, although policymakers expected poor results, the external monitoring of Colombia's education quality efforts was more important: *“I always found*

¹³ The armed conflict in Colombia, which began in the 1960s, has lasted for over five decades, making it one of the longest in Latin America. It stems from deep social inequalities, disputes over land, and political exclusion, involving actors such as guerrilla groups (e.g., FARC, ELN), paramilitaries, drug cartels and the military.

¹⁴ The Progress in International Reading Literacy Study (PIRLS) and Trends in Trends in International Mathematics and Science Study from from the International Association for the Evaluation of Educational Achievement (IEA).

assessments extremely relevant. When you manage such a large budget, the only way to know if you are going the right way is with high-quality assessments” (Interview, MOE 9). Underperformance critiques were not strong enough to weaken Colombia’s participation: *“Every time, after poor results, we have a huge scandal. In 2006, we performed at the bottom, with similar results than Argentina. I was proud of that, but Argentine’s were fuming!”* (Interview, MOE 9).

Moreover, subsequent PISA turmoils created growing dissatisfaction among Latin American participants: *“Some countries were considering stepping down from PISA and want us to join, but that was out of the question for us”* (Interview, MOE 9). Undoubtedly, appearing in PISA rankings symbolised belonging to the international community, a legitimising stamp for Colombia – more relevant than any possible scandal.

Undoubtedly, the OECD membership was a turning point in this story, with great significance for Colombia. During the accession process, the intense policy exchange between Colombian and OECD authorities opened room for lesson-drawing and policy transfer. Whereas accessing countries have to follow strict economic or environmental regulations (Carroll, 2014), the OECD only requires minor adjustments in education without imposing significant sub-sector policies. Thus, the novel requirement of Colombia’s education sector review created opportunities for soft influence mechanisms, mutual dialogue, and lesson-drawing.

Interestingly, when shown network images on OECD centrality and asked about their view on the organisation’s role in Colombia, responses changed based on whether the country had begun its accession process. Before 2011, officials had willingly joined PISA and saw OECD’s standards as an aspiration rather than a pressuring factor. Indeed, Colombia officials view the OECD as a policy forum, embracing the country’s membership as: *“an opportunity to learn best practices from developed countries and also to share yours. We want to learn from the best; that’s our*

country's aspiration" (Interview, NDP 2). However, authorities also noted the different stakes upon belonging to the organisation. For some, it also meant raising the bar: *"It changed people's mindset. If you compare within Latin America, Colombia always performs on average. However, when comparing to OECD countries, you are at the bottom, and you need to catch up"* (Interview, MOE 17).

Between 2013 and 2016, an OECD committee closely scrutinised Colombia's education sector. The review process involved gathering data about schools and students and multiple stakeholder interviews with national, regional, and local officials, NGOs, private sector representatives, and teacher union leaders (OECD, 2016). From the Minister's perspective, it was an insightful and demanding experience. According to interviewees, far from dictating a pre-established course of action, the OECD was concerned with understanding Colombia's education challenges.

During the process, Colombia officials and OECD experts exchanged views, documents, and ideas, underpinning the policies outlined in the 2014-2018 NDP. OECD's suggestions were not just any advice for education officials: *"We used OECD reports from 2012 and 2013, as well as drafts of the OECD's report as a Bible... We used them for policy planning. For us, OECD's recommendations were critical"* (Interview, MOE 15). This narrative contrasts that of authorities prior to the accession process, signalling the importance of institutional ties as a driver of policy influence.

The resulting product was one of OECD's (2016) key instruments, the Review of National Education Policies, with plenty of policy suggestions for Colombia. For OECD staff, recommendations should be tailored and feasible for its members: *"It makes no sense to suggest unfeasible recommendations. We had a high-level vision of what direction would suit Colombia.*

If you check OECD's report recommendations, we tried giving concrete examples while offering an agenda for the next ten years" (Interview, IO 3). This quote illustrates OECD's brokerage strategy, cautious of the local possibilities when giving policy advice (Seitzer et al., 2023).

Brokers' wisdom behind policy transfer: A tailored advice

Wearing its policy broker hat, two OECD recommendations showcase its context-sensitive approach to promoting SAWA reforms: the uses and consequences associated with LSLAs and the importance of having a national curriculum. Undoubtedly, LSLAs are at the core of SAWA reforms. However, whether there are hard or soft consequences associated with them, and the differences in testing coverage, frequency, and policy usage will result in different accountability regimes (Högberg & Lindgren, 2021). In every case, the appropriateness and feasibility of high-stakes or lower-stakes models are highly dependent on the countries' political environment and their compatibility with prior policies (Díaz Ríos, 2024).

During the second term of Santos' presidency (2014-2018), the Ministry of Education had an ambitious goal: making "Colombia the most educated in Latin America in 2025" (Policy Document 19#, 2015). For education officials, the improvement strategy was clear: *"I read tons of OECD papers... and we knew we needed to standardise the system and its procedures. Standardising meant refining and simplifying our learning standards and creating the Synthetic Index of Education Quality centred on SABER results"* (Interview, MOE 18). The Index was a composite measure indicating the overall school performance, where LSLAs' results had the largest weight (Baxter & León Cadavid, 2021). Its goal was to *"bring education quality at the centre of public debate"* (Interview, MOE 19), also releasing schools' results.

For the Index's annual publication, the standardisation plan required increased SABER's frequency, coverage, and use. Adopting an annual occurrence since 2012, in 2015 SABER shifted from sample to census coverage while linking school performance with financial incentives. These changes spurred heightened debates within the Minister's team, discussing the trade-offs between the new LSLAs features and their potential side-effects: *"We debated over economic incentives tied to SABER results... I lost that fight, because in that way we made the SABER high stakes, risking pervasive incentives... but we needed to show results quickly"* (Interview, MOE 20).

However, OECD officials warned Colombia about further negative consequences of high stakes, such as narrowing curricula and teaching to the test:

"Using a single test for many purposes is risky: it might not adequately capture what needs to be measured and can lead to distortions such as teaching to the test, which can undermine and invalidate the whole evaluation system (OECD, 2013b). Mexico's recent experience with its national external assessment provides a cautionary tale on the risks of attaching high stakes to formative assessments (Martínez, 2015)." (2016, p. 183).

A second debate centred on SABER coverage, whether it should remain a student sample within all schools or reach every student for more precise performance control over the years. Individualised student assessment data would enable attaching direct consequences, a concern among certain MOE members. In contrast, a second group aimed to provide as much information as possible to schools to foster data ownership, individualised student support and school improvement strategies. In 2017, SABER was sat by every student in grades 3, 5, and 9, enabling their traceability over time. However, the OECD's warning about the risks of high-stakes testing

prevented officials from attaching direct consequences to students' performance. Shortly after, due to technical and financial constraints, the individualised coverage was withdrawn in 2018.

Context matters: The limits of lesson-drawing

The second and perhaps most important OECD recommendation was to create a national curriculum for Colombia's school system. In its SAWA model, the curriculum is pivotal to standardised learning practices and for building assessment instruments (Verger & Fontdevila, 2025). During the education sector review, the OECD (2016) found significant disparities in teaching strategies, contents, and resources among Colombian schools. These differences partly responded to Colombian teachers' vast autonomy and uneven training, combined with extensive school resource inequalities (OECD, 2016; Radinger et al., 2018). For Colombia officials, school autonomy was problematic for education quality: "*When schools have high competencies like in Singapur, high autonomy works well. Colombia is quite unequal in that regard*" (Interview, MOE 18).

Hence, in a diverse and unequally resourced education system, OECD staff concluded that the absence of such an important roadmap as a national curriculum was harmful to schools and students, encouraging Colombia to embark on its development:

"The curriculum is the cornerstone of a coherent approach to evaluation and assessment (OECD, 2013b). In the absence of detailed national learning objectives, the ICFES has developed its own testing standards to externally assess the performance of schools. (...) The development of a national curriculum could help Colombia to ensure that its students acquire the full range of knowledge, skills, and values that all citizens hold to be important" (OECD, 2016, p. 157).

To make its recommendation evidence-based, the OECD looked beyond regional cases and drew lessons from two countries with similar teaching autonomy and decentralised governance: New Zealand and the United States. First, the OECD argued that among all its members, only New Zealand combines a similar degree of school autonomy without a national curriculum (see OECD, 2016, p. 153). Nonetheless, New Zealand compensates this mix with national guidance, which “sets the direction for student learning (i.e. vision, values, key competencies, learning areas, achievement objectives and principles) and provides guidance to schools as they design and review their own curricula” (OECD, 2016, p. 153). Furthermore, the OECD (2016) sustains its support to national curricula even in highly autonomous contexts as: “the existence of a national curriculum does not mean there is no room for local adaptation” (p.153), as occurs in other decentralised countries like Argentina or Mexico.

After making a case for a national curriculum, the OECD proposed the United States’ common learning standards as a suitable example for Colombia. Introduced in 2010, the “Common Core” provided national guidelines on core subjects while maintaining state and school autonomy to tailor their learning objectives, preventing political backlash (OECD, 2016, p. 153). In this case, the OECD goes beyond knowledge brokerage but offers direct advice on policy choices for its members (Bali et al., 2019; Steiner-Khamsi et al., 2024).

The need and rationale for a national curriculum were shared among top-rank decision-makers, who attempted its development in partnership with the University of Los Andes first and the University of Antioquia later. Official publications document the government’s efforts in developing a national curriculum (see Ministerio de Educación Nacional, 2018, p. 94). However, the project sank before dawn. A project officer said: “We worked for two years to create a national

curriculum on maths, literacy, and sciences. It was not meant to be compulsory for school, just suggested... but it was never implemented” (Interview, GOV 17).

Allegedly, a national curriculum sparks controversy in Colombia by dealing with heightened policies, like teachers’ autonomy, or affecting sensitive topics, like the country’s armed conflict. Whenever a project is on the table, stakeholders ask: How will the conflict’s history be narrated? How do you harmonise regional differences between urban schools in Bogota and rural ones in El Chocó? Should reading follow the phonological or the dialogical method? Furthermore, the curriculum is seen as a constraint to teachers’ autonomy, with fierce pushback from teachers’ unions. Upon Minister Parody’s dismissal in 2016, the project was abandoned. Thus, despite consensus among government officials or the appropriateness of OECD’s advice, as long as tensions are not settled, the curricula remain a distant dream.

Discussion and conclusion

In the era of evidence-based policymaking, countries face growing pressure to consume available data and knowledge to inform their policy decisions (Grek et al., 2020; Lingard, 2013). In line with this, Colombia’s use of citations in policy documents grew at a fast pace in the last two decades. However, it displays a ‘staging’ use of evidence, understood as performative compliance with external demands. As a theatrical routine, Colombia can claim its policy planning is evidence-based, creating a *facade* of rationality projected to the outside world (Steiner-Khamsi, 2021). However, after scratching the surface, we observe a gap between scientific production and governments’ knowledge use. Similar to prior studies, Colombia demonstrates a low engagement with scientific knowledge, relying heavily on data from the bottom of the evidence hierarchy:

statistical reports, own-published reports, and grey literature (Nutley et al., 2007; Steiner-Khamisi et al., 2019).

What I call a staging strategy is more prominent after unpacking the sharp detachment between the cited knowledge and the lessons drawn behind SAWA policies. This is more relevant as SAWA policies are not only the OECD's preferred reform but the main direction in Colombia's education since the 2000s (Author, 2024). As observed, Colombia's policy documents repeatedly cite international evidence like PISA data for problem-framing purposes. Arguing a quality crisis due to students' underperformance, documents tap into international sources with legitimising purposes (Santos & Centeno, 2023; Takayama, 2010). In parallel, learning standards, assessments, and performance incentives are the preferred solutions to these policy problems (Béland & Howlett, 2016), although documents do not mention concrete examples or specific lessons from other countries. Remarkably, this is not the case when discussing other reforms, such as higher education financing or extending school day length, where evidence endorsed the adopted policy direction. Arguably, externalising to international policies and standards helps local policymakers advance contested shifts as increased testing or learning standards promoted by SAWA policies.

The analysis of whose and what evidence was used revealed the predominant OECD role in Colombia's knowledge network, endorsed by decision-makers. First, the country's preference for IOs knowledge –demonstrated in Table 2– is unsurprising. Prior studies have uncovered that experts' knowledge often gets lost in 'political translation', while IOs' knowledge predominates in countries' policy documents (Steiner-Khamisi et al., 2019). However, these studies analyse cases where scientific or expert commissions are ingrained in the policy cycle. In contrast, experts' role varies in policy-planning in Colombia, having a more central position in long-term planning documents (see documents #3, #8, #9, #18, and #22 in Annex 1) in contrast to NDP and ESP,

where political agendas are prioritised over the surveyed evidence (Barón, 2013; Zapata-Cortés, 2020).

Nonetheless, the growing OECD centrality as a knowledge broker cannot be detached from Colombia's membership. Whereas Colombia's ties with the OECD trace back to its 2006 PISA participation, the organisation's weight within its education sector changed after the accession process in 2011. Prior to the membership bid, authorities saw PISA participation as a stamp for the countries' improvement efforts. Participating meant legitimising Colombia's belonging to the global education polity, a strong enough reason to endure any scandal for the countries underperformance. Beyond legitimising purposes, PISA also benchmarked whom to compare with and learn from (Grek et al., 2020). However, authorities felt no pressure in following OECD's advice whatsoever, a situation that changed for following administrations.

In this sense, of Colombia's education sector review (2013-2016) meant a unique opportunity for the OECD to influence policy reforms.. As mentioned, the education sector analysis entailed a tight relationship between OECD staff and Colombia's authorities. Despite the OECD cannot impose educational reforms to new members (Carroll, 2014), the close scrutiny of the country's policies it created room for knowledge exchange and lesson-drawing between Colombia and OECD staff. Importantly, Colombian authorities and OECD staff were ideologically aligned regarding potential policy reforms and obstacles, facilitating their agreement on sensitive matters like school autonomy or standardized testing. OECD's context-sensitive approach to policy brokerage is evidence in its exhaustive 300-page long "Education in Colombia" (Niemann & Martens, 2018; Seitzer et al., 2023). In other words, beyond pre-conceived models like SAWA reforms, the profound understanding of member countries' specificities allows the OECD to offer

fine-tuned alternatives more likely to persuade decision-makers, and, consequently, to drive voluntary policy transfer.

Certainly, at the core of this study lays the relationship between evidence use, lesson-drawing, and policy transfer. Opposed to learning as a rationalistic assessment of ‘all’ existing evidence (Meseguer, 2005), this study demonstrates that lesson-drawing in public policy-making is contingent not only on limited pieces of available information for decision-makers but also on political viability and ideological preferences for certain knowledge sources (Weyland, 2007). In this transaction, knowledge brokers not only bridge science and politics but, more importantly, circumscribe the available evidence and possible lessons to learn from (Bali et al., 2019; Steiner-Khamsi et al., 2024). As demonstrated, Colombia’s decision-makers relied on the OECD knowledge and advice, assigning significant weight to their transferability assessments upon surveyed lessons (Rose, 1991; Weyland, 2007).

The subsequent recalibration of Colombia’s LSLAs and complexities around the national curriculum symbolises the intersection of lesson-drawing and contextual compatibility behind policy transfer. During the first investigated period (2002-2010), the Chilean SIMCE served as a model for Colombia’s LSLA reforms. Chile, praised for its PISA performance, lent Colombia its technical know-how and suggested the independence of ICFES as a quality monitoring agency. However, Colombia limited its borrowing to compatible lessons, as SIMCE’s schools’ rankings or performance payments fostering marketised dynamics were unlikely to prosper in Colombia (Díaz Ríos, 2024).

Years later, Colombia officials enlarged SABER’s relevance by increasing its frequency, coverage, and uses. The intensification of SABER, unfeasible a decade ago, was built upon those prior adjustments (Author, 2024). Tensions arose among officials who endorsed adopting high-

stakes and competitive forces for school improvements and others who feared their potential side effects. Acknowledging the quality assurance nature of Colombia's SAWA policies (Baxter & León Cadavid, 2021; Díaz Ríos, 2024), the OECD warned of the negative Mexican experience with high-stakes LSLAs. Showcasing its horizontal policy brokerage strategy (Seitzer et al., 2023), the OECD persuaded Colombia officials to adopt a less controversial use of its LSLAs, which could enlarge its acceptance among local stakeholders.

Finally, the push for national curricula exemplifies how political and institutional legacies are key determinants in policy transfer. The OECD's advice, argued on relatable examples and lessons from New Zealand and the United States experiences, was endorsed by Colombia's decision-makers. However, the sole idea of a national curriculum sparked multiple tensions among relevant stakeholders. During Santos' term, the peace agreement with armed groups was a source of heightened political disputes. Furthermore, SABER and teachers' evaluation reforms unleash disputes with the teachers' union (Baxter & León Cadavid, 2021), draining the Minister's political capital to face the curriculum reform (Author, 2024). Therefore, international support was insufficient amid the government's scarce domestic support ahead of the new curriculum project, resulting in a failed policy transfer.

In conclusion, the study has provided a novel account of the complex linkages between evidence use, brokerage, and lesson drawing, resulting in different transfer outcomes. In this vein, as pioneered by Baek (2021), the paper signals the importance of mixed-methods design, combining the analysis of document citations and content with other sources to examine and unpack the multiple layers underpinning policy reforms. Second, the newly required education sector review as part of OECD membership opens room for a new research agenda in policy brokerage and global policy transfers by delving into how this key organisation seeks to influence

countries worldwide. Given the new membership of Costa Rica (2021) and ongoing accession processes of Peru, Brazil, and Argentina, further questions arise regarding how and why the OECD is gaining influence in Latin America.

References

- Author. (2024). Details were omitted to ensure anonymity during the peer-review process.
- Baek, C. (2022). Internalising externalisation: Utilisation of international knowledge in education policymaking. *British Journal of Sociology of Education*, 43(1), 159–176.
<https://doi.org/10.1080/01425692.2021.2007357>
- Bali, A. S., Capano, G., & Ramesh, M. (2019). Anticipating and designing for policy effectiveness. *Policy and Society*, 38(1), 1–13.
<https://doi.org/10.1080/14494035.2019.1579502>
- Barón, C. B. (2013). La planificación del desarrollo: ¿problema económico, político o jurídico? *Opinión Jurídica: Publicación de la Facultad de Derecho de la Universidad de Medellín*, 12(24), 169–187.
- Baxter, J. G., & León Cadavid, M. C. (2021). From Political Intentions to Structural Interventions A Review of Two Decades of Education Policy Reforms in Colombia. In A. Rivas (Ed.), *Examining Educational Policy in Latin America* (pp. 96–108). Routledge.
- Béland, D., & Howlett, M. (2016). How Solutions Chase Problems: Instrument Constituencies in the Policy Process. *Governance*, 29(3), 393–409. <https://doi.org/10.1111/gove.12179>
- Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network Analysis in the Social Sciences. *Science*, 323(5916), 892–895. <https://doi.org/10.1126/science.1165821>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Cardoso, M. E. (2020). Policy evidence by design: International large-scale assessments and grade repetition. *Comparative Education Review*, 64(4), 598–618.
<https://doi.org/10.1086/710777>

- Carroll, P. (2014). Policy Transfer and Accession: A Comparison of Three International Governmental Organisations. *Journal of Comparative Policy Analysis: Research and Practice*, 16(3), 280–296. <https://doi.org/10.1080/13876988.2013.836782>
- Centeno, V. G. (2021). The OECD: Actor, arena, instrument. *Globalisation, Societies and Education*, 19(2), 108–121. <https://doi.org/10.1080/14767724.2021.1882958>
- Cowen, R. (2009). The transfer, translation and transformation of educational processes: And their shape-shifting? *Comparative Education*, 45(3), 315–327. <https://doi.org/10.1080/03050060903184916>
- Creswell, J. W., & Clark, V. L. P. (2018). *Designing and conducting mixed methods research approach*. SAGE Publications.
- Díaz Ríos, C. M. (2020). The Role of Policy Legacies in the Alternative Trajectories of Test-Based Accountability. *Comparative Education Review*, 64(4), 619–641. <https://doi.org/10.1086/7110767>
- Díaz Ríos, C. M. (2024). *Translating Global Ideas: How Policy Legacies and Domestic Politics Shape Education Governance in Latin America*. SUNY Press. <https://sunypress.edu/Books/T/Translating-Global-Ideas2>
- Díaz Ríos, C., & Urbano-Canal, N. (2023). The World Bank and Education Policy in Colombia: A Comparative Analysis of the Effects of International Organizations' Learning on Domestic Policy. *Journal of Comparative Policy Analysis: Research and Practice*, 25(1), 101–117. <https://doi.org/10.1080/13876988.2021.1991796>
- Dolowitz, D. P., & Marsh, D. (2000). Learning from Abroad: The Role of Policy Transfer in Contemporary Policy-Making. *Governance*, 13(1), 5–23. <https://doi.org/10.1111/0952-1895.00121>

- Dunlop, C. A. (2014). The Possible Experts: How Epistemic Communities Negotiate Barriers to Knowledge Use in Ecosystems Services Policy. *Environment and Planning C: Government and Policy*, 32(2), 208–228. <https://doi.org/10.1068/c13192j>
- Edwards, D. B., Moschetti, M. C., Grillet, P., Quilabert, E., & Caravaca, A. (2023). *Descentralización y privatización educativa en Colombia: Facilitadores y tensiones*. Education International.
- Esper, T. (2024). Shifting Paradigms: The Trajectory of School Autonomy and Accountability Policies in Colombia. *Nordic Journal of Studies in Educational Policy*.
- Grek, S. (2009). Governing by numbers: The PISA ‘effect’ in Europe. *Journal of Education Policy*, 24(1), 23–37. <https://doi.org/10.1080/02680930802412669>
- Grek, S., Maroy, C., & Verger, A. (2020). Introduction: Accountability and datafication in education: Historical, transnational and conceptual perspectives. In S. Grek, C. Maroy, & A. Verger (Eds.), *World Yearbook of Education 2021: Accountability and Datafication in the Governance of Education*. Routledge. <https://doi.org/10.4324/9781003014164>
- Haas, P. M. (1992). Introduction: Epistemic Communities and International Policy Coordination. *International Organization*, 46(1), 1–35.
- Högberg, B., & Lindgren, J. (2021). Outcome-based accountability regimes in OECD countries: A global policy model? *Comparative Education*, 57(3), 301–321. <https://doi.org/10.1080/03050068.2020.1849614>
- Hossain, M. (2022). Diffusing “Destandardization” Reforms across Educational Systems in Low- and Middle-Income Countries: The Case of the World Bank, 1965 to 2020. *Sociology of Education*, 95(4), 320–339. <https://doi.org/10.1177/00380407221109209>
- Karseth, B., Wahlström, N., & Pizmony-Levy, O. (2025). Tracing Policy Networks in Education

Reforms: A Comparative Study of Norway and Sweden. *Under Review*.

Lingard, B. (2013). *Politics, Policies and Pedagogies in Education. The Selected Works of Bob Lingard*.

Martens, K., & Jakobi, A. P. (Eds.). (2010). *Mechanisms of OECD Governance: International Incentives for National Policy-Making?* Oxford University Press.

McDonnell, L. M., & Weatherford, M. S. (2013). Evidence Use and the Common Core State Standards Movement: From Problem Definition to Policy Adoption. *American Journal of Education*, 120(1), 1–25. <https://doi.org/10.1086/673163>

Meseguer, C. (2004). What Role for Learning? The Diffusion of Privatisation in OECD and Latin American Countries. *Journal of Public Policy*, 24(3), 299–325.

Meseguer, C. (2005). Policy Learning, Policy Diffusion, and the Making of a New Order. *The ANNALS of the American Academy of Political and Social Science*, 598(1), 67–82. <https://doi.org/10.1177/0002716204272372>

Ministerio de Educación Nacional. (2018). *Informe de gestión 2014-2018*. Ministerio de Educación Nacional. https://redes.colombiaaprende.edu.co/ntg/men/pdf/INFORME_DE_GESTION_MEN_2014_%202018_%20ISBN.pdf

Niemann, D., & Martens, K. (2018). Soft governance by hard fact? The OECD as a knowledge broker in education policy. *Global Social Policy*, 18(3), 267–283. <https://doi.org/10.1177/1468018118794076>

Nutley, S. M., Walter, I., & Davies, H. T. O. (2007). *Using evidence: How research can inform public services*. Policy Press.

- OECD. (2013). *OECD 2013. ROADMAP FOR THE ACCESSION OF COLOMBIA TO THE OECD CONVENTION*. OECD.
[https://one.oecd.org/document/C\(2013\)110/FINAL/En/pdf](https://one.oecd.org/document/C(2013)110/FINAL/En/pdf)
- OECD. (2014). *Colombia: La implementación del buen gobierno*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/governance/colombia-implementando-la-buena-gobernanza_9789264202351-es
- OECD. (2016). *Education in Colombia*. OECD Publishing. <https://www.oecd-ilibrary.org/docserver/9789264250604-en.pdf?expires=1663001166&id=id&accname=ocid177456&checksum=199C36D991FF309DE8425BE098B830B9>
- OECD. (2021). *Summary of Colombia's OECD Post-Accession Progress*.
- Padgett, T., & Otis, J. (2012, December 4). The Colombian Comeback. From nearly failed state to emerging global player – in less than a decade. *TIME*, 179(16).
<https://content.time.com/time/covers/europe/0,16641,20120423,00.html>
- Parcerisa, L., & Falabella, A. (2017). La Consolidación del Estado Evaluador a Través de Políticas de Rendición de Cuentas: Trayectoria, Producción y Tensiones en el Sistema Educativo Chileno. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*, 25, 1–24.
- Parkhurst, J. (2017). *The Politics of Evidence: From evidence-based policy to the good governance of evidence*. Taylor & Francis.
<https://library.oapen.org/handle/20.500.12657/31002>
- Pawson, R. (2006). *Evidence-based Policy*. SAGE Publications Ltd.
<https://doi.org/10.4135/9781849209120>

- Pizmony-Levy, O., & Baek, C. (2022). Exploring the Architecture of Policy Knowledge: A Methodological Note. In B. Karseth, K. Sivesind, & G. Steiner-Khamsi (Eds.), *Evidence and Expertise in Nordic Education Policy* (pp. 59–76). Springer International Publishing. https://doi.org/10.1007/978-3-030-91959-7_3
- Porto de Oliveira, O., & Osorio Gonnet, C. (2023). Comparative Transfer Analysis: Policy Instruments, Space and Time. *Journal of Comparative Policy Analysis: Research and Practice*, 25(4), 367–384. <https://doi.org/10.1080/13876988.2023.2235308>
- Prell, C. (2012). *Social network analysis: History, theory & methodology*. SAGE.
- Radinger, T., Echazarra, A., Guerrero, G., & Valenzuela, J. P. (2018). *OECD Reviews of School Resources: Colombia 2018*. OECD. <https://doi.org/10.1787/9789264303751-en>
- Rose, R. (1991). What is Lesson-Drawing? *Journal of Public Policy*, 11(1), 3–30. <https://doi.org/10.1017/S0143814X00004918>
- Santos, Í., & Centeno, V. G. (2023). Inspirations from abroad: The impact of PISA on countries' choice of reference societies in education. *Compare: A Journal of Comparative and International Education*, 53(2), 269–286. <https://doi.org/10.1080/03057925.2021.1906206>
- Schrefler, L. (2010). The Usage of Scientific Knowledge by Independent Regulatory Agencies. *Governance*, 23(2), 309–330. <https://doi.org/10.1111/j.1468-0491.2010.01481.x>
- Seitzer, H., Baek, C., & Steiner-Khamsi, G. (2023). Instruments of lesson-drawing: Comparing the knowledge brokerage of the OECD and the World Bank. *Policy Studies*, 0(0), 1–21. <https://doi.org/10.1080/01442872.2023.2220282>
- Shipan, C. R., & Volden, C. (2008). The Mechanisms of Policy Diffusion. *American Journal of Political Science*, 52(4), 840–857.

- Sorensen, T. B., Ydesen, C., & Robertson, S. L. (2021). Re-reading the OECD and education: The emergence of a global governing complex – an introduction. *Globalisation, Societies and Education*, 19(2), 99–107. <https://doi.org/10.1080/14767724.2021.1897946>
- Sriprakash, A., & Mukhopadhyay, R. (2015). Reflexivity and the politics of knowledge: Researchers as “brokers” and “translators” of educational development. *Comparative Education*, 51(2), 231–246.
- Steiner-Khamsi, G. (2021). Externalisation and structural coupling: Applications in comparative policy studies in education. *European Educational Research Journal*, 20(6), 806–820. <https://doi.org/10.1177/1474904120988394>
- Steiner-Khamsi, G., Karseth, B., & Baek, C. (2019). From science to politics: Commissioned reports and their political translation into White Papers. *Journal of Education Policy*, 00(00), 1–26. <https://doi.org/10.1080/02680939.2019.1656289>
- Steiner-Khamsi, G., Martens, K., & Ydesen, C. (2024). Governance by numbers 2.0: Policy brokerage as an instrument of global governance in the era of information overload. *Comparative Education*, 0(0), 1–18. <https://doi.org/10.1080/03050068.2024.2308348>
- Takayama, K. (2010). Politics of Externalization in Reflexive Times: Reinventing Japanese Education Reform Discourses through “Finnish PISA Success.” *Comparative Education Review*, 54(1), 51–75. <https://doi.org/10.1086/644838>
- Vargas-Alzate, L. F. (2021). La adhesión de Colombia a la OCDE como una estrategia de inserción en un contexto desfavorable. *Oasis*, 35, Article 35. <https://doi.org/10.18601/16577558.n35.07>
- Verger, A., & Fontdevila, C. (2025). They are Born, Age and Survive. The Construction of Policy Models by International Organizations: The Case of the OECD’s School

Autonomy with Accountability. *Comparative Education Review*.

Verger, A., Fontdevila, C., & Parcerisa, L. (2019). Constructing School Autonomy with Accountability as a Global Policy Model: A Focus on OECD's Governance Mechanisms. In C. Ydesen (Ed.), *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex* (pp. 219–243). Springer International Publishing.
https://doi.org/10.1007/978-3-030-33799-5_11

Verger, A., Fontdevila, C., Rogan, R., & Gurney, T. (2019). Manufacturing an illusory consensus? A bibliometric analysis of the international debate on education privatisation. *International Journal of Educational Development*, 64(April 2017), 81–95.
<https://doi.org/10.1016/j.ijedudev.2017.12.011>

Volden, C., Ting, M. M., & Carpenter, D. P. (2008). A Formal Model of Learning and Policy Diffusion. *The American Political Science Review*, 102(3), 319–332.

Weyland, K. G. (2007). *Bounded rationality and policy diffusion: Social sector reform in Latin America* [Electronic resource]. Princeton University Press. <http://www.columbia.edu/cgi-bin/cul/resolve?clio13206175>

Ydesen, C., Milner, A. L., Aderet-German, T., Gomez Caride, E., & Ruan, Y. (2023). *Educational Assessment and Inclusive Education: Paradoxes, Perspectives and Potentialities*. Palgrave Macmillan.

Zapata-Cortés, O. L. (2020). Reflexión sobre los planes de desarrollo en Colombia. *Bitácora Urbano Territorial*, 30(3), 233–246. <https://doi.org/10.15446/bitacora.v30n3.86811>

Zapp, M. (2017). The World Bank and Education: Governing (through) knowledge. *International Journal of Educational Development*, 53, 1–11.
<https://doi.org/10.1016/j.ijedudev.2016.11.007>

Annex I. Full list of coded documents by reform wave 2002-2022

#	Document Name	Year	Author	Reform Wave	Type of document	References	Pages Per Document
1*	National Development Plan 2002-2006: "Towards a Communitarian State"	2003	National Planning Department	1	Policy planning document	6	23
2	Education Sector Plan 2002-2006	2003	Ministry of Education	1	Policy planning document	7	40
3	Vision Colombia 2019. A proposal	2005	National Planning Department	1	Policy knowledge document	333	384
4	Colombia's goals and strategies to achieve the 2015-MDGs	2005	National Council for Social and Economic Policy	1	Policy knowledge document	8	70
5	Towards an equitable and inclusive Colombia	2005	ECLAC & National Planning Department	1	Policy knowledge document	14	31
6	Balance of the Ten-Year Education Plan 1996 - 2005	2006	Ministry of Education	1	Policy knowledge document	14	48
7*	National Development Plan "Development for All" 2006-2010	2007	National Planning Department	1	Policy planning document	25	62
8	Education Vision for Colombia 2019. Discussion document.	2006	Ministry of Education	1	Policy knowledge document	13	18
9	Ten-year Education Plan 2006-2016	2006	Ministry of Education	1	Policy planning document	3	60
10	Education Sector Plan: "Education Revolution" 2006-2010	2007	Ministry of Education	1	Policy planning document	16	68
11	Actions and lessons: Education Revolution 2002-2010	2010	Ministry of Education	1	Policy knowledge document	114	266
12	CONEPS #3674. Guidelines for the strengthening of the	2010	National Council for Social and Economic Policy	2	Policy knowledge document	11	91

human capital development system							
13*	National Development Plan 2010-2014: Prosperity for All	2011	National Planning Department	2	Policy planning document	26	20
14	Education Sector Plan: 2010-2014	2011	Ministry of Education	2	Policy planning document	14	110
15	Approval and recommendations to request an international loan for the implementation of the Education Sector Plan	2011	National Council for Social and Economic Policy	2	Policy knowledge document	6	19
16	Amendment to CONPES document: Colombia's goals and strategies to achieve the 2015-MDGs	2011	National Council for Social and Economic Policy	2	Policy knowledge document	14	55
17*	National Development Plan 2014-2018: everyone for a new country		National Planning Department		Policy planning document	33	31
18	Education Sector Plan 2014-2018: Colombia, the most educated by 2025. Strategic guidelines for Ministry of Education's National Education Policy	2015	Ministry of Education	2	Policy planning document	6	125
19	Ten-year Education Sector Plan 2016-2026. The Pathway for equity and Quality	2011	Ministry of Education	2	Policy planning document	21	82
20	Strategy for the implementation of 2030 Sustainable Development Goals in Colombia	2019	National Council for Social and Economic Policy	2	Policy knowledge document	30	74
21	Monitoring and evaluation strategy for Colombia's ten-year education sector plan	2019	Ministry of Education	2	Policy knowledge document	14	42

2016-2026

22	A vision for Colombia 2050: discussion about the country's future	2022	National Planning Department	3	Policy knowledge document	308	273
23*	National Development Plan 2018-2022: A pact for Colombia, a pact for equity	2019	National Planning Department	3	Policy planning document	114	39
24	Education sector plan 2018-2022: a pact for equity, a pact for education	2021	Ministry of Education	3	Policy planning document	28	335
25	"Todos a aprender". Technical note.	2022	Ministry of Education	3	Policy planning document	55	83
26**	Education Quality: the road to prosperity. Administration Report 2010-2014	2014	Ministry of Education	2	Policy knowledge document	None	-
27**	Colombia, the most educated by 2025. Administration Report 2014-2016	2017	Ministry of Education	2	Policy knowledge document	None	-
28**	Ministry of Education Administration Report 2014-2018	2019	Ministry of Education	2	Policy knowledge document	None	-

* Only citations from a limited number of pages concerning education policies were coded

** Documents #26, #27 and #28 were included in document content analysis in spite of not having references for the bibliometric analysis.

4.3. Article 3: The politics of standardized assessments: A multi-scalar analysis of a global reform in Argentina

Summary

Journal: Journal of Education Policy

Article Status: Under Review, submitted 10/21/2024

Article's Table of Content

Introduction

The agents behind policy transfer

Politics, beliefs, and coalitions in policy reforms

Mechanism-based explanations of policy change

Methods: A process-tracing approach

Standardized assessments as a stone in your shoe: A national downgrading (2003-2015)

Buenos Aires (2007-2019): A political rivalry with national outreach

NLSAs comeback (2016-2019): A push for performance-based accountability in education

A turbulent time (2019-2023): Assessments between the covid-crisis and a fading authority

Córdoba (2007-2023): Accommodating policy beliefs and political benefits

Mendoza's testing resurgence (2019-2023): Between a learning crisis and global influences

Policy coalitions, cross-scalar dynamics, and the evolution of global reforms

The reawakening of a global script: the non-continuity of the temporal dimension

Policy variation through mechanistic processes Policy Variation Through Mechanistic Processes

Conclusion

References

The politics of standardized assessments: A multi-scalar analysis of a global reform in Argentina

Abstract

The global spread of policy reforms has intrigued scholars for decades. Since the early 1990s, large-scale learning assessments (LSAs) have become a cornerstone of educational policy worldwide, echoing discourses on accountability and quality monitoring. Yet, countries not only adopt and utilize LSAs in varied ways, but the evolution of LSAs also follows distinct trajectories. Moving beyond the adoption of LSAs, this study focuses on an under-explored aspect of policy transfer: why and how global reforms become institutionalized. This article sheds light on the mechanisms behind these processes by examining the adoption and development of LSAs in Argentina over time and across multiple levels. Data combines semi-structured interviews (n=45) and document analysis (n=23) for analyzing the diffusion of LSAs across four levels of governance: the national government and the provinces of Buenos Aires, Córdoba, and Mendoza. Combining a policy coalitions approach with mechanism-based explanation of policy change, the paper demonstrates: (i) the importance of cross-scalar political dynamics across the global, national, and provincial spaces in the transfer process; (ii) how global scripts are repurposed and re-introduced in light of continuous contentious politics; (iii) the role of policy agents and institutional constraints in enabling or blocking policy reforms.

Keywords: global reforms, policy transfer, temporal dimension, policy change, policy mechanisms

Introduction

Nowadays, large-scale learning assessments (LSAs) belong to the typical policy *repertoire* of education systems worldwide. Since the 1990s, LSAs spread was driven by international pressure to monitor education quality (Furuta, 2022; Kamens & Benavot, 2011), giving rise to a global “testing culture” (Smith, 2016). Nonetheless, assessments are deeply political tools, rendering heterogeneous models worldwide serving other interests beyond quality monitoring. Some foster market-like models where assessments serve school choice and competition (Parcerisa et al., 2022), and others promote LSAs to dynamize continuous improvement cycles and evidence-based decision-making (Clarke & Luna-Bazaldua, 2021; Pinkasz, 2022). Of course, opposing voices see LSAs as neoliberal means to delegitimize teachers’ social status and working conditions, advocating for non-standardized tools such as formative evaluations, among others (Bocchio et al., 2023; Oliveira & Feldfeber, 2024). Thus, there are different global “scripts” on LSAs’ goals, uses, and features, whose local imbrications are explained by the institutional, cultural, and political landscapes of the adopting contexts (Díaz Ríos, 2024; Verger et al., 2019; Ydesen et al., 2023).

LSAs mushroomed in Latin America under the 1990s decentralization reforms, regionally promoted by the World Bank and Inter-American Development Bank (Hossain, 2023). However, since then, two interlocking phenomena have shaped LSAs’ evolution in the region. First, the effects of decentralization reforms triggered inner cross-scalar disputes over state legitimacy, institutional autonomy, and political payoffs derived from education policies (Benveniste, 2002; Falleti, 2010). Second, the global rise of performance-based accountability pushed countries to increase monitoring and testing for education governance (Grek et al., 2020; Rivas & Sanchez, 2022). In response to this twofold process, countries like Mexico, Brazil, and Colombia moved towards test-based accountability models to surveil schools (Esper, 2024; Moreno Salto, 2023;

Termes & Pagès, 2024). In contrast, Chile attenuated LSAs' stakes due to students' protests, while it remains a core instrument in school governance (Falabella & Ramos Zincke, 2019). Clearly, LSAs' endurance in education systems is far from static, making their sustained centrality an intriguing puzzle.

Why countries buy into global reforms and how they are translated among different contexts have been core questions for comparative policy studies (Béland et al., 2018; Benson & Jordan, 2011; McCann & Ward, 2012). However, reforms destiny after adoption has received far less attention (Porto de Oliveira & Osorio Gonnet, 2023; Steiner-Khamsi et al., 2024). In long-lasting global policies such as LSAs, the international pressures and local actors that enabled its establishment will continuously seek to secure or alter the newly introduced policy. The effects of these ongoing processes yield different results: certain global reforms stick while others are abandoned. Hence, conceptualizing policy transfer as a continuous, this study examines what drives the adoption and institutionalization of LSA using the Argentine case.

The appeal to investigate the adoption and evolution of LSAs' in Argentina is threefold. Firstly, Argentina introduced its national LSA (NLSA) in 1993, along with the state restructuring and decentralization reforms (Rhoten, 2000). Decentralization rendered the National Ministry of Education (NME) 'school-less,' shrinking the nation-state's role to fewer functions, including evaluating the system's quality (Derqui, 2001).¹⁵ Unlike other Latin American countries, Argentina's NLSA was neither used for student certification nor ruled under an independent agency like Chile or Mexico (Diker et al., 2023). However, LSAs' uses and features have been

¹⁵ The Civilian-Military government (1976-1983) began a progressive dismantling of the 'teaching state', consolidated under Carlos Menem presidency (1989-1999). (Narodowski et al., 2016)

constantly reshaped in light of contradictory interests, becoming a political battleground for opposing coalitions (Benveniste, 2002; Moyano & Rodrigo, 2022; Rodríguez et al., 2018).

Second, Argentina is among the earliest members of international assessment programs, such as the UNESCO Regional Study 1997 and the OECD-PISA survey from 2000 onwards. Third, Argentina also presents sub-national LSAs thanks to its federal structure. During the 1990s, 15 of 24 provinces developed their domestic LSAs (PLSAs), although all except one were discontinued in the early 2000s after a severe economic and social crisis (DINIECE, 2003). However, a second wave of PLSA adoption began in 2011 and took off after the pandemic in 2020 (Ministerio de Educación de la Nación, 2023). This ‘second wave’ of provincial assessments coincided with the PISA participation from three jurisdictions: the Autonomous City of Buenos Aires,¹⁶ and the Provinces of Córdoba, and Mendoza.

Keeping this in mind, the study offers a multi-scalar and historical analysis of LSAs’ diffusion over time and across four transfer spaces in Argentina. Transfer spaces go beyond context, referring to the concrete policy subsystems and environments in which reforms are introduced, involving the agents –knowledge brokers, policy entrepreneurs, epistemic communities, etc.– and policy networks formed across scales and localities (Beech & Artopoulos, 2016; Porto de Oliveira & Osorio Gonnet, 2023). As a federal and decentralized country, vertical and horizontal dynamics between the national and provincial levels are essential in policy change, as each level portrays divergent policy goals and agendas (Beech et al., 2023; Termes & Pagès, 2024). Hence, I analyze the national space with its NLSA and three subnational cases: Buenos Aires, Córdoba, and Mendoza. Two research questions guide this paper: (1) What have been the

¹⁶ Thereafter, Buenos Aires.

main drivers of the evolution of LSAs in Argentina at both national and sub-national levels since the 2000s?; and (2) How have policy coalitions, cross-scalar dynamics, and local political actors influenced the development of LSAs in response to different global scripts? Data for this study consists of semi-structured interviews (n=45) and secondary sources, including policy documents and briefs, legislation, and academic literature.

This paper studies the adoption and changing trajectory of LSAs global scripts in Argentina. The argument of this paper is twofold. First, the study demonstrates that external pressures from international organizations or global norms can be necessary for transnational policy transfer but never sufficient to explain adoption (Marsh & Sharman, 2009; Verger et al., 2018). Local receptiveness is required in every case, making policy coalitions central to the transfer process (McCann & Ward, 2012; Steiner-Khamsi, 2021). Second, seeking to overcome methodological nationalism (Wimmer & Glick Schiller, 2003), the Argentine case shows the effects of mutually influencing dynamics between the transnational, national, and provincial levels in explaining why and how traveling policies are implemented, dismantled or repurposed across different settings.

Table 1. The four LSAs – Argentina, Buenos Aires, Córdoba & Mendoza

P = Pilot

	9 2	9 3	9 4	9 5	9 6	9 7	9 8	9 9	0 0	0 1	0 2	0 3	0 4	0 5	0 6	0 7	0 8	0 9	1 0	1 1	1 2	1 3	1 4	1 5	1 6	1 7	1 8	1 9	2 0	2 1	2 2	2 3																	
	Large Scale Learning Assessment																																																
ARGENTINA																																																	
BUENOS AIRES																																																	
CÓRDOBA																					Pilot Phase																												P
MENDOZA																																																	
	Participation in PISA																																																
ARGENTINA																																																	
BUENOS AIRES																																																	
CÓRDOBA																																																	
MENDOZA																																																	
	Government's Ideological Orientation																																																
ARGENTINA	Right									T*	Center-Left										Center-Right					Center-Left																							
BUENOS AIRES	Center-Right						Center-Left						Center-Right																																				
CÓRDOBA	Center-Right						Centrist						Center																																				
MENDOZA	Centrist						Center-Right						Center-Right																																				

T*= Transition governments after the 2001 crisis.

The agents behind policy transfer

In policy studies, scholars have argued for decades on the effects of structure versus agency in explaining policy diffusion and the interlink therein (Nullmeier & Kuhlmann, 2022). Beyond this historical dichotomy, the current state of the art sees macro phenomena like economic coercion or global norm-setting as complementary to explanations centered on agents' interests, power

struggles, and political conflict (Steiner-Khamsi, 2021). We see policy diffusion and its temporal recursivity as a dialectical process between global ideas, models, and agents interacting with domestic interests, institutional legacies, and political factors (Verger et al., 2018; Wimmer, 2021). In this sense, the transfer might be triggered by coercive pressure from international donors (Hossain, 2023). However, adopting localities are far from powerless: agents should buy into transferred policies, which will face resistance and endorsement among the different groups affected by the new reform (Díaz Ríos, 2024). Hence, focusing on the receptive part of the transfer process, I combine three theoretical lenses to explain the adoption and evolution of global policies: the Advocacy Coalition Framework (Jenkins-Smith et al., 2019), the Multiple-Streams Framework (Kingdon, 2014), and mechanism-based explanations of policy change (Nullmeier & Kuhlmann, 2022).

Politics, beliefs and coalitions in policy reforms

Undoubtedly, every new policy is inscribed in a specific context with its own history. In other words, countries have policy legacies that structure their norms, rituals, and agents' beliefs, creating power trade-offs among interest groups, often rendering them 'winners' and 'losers' (Díaz Ríos, 2024; Mahoney & Thelen, 2010). Political parties, unions, corporations, as well as policy entrepreneurs, experts or consultants, are in constant dispute to maintain or alter the existing power structures (Béland et al., 2018; Kingdon, 2014). Thus, to influence policy, actors organize into different coalitions based on their shared beliefs and motivations (Sabatier & Weible, 2007). Coalitions' beliefs render a remarkably stable three-tier system (Henry et al., 2022). First, deep core beliefs are fundamental normative orientations –i.e., the role of government or markets, left or right-wing politics, etc.–, which are not policy-specific but applicable to multiple subsystems –

i.e., education, economy, environment, etc– (Sabatier & Weible, 2007). Then, policy core beliefs are subsystem-related and encompass an “overall assessment of the seriousness of the problem, basic causes of the problem, and preferred solutions for addressing the problem” (Jenkins-Smith et al., 2019, p. 96). Finally, *secondary beliefs* are subsidiary to advancing core beliefs, also seen as the protective belt in Lakato’s terms (Jenkins-Smith et al., 2019). Importantly, coalition membership is not defined by institutional affiliation, like belonging to political parties, unions, or other interest groups. Instead, coalitions are cross-class formed by multiple agents with different institutional affiliations but connected by shared beliefs within the subsystem (Henry et al., 2022; Jenkins-Smith et al., 2019).

In each policy subsystem, such as education, coalitions frame policy problems, promote solutions, and mobilize political support to influence the public agenda (Mukherjee & Howlett, 2015; Weible, 2018). Using the Multiple Streams Framework logic, by coupling the three streams, coalitions push to materialize their beliefs into specific policies (Henry et al., 2014; Kingdon, 2014). In this endeavor, coalitions might benefit from global discourses to convey policy problems – known as externalization (Steiner-Khamsi, 2021), or forge alliances with international actors to channel resources favoring their choices. Likewise, coalitions learn from their own experience or that of others, altering their beliefs and opening room for new reform attempts (Henry et al., 2022). Of course, several factors beyond coalitions’ control can shake the public agenda, pressing for change. External shocks, like economic crises, can alter priorities and unleash unforeseeable political dynamics, reshuffling stakeholders’ stances ahead of specific topics (Sabatier & Weible, 2007). Likewise, political scandals, policy failure, or government turnover might impact the public agenda or well shift coalitions’ ability to influence government decisions (Jenkins-Smith et al., 2019).

The central argument conveyed in this paper is that LSAs' transfer in Argentina –their adoption, dismissal, re-emergence, or repurposing– is driven by competing coalitions' ability to influence the public agenda in favor of their policy preferences. Coalitions' mastery in dominating the political arena and imposing their policy preferences is directly linked with the observed trajectory. Schematically, coalitions are more or less aligned with different global scripts on LSAs, from which they arguably emulate repertoires, actions, and discourse. These scripts can be labeled as (i) Pro-Marketizers, (ii) Anti-Standardized Assessments, and (iii) Quality Assurance, as detailed in Table 2.

First, Pro-Marketizers advance LSAs' to produce performance data as a resource for school choice and competition, often supporting the deregulation of the education sector to emulate market-like conditions (Parcerisa et al., 2022). Hence, performance-based accountability through high-stakes testing is instrumental in advancing their larger policy goals for creating an educational market (Díaz Ríos, 2024; Verger et al., 2024). On the other corner, the Anti-Standardized Assessment script considers standardized tests a privileged tool of neoliberal dominance, portraying a de-contextualized image of complex learning processes while serving stigmatizing purposes for public schools, families, and teachers (Bocchio et al., 2023; Oliveira & Feldfeber, 2024). Opponents of LSAs belittle participation in international tests and are suspicious of standardized assessments regardless of their stakes (The Guardian, 2014). Finally, somewhat in between the first two scripts, a third template can be called 'Quality Assurance.' In this case, LSAs are considered instruments capable of enhancing quality via continuous monitoring, comparisons, and data analysis (UNESCO, 2019). In contrast to the Pro-Market template, the Quality Assurance script emphasizes evidence-based policy-making, dismisses rankings or competitive incentives,

and aims to complement LSAs with other assessment types or information tools, such as education monitoring systems –EMIS– (Clarke & Luna-Bazaldua, 2021; Verger et al., 2019).

Table 2: Policy Coalitions in their three-tier belief system

	Coalitions		
	Anti-Assessments	Quality Assurance	Pro-Marketizers
Deep core beliefs	<p>Promoter of welfare State policies and naturally opposed to the private sector involvement in public affairs</p> <p>Neoliberalism & market mechanisms detrimental to societal welfare and social justice goals;</p>	<p>Predominance of State role in public policy, although open to civic-society/private sector regulated participation;</p> <p>Improvement cycles including results evaluation, analysis, feedback, and successive action;</p>	<p>Private sector superiority; Small state role; Market mechanisms and competitions as social relations;</p> <p>Information for consumers' choice;</p>
Policy core beliefs	<p>Assessment should respond to both teaching goals and styles, decided by teachers, and not centralized</p> <p>LSAs are used for punishments, rendering de-contextualized measures of education quality;</p> <p>ILSAs as domination instrument from IOs</p>	<p>Educational assessment surpasses standardized testing, including several dimensions and multiple data points.</p> <p>Information for planning and improvement purposes</p> <p>ILSAs valuable as additional data to domestic tests</p>	<p>LSAs as a privileged instrument to monitor school quality</p> <p>Information for school choice & competition</p> <p>ILSAs key to economic prosperity</p>
Secondary beliefs	<p>Supporters of teachers' assessment practices</p> <p>Quality oversight spanning multiple datapoints on students and schools performance, resources and training</p>	<p>LSAs are one assessment component among other tools.</p> <p>Reduced frequency, low-stakes & sample-based</p>	<p>Performance incentives for competition & improvement</p> <p>Annual frequency, census-based & high-stakes</p>
Policy Problem	<p>Education quality improves with educational inclusion; Pervasive resource disparities among regions;</p>	<p>Education quality as a multi-layered problem: inequalities, low-skilled workforce, resource disparities, etc.</p>	<p>State bureaucracy, lack of public transparency, incentives, and competition</p>
Policy Alternatives	<p>Income or resource redistribution policies (i.e., conditional-cash transfers, netbooks, textbooks, etc.)</p>	<p>Information and Assessment mix: LSAs, student data, labor conditions, school resources</p>	<p>Standardized testing, school rankings, performance-based rewards</p>

Source: own elaboration

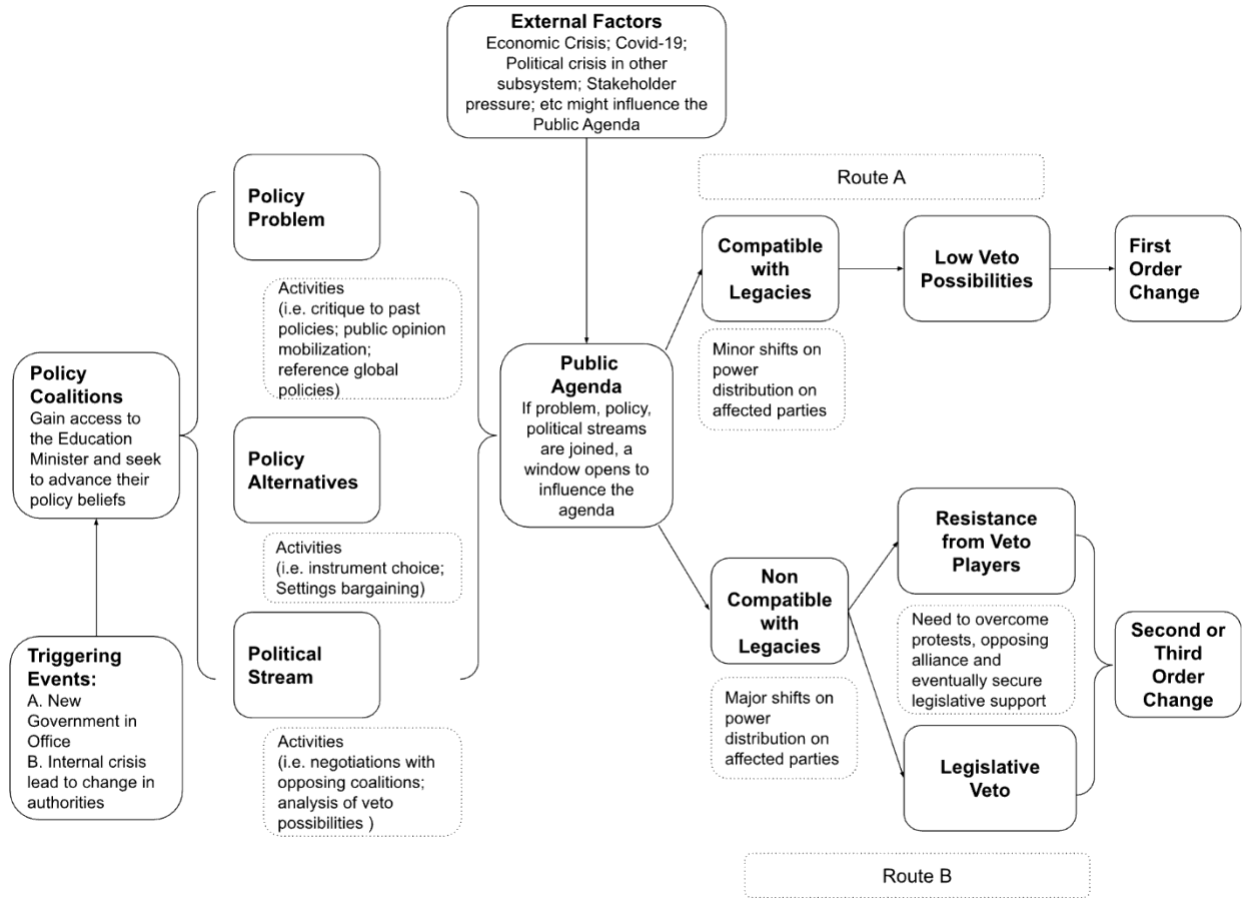
Coalition beliefs extend beyond their views on LSAs, including definitions of the state's role, the main policy problems, and how to solve them. Indeed, coalitions constitute ideal types that serve as a heuristic tool for analyzing the trajectory of LSAs'. As explained next, the degree of variation on LSAs results from the interaction between policy legacies, proposed changes, and veto ability from opposing coalitions.

Mechanism-based explanations of policy change

In recent years, what was called 'the mechanistic turn' in policy studies focused on exploring the black-box leading to policy change (Capano et al., 2019). Mechanisms are largely theorized and used in comparative policy studies, including economic coercion to drive structural reforms or normative emulation as a driver of organizational homogeneity (Marsh & Sharman, 2009; Wimmer, 2021). Yet, more often than not, how they exactly operate is underexplored (Nullmeier & Kuhlmann, 2022). Mechanisms are causal processes requiring explicit theorization of their composing entities –actors, organizations, or structures– that engage in specific activities transmitting the causal forces (Beach & Pedersen, 2019). As mechanisms are not observable, researchers rely on casual-process observations as if they were a smoking guns: they link agents, activities, and events, enabling causal inferences (Collier, 2011, p. 20; Mahoney, 2010). Importantly, mechanisms are not just another variable, as in variance-based case methods, where different variable values alter mechanistic outcomes (Mahoney, 2010). In contrast, mechanisms are viewed as invariant processes triggered if the cause and relevant scope conditions that allow its functioning are present (Beach & Pedersen, 2019).

Building on the view of coalitions as a primary driver of change, the theorized mechanism shown in Figure 1 has two parts and works as follows. The presence of a new government will

unleash disputes among coalitions over access to decision-making roles, such as the Minister of Education. In Argentina, national and provincial ministers can influence assessment policies on each scale. To translate beliefs into policies, coalitions will frame certain issues as problematic – i.e., low educational attainment– through several activities, including criticizing past government policies, referring to international standards –i.e., externalization–, or mobilizing public opinion. In parallel, they select their preferred policy solution and engage in political bargaining to secure support ahead of reforms. Coalition members will aim to couple the problem, policy, and political streams to influence the public agenda. Once a window for change occurs, two routes open up, as shown in the second part of the mechanisms. When policies are compatible with existing legacies, they provoke small shifts into power trade-offs and benefits among interest groups. Therefore, they concede scarce veto opportunities to opposing groups but render only minor policy changes. Alternatively, policies that challenge the ‘status quo’ or affect existing norms will face considerable resistance from affected stakeholders, possibly requiring ample political support within the legislature. Depending on the coalition’s ability to overcome veto points, reforms will be adopted as intended, with modifications, or rejected. Therefore, the outcome is contingent on how actors deploy their resources and strategies in each case.



Source: self-elaboration

Figure 1. Causal Mechanism for Policy Coalition Influence on LSAs

Methods: A process-tracing approach

This study follows a process-tracing (PT) approach to shed light on the unobserved drivers and mechanisms re-shaping standardized tests. In qualitative policy research, PT is inscribed in the ‘mechanistic turn,’ which focuses on connecting triggering causes leading to specific outcomes (Capano et al., 2019; Collier, 2011; Hall, 2013). PT is grounded in a deterministic ontology, meaning that things happen for a reason, and a probabilistic epistemology, as our certainty over the casual relationship lies in the quality of the produced evidence (Beach & Pedersen, 2019). Therefore, PT pays close attention to the sequence of events from a historical standpoint and

focuses on agents such as coalitions and their members over structure, namely institutions or political parties, as animators of the unobserved mechanisms (Nullmeier & Kuhlmann, 2022). Despite PT's potential to enlighten the black box between causes and outcomes in policy studies, only a few empirical studies have taken such an approach (Díaz Ríos, 2024; Löblová, 2018).

My design followed a theory-testing PT approach to account for how the Advocacy Coalition Framework and the Multiple-Streams approach explain policy change (Beach & Pedersen, 2019; Henry et al., 2022). Therefore, I selected the cases where the mechanisms could be present, at least in theory (Beach & Pedersen, 2019). I made a within-case analysis of the identical policy variation mechanism in four transfer spaces: the National level, the City of Buenos Aires, and the Provinces of Mendoza and Córdoba. The chosen provinces rank among the richest and have the most developed educational system, and all three participate in PISA as independent economies (see Table 1). The four cases have been governed by different political parties, from the center-left to the center-right spectrum, and they illustrate two contrasting LSA pathways. Argentina and Buenos Aires serve as cases of LSAs repurposing over time, while Córdoba and Mendoza showcase the re-emergence of a global script in recent years. Then, the critical contextual factors explaining policy variation include party politics, institutional frames, and external shocks such as economic crises, the Covid-19 pandemic, or significant political shifts.

Data for this study consists of semi-structured interviews (n=45) and secondary sources (n=37), including policy documents, briefs, legislation, and academic literature. As PT aims to rebuild the sequence of relevant events linking causes to outcomes, secondary sources such as academic publications or gray literature are integrated as part of valuable historical data (Beach & Pedersen, 2019). Interviews involved key informants, including former national education authorities (n=10) or at the evaluation agency (n=12), provincial educational authorities (n=11),

international organizations' consultants (n=3), academics, and policy experts (n=3), union members (n=4) and leaders of education-oriented civil society organizations (n=2). Furthermore, certain interviewees have worn different hats, i.e., serving at both provincial and national levels, working as consultants, etc. Interviews were verbatim transcribed and thematically coded (Saldaña, 2021), and identity is masked following ethical considerations.¹⁷ I analyzed both documents and interviews following a thematic content analysis, focusing on how and what problems were framed, what views agents have on LSAs, with whom they shared ideas, or why policies have changed (Braun & Clarke, 2006). Furthermore, from data analysis and guided by my knowledge of existing views on educational assessment, I inductively derived coalitions' beliefs and membership (Henry et al., 2022). Finally, I focused on identifying traces of mechanistic evidence to confirm or reject the guiding tenets of our policy variation mechanism, as described above (Collier, 2011; Beach & Pedersen, 2019).

Next, results from the four cases are presented in three historical periods corresponding to different national administrations: 2003-2015, 2015-2019, and 2019-2023. These distinct periods offer a clear-cut dynamic perspective for identifying the effects of contentious politics between the federal and provincial governments over time.

Standardized assessments as a stone in your shoe: A national downgrading (2003-2015)

¹⁷ The project was approved by the Institutional Review Board (IRB) of Teachers College, Columbia University (IRB #23-237). Participants agreed to the project informed consent, where ethical considerations include potential risk of identity recognition beyond the researcher actions given the limited number of people occupying their roles.

In the aftermath of a profound social, political, and economic crisis, the newly elected center-left Peronist government deployed a discursive and political confrontation with past neoliberal policies promoted by former right-wing presidents Carlos Menem (Peronist, 1989-1999) and Fernando De La Rúa (Radical Party, 1999-2001), shown in Table 1. If the 1990s epitomized ideas of privatization, decentralization, and a small state, the presidencies of Nestor Kirchner (2003-2007) and Cristina Kirchner (2007-2015) attempted the reinstallation of a welfare State, including the recentralizing of education governance. The enduring social and economic effects of the 2001 crisis were problematized as social inequalities and disparities in the need for compensatory state policies. However, the center-left Peronist coalition had weak foundations, with just 22,25% of electoral support.¹⁸ In search of a larger political and social base, Peronists allied with the national teacher's union CTERA¹⁹ (Chambers-Ju, 2024), which opposed national governments since 1983 and was radically against standardized testing. As an illustrative gesture signifying the emerging partnership, President Kirchner's first public intervention was solving a profound salary dispute in Entre Rios province, favoring teachers' demands (Obarrio, 2003). Teachers unions became crucial allies in the government's symbolic crusade against neoliberal policies, where IOs had a terrible reputation after the debt crisis.²⁰

Ahead of the enduring consequences of the 2001 breakdown, the government's agenda focused on schooling access and retention policies to tackle inherited and pressing social inequalities. During Kirchners' administrations, standardized assessments were a 'stone in the shoe,' becoming increasingly uncomfortable over time. The NLSAs, suspended in 2001 and 2002

¹⁸ Nestor Kirchner ended second behind former President Carlos Menem who obtained 24,45% of votes. However, Menem withdrew his candidacy from the run-off making Kirchner president.

¹⁹ Confederation of Education Workers of the Argentine Republic.

²⁰ During 2005, a left-wing block joined by Venezuelan President, Julio Chavez, among other regional leaders opposed the United State attempt to expand NAFTA led by President George W. Bush.

due to the economic crisis, resumed in 2003 with a downgraded role in the NME's agenda, moving from annual to bi-annual frequency responsive to the new assessment's goals: to provide educational authorities with a diagnosis to act upon. In parallel, the former Institute for the Development of Educational Quality (IDECE), "a semi-autonomous agency responsible for the administration of NLSAs" (Benveniste, 2002, p. 9), was replaced by the National Directorate of Educational Information and Education Quality ("DINIECE"). In 2006, Argentina also returned to PISA.²¹ After poor results, officials considered withdrawing from the test, a position shared by several government allies, such as CTERA and several minister officials.²² The NLSA's downgrading continued in 2010, when the government reduced testing frequency even further, now every three years. The decreased frequency, criticized by World Bank specialists,²³ proves the governments' discomfort with standardized testing.²⁴

Despite the decreased frequency, the evaluation mainly had technical adjustments during this period, including score scales and item designs (Moyano & Rodrigo, 2022). NLSAs were not seen as part of the solution to educational challenges, also evidenced in their scarce use for policy planning.²⁵ The political alliance with CTERA, fiercely opposed to standardized assessments, conditioned alternative pathways for NLSAs' uses. Furthermore, fiscal efforts in educational expansion showed stagnated test scores, yielding critiques over the government's neglect of education quality (Rivas & Sanchez, 2022).

²¹ Argentina, whose first PISA participation was in 2000, withdraw from the 2003 crisis due to the social and economic crisis.

²² Interview with former National Education Authority

²³ Interview with former Evaluation Agency Authority

²⁴ For a complete historical sequence of periodicity, coverage and subject changes see Diker et al., (2023).

²⁵ Interview with former Provincial Educational Authority

Arguably, the National Ministry of Education's ineptitude or unwillingness to craft a coherent narrative around NLSAs became its 'Achilles heel,' a target of fierce critiques from government opponents. During Kirchner's second term (2011-2015), Peronists faced several scandals due to official statistics manipulation, which created government mistrust (Llorente, 2016). The NME's decision to change LSAs assessment scales in 2013 (Ganimian, 2015) and scarce data availability on prior editions²⁶ were framed by opponents as data manipulation. The corollary of this saga was Argentina's exclusion from PISA 2015 results due to a sampling mismatch with prior editions, leveraged by the incoming government to criticize Kirchner's policies and impose a new agenda.

Buenos Aires (2007-2023): A political rivalry with national outreach

Since Mauricio Macri became Buenos Aires' mayor in 2007, he embodied a vivid political confrontation with the Peronist government led by Cristina Kirchner (2007-2015). Most of its policy decisions staged its rivalry with the national government. In 2010, Esteban Bullrich was appointed Buenos Aires Minister of Education (2010-2015), changing the city's educational agenda. Minister Bullrich shared with President Macri an elitist origin and a strong affinity for neoconservative political views, including a preference for private-sector methods (Matovich & Esper, 2023). Bullrich embodied pro-market values and beliefs, rapidly becoming a key entrepreneur in advancing performance-based accountability policies. For instance, Bullrich opened the door for the landing of "Teach for Argentina"²⁷ and, years later, for the Varkey Foundation (Matovich & Esper, 2023). Upon his arrival, he promoted a new standardized PLSAs

²⁶ NLSAs reports are available from 2009 onwards (Diker et al., 2023).

²⁷ Interview with Educational Expert

distancing from its predecessor's view,²⁸ who had created the new Directorate of Educational Assessment in December 2007. Thus, Buenos Aires launched two new tests in 2010. First, a 9th-grade assessment was abandoned two years later. Second, the Secondary School Finalization Examination (FESBA, in Spanish) continued until 2017. FESBA differed considerably from prior PLSAs in scope, targeted population, stakes, technical aspects, and policy goals. The exam was multiple-choice in Mathematics and Literacy, with census coverage on a yearly basis, although without any consequences for test-takers, teachers, or schools.

The repackaged assessments conveyed a quality assurance discourse, fostering information and testing for educational improvement. The annual test frequency was argued to monitor the achievement of the city's curriculum standards (Buenos Aires Ciudad, 2013, p. 2). However, FESBA was initially intended to act as a certification exam for secondary students, meaning a high-stakes test.²⁹ This non-trivial caveat reveals an unsolved tension between pro-market versus quality monitoring views within the city educational coalition. A minority group, led by Minister Bullrich, attempted to imprint test-based accountability reforms by increasing LSAs' stakes, publicizing school results, and rewarding top-performing schools.³⁰ However, these goals would find limits in different educational experts and assessment technocrats inside the Ministry, who ruled out stakes circumscribing testing for monitoring and improvement purposes.

Buenos Aires' assessments were integral to a broader accountability mix. In 2013, the city's new education quality head, Silvia Montoya, led the creation of two quality and equity indexes and a Primary Education Finalization Examination (PESBA, in Spanish).³¹ Then, in 2014,

²⁸ Interview former Provincial Evaluation Authority

²⁹ Interviews with former Provincial Evaluation Authorities

³⁰ Interview with former Provincial Evaluation Authority

³¹ Interviews with former Provincial Evaluation Authorities

a new teachers' assessment program was launched jointly with the Quality and Equity Evaluation Unit (UEICEE in Spanish), independent of the Education Ministry. Interestingly, despite Bullrich's attempts, Buenos Aires accountability system did not adopt market-like features. For instance, the indexes never became public-facing due to a twofold resistance. Internally, technocrats were against publicizing results. However, the most fierce opposition came from the Directorate of Private Schools and the Archdiocese of Buenos Aires.³² Results dissemination could unmask underperformance in private schools, jeopardizing the state subsidies received by the Catholic Church to run their private schools (Moschetti et al, 2023). Therefore, internal-facing indexes did not influence pro-choice dynamics between families and schools as originally intended. Likewise, the city quality unit's independence as an audit agency has been merely formal, as all executive directors since 2014 have worked in close association with the Minister of Education.³³ In sum, the regional accountability system rendered a quality assurance logic for education improvement rather than a market model with school competition and choice.

Beyond the official narrative, Buenos Aires' rationale for a revamped assessments and accountability policy responded to the political confrontation with the national government (López Leavy, 2024). The mix formed by student and teacher evaluations, quality indexes, an evaluation unit, and school reports contrasted with the National government rebuttal of NLSAs and educational accountability. Buenos Aires' assessment policies were publicly argued to favor curriculum monitoring and information empowerment by decoupling them from national decisions and timelines. However, assessments were a political tool. City leaders criticized NLSAs' discontinuity and doubted the results' validity in the context of severe critiques of statistical

³² Interview with Educational Expert

³³ Interviews with former Provincial Evaluation Authorities

manipulation. Hence, assessments granted Buenos Aires a powerful instrument for discursive rivalry with the national government. The city testing fever also included participation in ILSAs, such as PISA since 2012, and TIMS and PIRLS from the IEA since 2011. In 2016, Buenos Aires showed impressive improvement in PISA 2015 edition, which contrasted with the exclusion of Argentine results (discussed next). This contrast fed the political contestation, although the city's Minister admitted students were sensitized after suspicion of cheating (Clarín.com, 2016). Nonetheless, the deployed testing machinery sufficed Montoya to become head of UNESCO UIS.³⁴

In the years after, Macri's arrival as Argentina's president would influence Buenos Aires' testing policy, becoming more aligned with the national agenda. For instance, "FESBA" moved from 5th to 3rd year of secondary education as it overlapped with the national test,³⁵ and, in 2021, it relaxed its frequency to bi-annually. However, the city would not resign from this valuable governing instrument, which grants political and discursive autonomy from the national sphere.

NLSAs comeback (2016-2019): A push for performance-based accountability in education

After 12 years of Néstor Kirchner and Cristina Fernández de Kirchner's administrations, the elected *Cambiamos* party would position its assessment policy as a gateway from its predecessor. Esteban Bullrich jumped from Buenos Aires to the NME. *Cambiamos'* agenda merged a smaller

³⁴ Interview with Educational Expert

³⁵ Renamed Third Year of Secondary Education (TESBA, in Spanish) since 2017, had an annual occurrence until 2021.

national role in favor of federalism and a ‘quality turn’ as a guiding principle (Sanchez & Rivas, 2021). The national government problematized education quality as the most pressing problem of Argentina’s school system. It criticized inefficiencies from Kirchners’ welfare policies on educational inclusion and Peronists’ dismissal of data transparency and accountability. Therefore, educational assessments, information, and accountability signified a break from the past. The corollary of assessments used for partisan confrontation was Argentina's exclusion from PISA 2015. According to the OECD (2016), Argentina's 2015 sample was not comparable with prior cycles, while it also showed improved scores.³⁶ However, to maintain political strife and accuse Peronists of data manipulation and poor results, *Cambiamos* authorities advocated for Argentina’s exclusion.³⁷

During *Cambiamos* national tenure, assessment and accountability ideas dominated the Minister’s policy. In framing underperformance and information *vacuum* as policy problems, *Cambiamos* propelled a re-branded NLSA, now termed “APRENDER” as their policy solution. Educational data and performance information were stressed as essential inputs for much-needed improvement. Furthermore, the government created a National Secretariat of Educational Evaluation with more resources in the Ministry structure while a quality agency project was on file.³⁸ The NLSAs’ new aims were harmonized with the technical aspects of the tests. Whereas the assessment instrument did not change significantly despite its rebranding (López-Levy, 2024; Moyano & Rodrigo, 2020), APRENDER moved from a sample base to a census base, regained an annual frequency like in the 1990s, covered 6th and 12th grades, and included new school results’

³⁶ This information was replicated by multiple interviewees with access to results.

³⁷ Interview with National Evaluation Unit Officer

³⁸ Law Project N. 29891 from 2016 for creating an independent educational evaluation and information agency.

reports. Furthermore, APRENDER results became benchmarks in the government's strategic plan, although no stakes for schools were associated in case of underachievement. Nonetheless, NLSAs' periodicity and coverage were not sustained due to the economic crisis in 2018, combined with the technical challenges of producing NLSAs in such a short time.³⁹

Behind the new assessment and educational information policies laid an unresolved dispute between market-model supporters and advocates of quality improvement logic within *Cambiamos*. Inside the party, a highly influential but small group championed market-like assessment changes, opposed by other members with an educational background. This dispute came to light during a presidential address presenting APRENDER 2016 results. In his speech, President Macri criticized the legal impediment to publicizing results at the school level (Braginski, 2018). Subsequently, members of *Cambiamos* presented successive reform projects to National Education Law Article 97, without much political traction.⁴⁰ The publication of results would enable rankings and, more importantly, was seen as a weapon against teachers' unions, unleashing strong resistance from opposing groups (Feldfeber et al., 2018). Similar reforms were also truncated early, while 'successful' ones underwent sensitive alterations, trimming their competitive features. For instance, in 2017, the government promoted a new teachers' evaluation called "ENSEÑAR." This evaluation was initially intended to become a teacher assessment similar to Peru's merit-pay reform in 2013, but it was resisted internally by coalition members⁴¹ and externally by unions (Feldfeber et al., 2018). Additionally, *Cambiamos* attempted to transform the Evaluation Secretariat into an independent assessment agency, emulating Chile, Colombia, or Mexico.

³⁹ Interview with former National Evaluation Unit Authority.

⁴⁰ Project N. 0183-D-2018 from 2018. Presented by Cambiamos Congressman Ricardo and Banfi, available at: <https://www2.hcdn.gob.ar/proyectos/proyectoTP.jsp?exp=0183-D-2018>

⁴¹ Interview with educational leader in Cambiamos

Although the project gathered a larger consensus among *Cambiamos*' members, it also sank due to the government's weakness after the 2018 economic crisis.⁴²

A turbulent time (2019-2023): Assessments between the covid-crisis and a fading authority

The 2019 triumph of a Peronist-led political coalition was marked by its slogan “*return to do it better*,” promising not to repeat previous mistakes. Public distrust in government statistics had signed Néstor Kirchner and Cristina Fernández de Kirchner's administrations, placed at the top of the ‘lessons learned.’ In education, the political alliance with teachers' unions continued, as Nicolás Trotta, the newly appointed Minister, was the Dean of a union-led private university. Nonetheless, this time, the head of the evaluation unit endorsed ‘Quality Assurance’ views on assessments and information (Diker et al., 2023). In contrast to the early 2000s, low educational attainment had become an elephant in the room. Therefore, underperformance was argued on social inequities combined with an inefficient use of school data for decision-making. Upon arrival, the Ministry formed the Council of National Education Quality to demonstrate a different stance than the past Peronist administration. Likewise, the new Secretariat of Information and Educational Evaluation –replacing the Secretariat of Educational Evaluation– agreed on a national assessment and information plan with different stakeholders, including CTERA.⁴³ Interestingly, the new Secretariat's name and mission laid in between prior governments: standardized testing was one data source engrained in a larger information machinery, including school access data,

⁴² Interview with former National Education Authority

⁴³ Interview with former National Education Authority

retention, drop-out, socioeconomic segregation, or school material conditions, among others.⁴⁴ The new approach was rubricated in a federal assessment policy in agreement with provinces.⁴⁵ The “Assessment Plan 2021-2022” stabilized APRENDER frequency and coverage to effectively serve planning and monitoring purposes, fostering a new “Federal Information Educational Network” to integrate test results with school performance and students’ trajectory data into a centralized EMIS named “Integral System of Digitalized Educational Information” (SiNIDE, in Spanish).⁴⁶ Furthermore, the National Ministry of Education deployed a set of formative assessment tools through a digital platform during Covid-19 to provide schools with resources to monitor student progress.

However, the Ministry’s objectives were sabotaged by the Covid-19 crisis. For the most part of 2020, Argentine schools were closed. In 2021, the Ministry led a national agreement at the Federal Education Council⁴⁷ for schools reopening, including unions’ support.⁴⁸ Nonetheless, without prior notice, the President announced on television that school closure would continue until vaccination reached higher levels,⁴⁹ leading to the Minister’s resignation shortly after. Buenos Aires and Mendoza disobeyed the National government and reopened their schools. In a similar scenario, NLSAs were canceled in 2020, but the National Ministry of Education planned to apply them in 2021. After Minister Trotta’s resignation, new authorities initially decided to postpone APRENDER in 2021 in line with the president’s order (Vallejos, 2021). However, the cancellation

⁴⁴ Interview with former Head of Assessment Unit

⁴⁵ Ministerio de Educación. (2023). *Plan Nacional de Evaluación Educativa 2023-2024*. Available at: <https://siteal.iiep.unesco.org/bdnp/4312/plan-nacional-evaluacion-educativa-2023-2024>

⁴⁶ SiNIDE was created in 2012 by National Ministry of Education Resolution N° 1.041/12

⁴⁷ The Federal Education Council was created in 1993 as a result of the Federal Education Law. It is chaired by the National Minister of Education, and formed by the 24 jurisdictions, and representatives from National Public and Private universities. Since 2006, its resolutions are mandatory for provinces.

⁴⁸ Interview with former National Education Authority

⁴⁹ Interview with former Provincial Education Authority

attempt triggered a powerful political backlash in most provinces, demanding the assessment to measure the effect of school closure on students' learning.⁵⁰ Provincial pressure triumphed, resulting in APRENDER re-installation in 2021.

In broad terms, the new Ministry authorities maintained a similar orientation to their predecessors. During the whole period (2019-2023), educational evaluation aimed to assess the school system integrally, including teaching and learning; students' school trajectories, school organization; and the material and institutional resources.⁵¹ Officials would focus mainly on developing the overdue education information system, although certain provinces were far ahead in the process. Clearly, the National strategy on testing and accountability aligned with a quality assurance logic, aimed at creating multiple data points from different instruments to use them as systemic improvement inputs. However, the president's sudden and unconsulted changes hindered the Ministry authority, incapable of articulating an effective federal evaluation policy. Put differently, national inconsistencies and the urge to monitor covid-10 effects gave provinces an excuse to advance their assessment policy.

Córdoba (2007-2023): Accommodating policy beliefs and political benefits

The case of Córdoba lies in between cross-scalar political confrontations and the appropriation of the quality assurance discourse. In political terms, Córdoba is a *rara avis* in the Argentine

⁵⁰ Interview with former Provincial Education Authority

⁵¹ Argentina Federal Evaluation Plan 2023-2024.

ecosystem. Since 2003, it has been governed by the local branch of the Peronist party. However, against common sense, it was politically opposed to national Peronist administrations (2003-2015 & 2019-2023) and closer to *Cambiamos* (2015-2019). Another singularity in Cordoba has been the 16-year continuity of the same Minister and Secretary of Education between 2007-2023. In this context, what characterized the province's educational administration has been its ability to accommodate a changing national environment. While the introduction of PLSAs occurred in 2019, the Provincial Evaluation of Learning Attainment (PRISMA, in Spanish) origins trace back to an institutional auto-evaluation for schools implemented in 2012.⁵² The auto-evaluation was thought of as a reflective exercise for schools to analyze key dimensions, such as (i) students' trajectories; (ii) teaching and the academic regime; (iii) the school's pedagogical approach; (iv) labor organization and teaching conditions (Gobierno de Córdoba, 2013). Then, in 2013 Córdoba's PLSAs began as a non-representative sample-based mathematics and literacy test. The instrument was built by disciplinary specialists in consultation with school teachers, yet the Ministry of Education made internal use of results. In 2019, the exam was re-launched, adopting most of its current features. The test was redesigned with psychometricians from Córdoba's National University (Instituto de Investigaciones Psicológicas, 2021). Furthermore, since 2019, PRISMA has covered a representative sample of 3rd-year students in primary and secondary schools with a bi-annual occurrence, but it has never had consequences associated with it. After Covid-19, PRISMA 2021 showed worse results compared to 2019. In response, in 2022, authorities developed a census-based pilot assessment covering last year's secondary students to measure the effects of the pandemic.

⁵² Interview with provincial education authority.

In Córdoba, the PLSA's re-emergence was at the crossroads of governance and political reasons. From a governance standpoint, Córdoba authorities learned from Buenos Aires the benefits associated with owning their test.⁵³ The assessment was a powerful instrument for monitoring schools and articulating an educational improvement discourse ahead of teachers and families. As the NLSAs covered grades 6th and 12th, Córdoba's focused on grades 3rd and 9th. Despite its low stakes, the PLSAs spurred suspicion from school teachers and principals, particularly union leaders.⁵⁴ Overcoming this initial resistance was key in installing the assessment, aided by its low consequences.⁵⁵ For provincial authorities, counting on students' performance data halfway through their school trajectories was vital to react⁵⁶. Results were used privately, uploaded into the provincial EMIS, and accessible by principals and teachers.

Beyond technical or governance rationalities, political drivers explain Córdoba's test re-emergence. Clearly, the provincial space was influenced by national politics. *Cambiamos* endorsement of international assessment discourses promoting a nationwide 'quality turn' shifted the domestic views on LSAs. Undoubtedly, the *Cambiamos* tenure (2015-2019) succeeded in installing a new assessment rhetoric, portrayed as a synonym of transparency and accountability. Hence, the nationally-advanced assessment culture resonated locally, mobilizing provincial authorities to leverage the new norms.⁵⁷ Emulating Buenos Aires, Córdoba became the second province to participate in PISA in 2018. Then, in 2019, they revamped their PLSA pilot test into its current form.

⁵³ Interview with provincial education authority

⁵⁴ Interview with union members and provincial teachers

⁵⁵ Interview with provincial union member

⁵⁶ Interview with provincial education authority

⁵⁷ Interview with experts

Owning data from PLSAs and international tests reinforced the political differentiation with the national government, particularly after Kirchners' return in December 2019. The Peronist comeback, tainted by an adverse history with assessments, represented an opportunity for Córdoba's authorities to show its independence from national policies. Local stakeholders criticized the continuous changes in NLSAs frequency, coverage, or grades among national administrations, as well as the absence of federal agreements around educational assessment and information.⁵⁸ Therefore, developing their own PLSAs aided local politicians in crafting an educational accountability and improvement narrative, decoupling from national decisions over which provincial authorities have little or no say.⁵⁹

Mendoza's testing resurgence (2019-2023): Between a learning crisis and global influences

Next to the Andes, the re-emergence of LSAs in Mendoza embodied the provincial response to a pressing learning crisis. Arguably, Mendoza's performance in national tests was not too different from other provinces. Rather than an external shock, the crisis was triggered by reaching a tipping point. Since 2016, literacy scores in Argentina's NLSAs concerned local authorities. Then, UNESCO's 2019 Regional Comparative and Exploratory Study portrayed a more worrisome picture, blowing off local alarms (Banco Mundial & Gobierno de Mendoza, 2023). Pressure to act grew as influential civil society organizations warned that "students do not understand what they read" (Argentinos por la Educación, 2021). Initially, low literacy rates led to an ambitious provincial literacy plan in 2017, followed by an innovative standardized assessment titled

⁵⁸ Interview with provincial authority

⁵⁹ Interviews with several national and provincial authorities.

“Reading Fluency Census” to monitor students’ reading skills starting in 2019 (Banco Mundial & Gobierno de Mendoza, 2023). The Reading Census occurs three times a year during the fall, winter, and spring. Each application round consists of a controlled sample led by external test-takers, used for policy decisions –i.e., updates on the provincial reading plan–, and a census assessment managed by school teachers. Additionally, coverage of school years grew over time. Reading Census results are uploaded to Mendoza’s EMIS platform (GEM+) and accompanied by a school report. The Census received World Bank endorsement as a best practice to recover from the pandemic, emulated by Buenos Aires since 2024.

In 2021, the national government’s intentions to cancel APRENDER led Mendoza’s authorities to deploy a new test, the “Provincial Learning Survey” (RPA in Spanish).⁶⁰ The assessment was built off existing test items in Mendoza’s EMIS platform, elaborated in close collaboration with Chilean evaluation agency experts.⁶¹ In a few months, the new assessment was readily available and deployed in each school through the EMIS. Aligned with a quality assurance view of educational monitoring and planning, the test covered a representative school sample, had no consequences, and never published its results. Like Córdoba, Mendoza officials aimed to complement their PLSAs with the national test APRENDER while stressing the importance of counting with their information independently from national turmoil or orientation turns. Indeed, education statistics and school information became another battleground between the National Ministry of Education and the provinces. While the national authorities required provincial data to build the national EMIS, it offered nothing in exchange for provinces with their systems like

⁶⁰ Interview with Provincial Education Authorities

⁶¹ Interview with Provincial Education Authorities

Mendoza or Córdoba.⁶² Put differently, Buenos Aires, Córdoba, and Mendoza received no benefit from transferring their data while risking political usage or funds cuts from the national government.⁴⁷

Lastly, as part of Mendoza's educational evaluation plan, the province joined PISA in 2022. For provincial authorities, PISA offered valuable data on different aspects of those monitored by national and provincial assessments, allowing a cross-national comparison with cases like Santiago de Chile, considered more similar to their reality than Buenos Aires.⁶³ More so, PISA results were included as a provincial benchmark in the literacy plan – students were expected to meet OECD standards by 2025.⁴⁸ For Mendoza's education leaders, more information and assessments made no harm but instead gave authorities valuable data to inform their decisions. Likewise, it granted them independence from the national landscape, crafting a political narrative with education as a policy priority.

Policy coalitions, cross-scalar dynamics, and the evolution of global reforms

The analysis of LSAs in Argentina helps us understand the role of the temporal and spatial dimensions jointly with institutional and political drivers behind the adoption and evolution of global scripts. An initial reflection signals that macro or global developments might be necessary but insufficient to explain diffusion (Béland et al., 2018; Verger et al., 2018; Marsh & Sharman, 2009). Across different transfer spaces, this study demonstrates the unequivocal role of policy coalitions in creating the conditions for borrowing or transfer (Steiner-Khamsi, 2021). Starting at

⁶² Interview with Provincial Education Authorities

⁶³ Interview with Provincial Education Authorities

the national level, Kirchners' administrations (2003-2015) seized the epochal chorus against neoliberalism in the early 2000s to resist ideas of performance-based accountability (Díaz Ríos, 2024). The government's feisty stance ahead of IOs was echoed by other southern cone countries like Brazil, Uruguay, Chile, and Venezuela. However, over the years, many such countries have shifted their posture on LSAs in light of a rising global educational accountability discourse (Pinkasz, 2022; Rivas & Sanchez, 2022; Termes & Pagès, 2024). Nonetheless, Kirchners' political alliance with the teacher unions rendered Argentina 'closed' to the hegemonic accountability discourse, blocking any attempt to increase testing centrality or consequentiality (Chambers-Ju, 2024). Therefore, NLSAs exhibited almost a ceremonial role during those years, with scarce engagement with the theory of change linking testing and educational improvement. Analogously, the government belittled PISA participation. Because abandoning national or international testing altogether would have generated widespread contention, the Kirchners' administration opted for minimizing their centrality in the educational agenda. However, this fed the confrontation with opposing coalitions in favor of more testing and accountability, particularly in Buenos Aires.

The cross-scalar and temporal effects derived from coalitions' adherence to different LSAs are depicted in Figure 2. The y-axis shows the three main scripts organized as a continuum, while LSAs' trajectories are plotted on the x-axis, taking different values according to the increase or decrease in their frequency, centrality, consequentiality, and mix with other accountability policies, such as indexes, rankings or ILSAs engagement. As discussed, when the national government dismissed accountability ideas, Buenos Aires politicians embraced them to distance themselves from Kirchners and to project themselves as performance-based accountability champions into the global arena.

In the Argentine context, Buenos Aires became the lead adopter of the global accountability discourse, spreading it countrywide (Rodríguez et al., 2018). The city transfer space was ruled by agents with different policy views and political aspirations holding sufficient governance autonomy to depart from the national agenda. Since 2010, Buenos Aires redefined its LSAs policy, emulating performance-based accountability ideas. As a good student, Buenos Aires' authorities transformed their qualitative-oriented LSAs into a traditional standardized test with annual frequency and universal coverage. Testing centrality included their new independent education evaluation unit, school reports, and quality indexes – pretty much IOs' full testing repertoire (Clarke & Luna-Bazaldua, 2021; Rodríguez et al., 2018). Despite higher-stakes features were dropped due to incompatibility with the institutional and political landscape, the city LSAs were clearly repurposed in conformity with performance-based accountability practices.

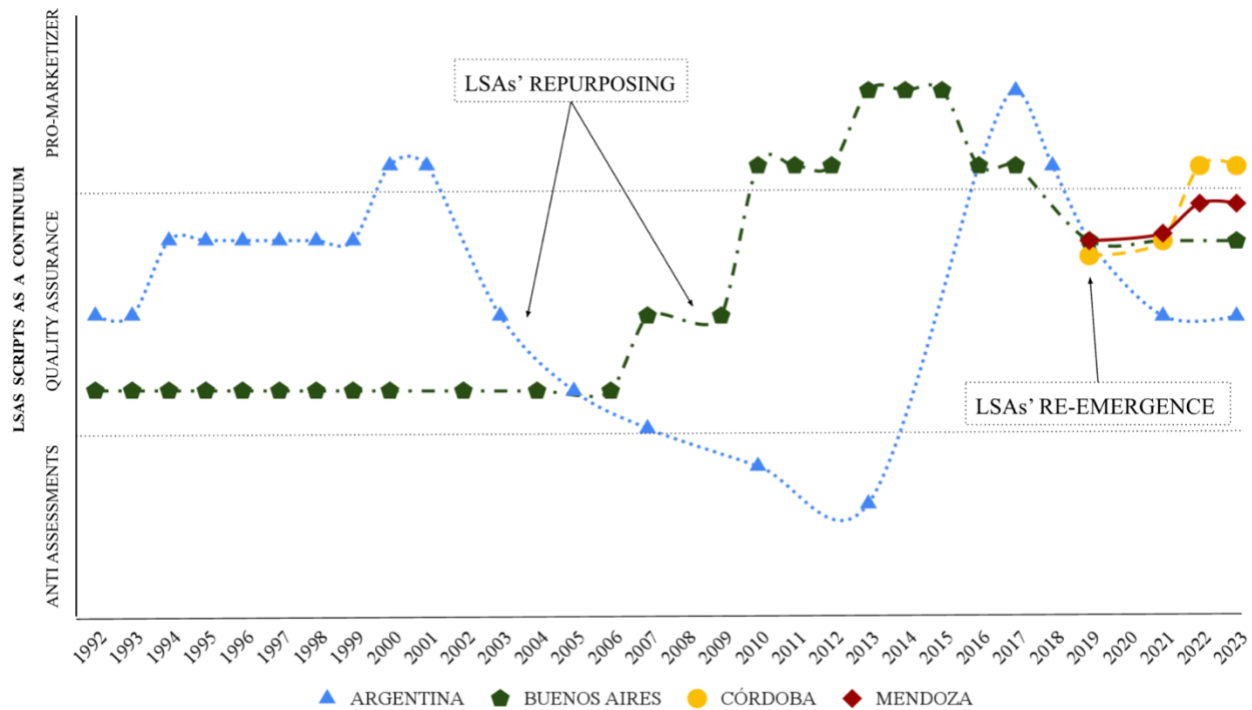
Buenos Aires' leadership on educational accountability was not limited to domestic assessments: it also covered the transnational sphere. The city pioneered joining PISA as an independent territory in 2012, and into the International Association for the Evaluation of Educational Achievement programs in 2013. This scale-jumping dynamic reveals not only the political independence of states in federal systems, but the boundless influencing aspirations of IOs (Beech et al., 2023). We have long acknowledged the mistake of constraining policy transfer to national frontiers. However, not only do local politicians seek benefits from global platforms like PISA, also IOs find ways to bypass national governments when they disregard their agendas due to political or ideological differences, enlarging IOs influencing capacity. On the one hand, PISA participation fed the city antagonism with the national government. On the other hand, it hid the personal ambitions of a key actor: in 2014, the head of the city's assessment unit became director of UNESCO's Institute of Statistics. Hence, in the rursposing of Buenos Aires assessment,

we can distinguish how the global, national, local, and even micro scales mutually influence each other when borrowing a global script.

At the national level, *Cambios*' disembarkment in December 2015 shifted the political climate shaking the education sector. If during Kirchner's terms educational accountability had little penetration, the new government made it the cornerstone of its educational agenda (Rodrigo & Moyano, 2023). At the national level, *Cambios*' scaled up the Buenos Aires assessment playbook. As shown in Figure 2, when *Cambios* simultaneously ruled the country and city of Buenos Aires between 2015 to 2019, their assessment policy was harmonized, and the political rivalry was de-escalated. The accountability discourse sought to echo OECD's views, as becoming a member was one of *Cambios*' policy priorities. All in all, Argentina's government gestures aimed to regain global legitimacy with Western powers, deemed lost due to Kirchner's alignment with BRICS countries – Brazil, Russia, India, China and South Africa.

Despite its several reform attempts, *Cambios* was more successful in its discursive shift over assessments and education quality than in materializing or consolidating the changes on the ground. Indeed, the 'refundational' view on assessment portrayed a symbolic break with Kirchner's policies while depicting CTERA as responsible for educational stagnation. Nonetheless, most of the intended reforms, such as publicizing schools' results, replicating an independent evaluation agency, or assessing teachers, failed due to insufficient political support. Likewise, the OECD rejected Argentina's candidacy early on due to its unstable economic situation. Nonetheless, *Cambios*' term (2015-2019) succeeded in instilling an educational accountability discourse nationwide, whose effects would transpire into future federal and provincial policies. For instance, the restitution of LSAs' centrality within the Argentine context left little room for the returning Peronist coalition (2019-2023) to overlook national and even

international assessments as in the past. Resisting harsher models of test-based accountability, the national government embraced educational assessment as a more comprehensive process combining data from standardized tests, school internal efficiency, and students' backgrounds. The shock produced by Covid-19 created a feeling of urgency, as students' attainment showed alarming signs of decline. The government's long-term school closure and hesitancy to resume the NLSAs in 2021 were a breaking point with provinces. Beyond Buenos Aires, Córdoba, and Mendoza, another handful of provinces adopted a new standardized assessment (Ministerio de Educación, 2023). After three decades of inconsistencies, interruptions, and shifting goals, one thing seems clear: the national government's legitimacy to articulate a federal system has weakened at the expense of provincial power, intensifying the effects of the 1990s decentralization reforms where provinces become increasingly autonomous (Falleti, 2010) and, consequently, school systems more heterogenous countrywide.



Source: self-elaboration.

Figure 2. The adoption, institutionalization and re-emergence of LSAs in Argentina

The reawakening of a global script: the non-continuity of the temporal dimension

The cases of Córdoba and Mendoza illustrate how time affects policy trajectories in non-linear ways: after decades of being dormant, national and provincial political developments reanimated the global testing script. In Córdoba, the assessment agenda shifted gradually along with changing epochal discourses. Starting with a pilot PLSAs in 2012, a mainstream standardized test was only consolidated in 2019. In this sense, Córdoba’s politicians’ testing ideas were shaped by the rebranding and repurposing of the national assessment APRENDER since 2016. The local coalition accommodated their beliefs to the nationally promoted assessment culture. Recognizing the political and policy advantages of assessment ownership, authorities also decided to join PISA in 2018. Arguably, Buenos Aires led the ground for other provinces to follow, producing a

horizontal contagion effect among provincial policy coalitions. Horizontal emulation is common in federal systems, particularly between transfer spaces which shared political aspirations or institutional norms (Beech et al., 2023; Steiner-Khamsi et al., 2024). Nonetheless, it is rare to find three independent samples in an impoverished and middle-sized country with a standing PISA participation since 2000. Put differently, Buenos Aires, Córdoba, and Mendoza's PISA participation seem heavily influenced by political aspiration rather than technical or policy rationales.

Returning to our prior point, the consolidation of Córdoba's PLSAs occurred after Covid-19. National hesitations and the cancelation attempt of APRENDER in 2021 fueled provincial coalitions' conviction for owning their assessment. Undoubtedly, testing rendered a legitimizing stamp and a powerful governance instrument to craft a political narrative around educational quality.

Lastly, Mendoza's re-adoption of standardized testing in 2019 also occurred in response to a shifting environment, demonstrating the discontinuous nature of the temporal dimension in policy reforms. Assessment and accountability ideas have gained ground in Argentina's ecosystem in recent years, making it more prone to global testing discourses. In this context, the combined effects of students' impoverished performance due to school closure with the potential suspension of APRENDER in 2021 paved the way for a new stance ahead of standardized assessments. In 2019, poor literacy results gave birth to Mendoza's reading census. Then, in 2021, announcements of NLSAs suspension triggered the adoption of the 'Provincial Learning Survey.' Once in effect, PLSAs delegitimized NLSAs' function for education governance, rendering national policies a distant matter for provinces' interests and concerns. Remarkably, the concatenated effects of different events at national and provincial level modify stakeholders' perception on assessments,

opening room for a new direction unthinkable years ago. In other words, our sequential analysis elucidates the often subtle environmental change and slow political build-up leading to policy transfer or policy change over time.

Policy variation through mechanistic processes

The detailed historical narrative of events and activities shaping the adoption and evolution of LSAs in Argentina set the stage for examining when, how, and why political action can produce policy change. Drawing from recent efforts in policy studies to unpack the policy mechanisms (Weible, 2018; Mukherjee & Howlett, 2015; Capano et al., 2019), the study has shown that LSAs trajectories in Argentina have been driven by policy coalition ability to, first, influence the public agenda and, secondly, overcome resistance raised by the proposed reforms. The main aspects of this two-phase mechanism, including the agenda-setting stage of policy reforms, and the opposition encountered depending on their compatibility or disruption related to existing legacies, are summarized in Table 3. Following Steiner-Khamsi et al. (2024), the table shows whether policies were adopted (A), partially adopted (P), suspended after initial adoption (S) or rejected (R).

Schematically, the process is repeated at the national and provincial levels. Driven by a new election or in response to an external shock (i.e., Covid-19), different coalitions raffled to secure access to decision-making positions. Then, coalitions followed a typical sequence for agenda setting: they engaged in problem-framing to create a need for reform (Kingdon, 2014; Jenkins-Smith et al., 2019), chose and tailored LSAs features to their goals, and sought to leverage political alliances to secure its adoption. For instance, after the 2001 crisis, Kirchners associated LSAs with IOs to delegitimize their use and centered on compensatory policies to the detriment of

tests, also rejected by unions and their political allies (Rivas & Sanchez, 2022). In contrast, Macri's administration exaggerated the PISA 2015 exclusion and exploited Kirchner's statistical scandals in favor of its pro-testing agenda, blaming unions for educational underperformance (Moyano & Rodrigo, 2018; Boschio, 2023).

Once on the public agenda, policies could follow two routes. If proposals were 'compatible' with existing norms and posed minor challenges to existing power structures, they would give little veto possibilities to opposing groups –i.e., unions, education NGOs, opposing political parties, etc.–. To illustrate, Kirchner's dilution of NLSA's centrality was criticized by local experts and World Bank members but endorsed by unions that were more politically relevant. Hence, opposing coalitions could do little against the progressive decrease of NLSAs' frequency, uses, and presence in the educational agenda. However, the government was not capable of withdrawing the NLSAs altogether or stopping participating in PISA, whose effects were deemed more problematic. This policy alternative would fall into our second pathway of 'incompatible policies.' Reforms that challenge existing norms and seek shifts in power relations among interest groups would likely face resistance from opposing stakeholders and, potentially, institutional veto points. The re-emergence of PLSAs exemplifies this second route. Initially, unions questioned new testing initiatives requiring a political effort from education authorities to persuade them. Although PLSAs meant a shift in testing logic, they did not create reputational or career stakes for teachers or students, as results were not publicly available. Furthermore, Buenos Aires laid the ground for Córdoba and Mendoza, who seized the city's example to justify the adoption. Arguably, Buenos Aires pioneer example played an essential role in the horizontal dissemination of PLSAs.

In contrast, certain policies are more heightened, such as the publication of school-level tests results. This project was rejected twice, first in Buenos Aires (2007-2015) and at the national

level (2015-2019), as each government failed to secure legislative support from opposing political parties. This failure is, nonetheless, far from definitive. Following Kingdon's (2014) words, pro-market advocates are like surfers waiting for the big wave. Not so long ago, Javier Milei's far-right presidency would have been unthinkable for most Argentinians. Upon his arrival in 2023, the country has undergone a large overhaul towards free markets and pro-privatization policies in almost every sector, including the adoption of school vouchers in 2024. Recently, a new project was filed in Congress to publicize test results.⁶⁴ Therefore, despite Argentina's education policy legacy seeming at odds with test-based accountability or school choice (Díaz Ríos, 2024; Ydesen et al., 2023), a significant shift in political power dynamics might break a weakening resistance.

Table 3: Policy Coalitions and Policy Streams

⁶⁴ María Eugenia Vidal, former Governor of the Province of Buenos Aires with a historical dispute with teachers' unions, introduced, once again, is pushing for enabling tests results publications. Project N. 0313-D-2024 Available at: <https://www4.hcdn.gob.ar/dependencias/dsecretaria/Periodo2024/PDF2024/TP2024/0313-D-2024.pdf>

Case	Dominant Coalition	Period	Policies in the public agenda	Links with Policy Legacies	Stakeholders Veto	Result	Type of policy change
Argentina	Anti Standardized Assessments	2003-2015	Reduced NLSAs centrality and intensity	Compatible	Few veto opportunities from opposing groups	Adoption	1st order Goals: unchanged: Quality assurance Instruments: unchanged Settings: minor
	Pro-Market / QA	2015-2019	Increased NLSAs centrality, coverage and intensity without high stakes	Compatible	Few veto opportunities from opposing groups	Adoption	1st order Goals: unchanged: Quality assurance Instruments: unchanged Settings: minor
			Publication of test results	Incompatible	Contested internally, opposed by unions and vetoed in Congress	Failed adoption	
			Independent Assessment Unit	Mostly Compatible	Lacked political support in Congress	Failed adoption	
			Teachers' evaluation	Mostly Compatible	Contested internally, opposed by unions	Partial adoption: Pilot test	
Quality Assurance	2019-2023	Stabilized NLSAs frequency & coverage;	Compatible	Agreed with provinces and unions	Adopted	1st order Goals unchanged: Quality assurance Instruments: unchanged Settings: minor	
Buenos Aires	Pro-Market / QA	2007-2015	PLSAs: annual, census, no-stakes	Mostly Compatible	Negotiated with unions	Adopted	2st order Goals: unchanged: Quality assurance Instruments: changed Settings: major
			Quality Indexes	Mostly Compatible	Veto to publicize results	Partial adoption: with changes to ideal version	
			Quality Assurance Agency (UEICEE)	Compatible	Few veto opportunities from opposing groups	Partial adoption with changes to ideal version	
			Participation in ILSAs	Compatible	Few veto opportunities from opposing groups	Adopted	

Córdoba	Quality Assurance	2007-2023	PLSAs: bi-annual, sample, no-stakes	Compatible	Negotiated with unions	Adopted	1st order Goals: unchanged: Quality assurance Instruments: unchanged Settings: minor
			Participation in ILSAs	Compatible	Few veto opportunities from opposing groups	Adopted	
Mendoza	Quality Assurance	2019-2023	PLSAs: bi-annual, sample, no-stakes	Compatible	Negotiated with unions	Adopted	1st order Goals: unchanged: Quality assurance Instruments: unchanged Settings: minor
			Participation in ILSAs	Compatible	Few veto opportunities from opposing groups	Adopted	

Note: A: adoption; P: Partial Adoption; F: Failed Adoption; S: Suspended after initial adoption

Conclusion

To circle back to the triggering point of this study, the Argentine case helped us look closer at policy transfer dynamics from a synchronic and diachronic perspective. The study has unpacked the ongoing influence of policy coalitions operating at different scales in the evolution of global reforms by steering at policy transfer as an ongoing phenomenon with different loops between local and global developments. From a distance, Argentina's national assessment seemed rather static over the years, nurturing claims of indifference due to its low-stakes features. However, this has proven far from true. The historical dispute of different policy coalitions animates an ongoing struggle to reshape assessment policies to serve political interests, not only domestically but also internationally. Likewise, the study has shown the continuous and discontinuous effects of the temporal dimension in policy reforms. Whereas 30 years after educational decentralization, the national state legitimacy weakens ahead of growing provincial autonomy, the sub-national reemergence of the global testing discourse occurred after decades of being dormant. These findings stress the importance of the temporal dimension in understanding policy transfer, opening new lines of research for examining how shifting timelines and resurgent policy discourses open avenues for educational reform.

References

- Aquilino, N., Suaya, A., Pomares, J., & Page, M. (2016). *The Deficits in Horizontal Accountability in Argentina: A Tale of two worlds* (Evidence and Lessons from Latin America) [Research Paper Series]. Centre for the Implementation of Public Policies that promotes Equity and Growth (CIPPEC). https://www.vippal.cippec.org/wp-content/uploads/2016/06/CIPPEC_DEFICITS.pdf
- Argentinos por la Educación. (2021). *#NoEntiendenLoQueLeen*. Argentinos por la Educación. <https://argentinosporlaeducacion.org/campana/campana-nacional-por-la-alfabetizacion/>
- Banco Mundial, & Gobierno de Mendoza. (2023). *De vuelta a lo básico: Evaluación de fluidez lectora y uso pedagógico de resultados para la recuperación de los aprendizajes. La Experiencia de la Provincia de Mendoza, Argentina*.
- Beach, D., & Pedersen, R. B. (2019). *Process-tracing methods: Foundations and guidelines* (Second Edition). University of Michigan Press.
- Beech, J., & Artopoulos, A. (2016). Interpreting the circulation of educational discourse across space: Searching for new vocabularies. *Globalisation, Societies and Education*, 14(2), 251–271. <https://doi.org/10.1080/14767724.2015.1025713>
- Béland, D., Foli, R., Howlett, M., Ramesh, M., & Woo, J. J. (2018). Instrument constituencies and transnational policy diffusion: The case of conditional cash transfers. *Review of International Political Economy*, 25(4), 463–482. <https://doi.org/10.1080/09692290.2018.1470548>
- Benson, D., & Jordan, A. (2011). What have we Learned from Policy Transfer Research? Dolowitz and Marsh Revisited. *Political Studies Review*, 9(3), 366–378. <https://doi.org/10.1111/j.1478-9302.2011.00240.x>

- Benveniste, L. (2002). The Political Structuration of Assessment: Negotiating State Power and Legitimacy. *Comparative Education Review*, 46, 89–118.
- Bocchio, M. C., Grinberg, S. M., & Villagran, C. A. (2023). La evaluación como política educativa en la sociedad managerial. Control externo, flexibilización normativa e in/exclusión. *Jornal de Políticas Educacionais*, 17(2).
<https://doi.org/10.5380/jpe.v17i2.93693>
- Bonal, X., Pagès, M., Verger, A., & Zancajo, A. (2023). Regional policy trajectories in the Spanish education system: Different uses of relative autonomy. *Education Policy Analysis Archives*, 31. <https://doi.org/10.14507/epaa.31.8031>
- Braginski, R. (2018, March 2). Macri pidió un “ranking” de escuelas y desató una polémica. *Clarín*. https://www.clarin.com/sociedad/macri-pidio-ranking-escuelas-desato-polemica_0_BJVXhM8_G.html
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Buenos Aires Ciudad. (2013). *Proyecto Jurisdiccional de Evaluación de finalización de Nivel Medio (2013)*. Dirección General de Evaluación de la Calidad Educativa.
<https://buenosaires.gob.ar/sites/default/files/media/document/2016/04/22/20450f08c51490574147688473017de94ff4bf8c.pdf>
- Capano, G., Howlett, M., Ramesh, M., & Virani, A. (Eds.). (2019). *Making Policies Work: First- and Second-order Mechanisms in Policy Design*. Edward Elgar Publishing.
<https://doi.org/10.4337/9781788118194>
- Chambers-Ju, C. (2024). *Mobilizing Teachers: Education Politics and the New Labor Movement in Latin America*. Cambridge University Press.

- <https://www.cambridge.org/core/books/mobilizing-teachers/3FE82A72A4EB7D78B9B60BF6940D074E>
- Clarín.com. (2016, July 12). *Pruebas PISA: el ministro Bullrich defendió el entrenamiento de los alumnos*. https://www.clarin.com/sociedad/pruebas-pisa-bullrich-defendio-entrenamiento_0_HJz3ZYS7g.html
- Clarke, M., & Luna-Bazaldúa, D. (2021). *Primer on Large-Scale Assessments of Educational Achievement*. World Bank. doi:10.1596/978-1-4648-1659-8.
- Collier, D. (2011). Understanding Process Tracing. *PS: Political Science & Politics*, 44(4), 823–830. <https://doi.org/10.1017/S1049096511001429>
- Derqui, J. M. (2001). Educational decentralization policies in Argentina and Brazil: Exploring the new trends. *Journal of Education Policy*, 16(6), 561–583. <https://doi.org/10.1080/02680930110087825>
- Díaz Ríos, C. M. (2024). *Translating Global Ideas: How Policy Legacies and Domestic Politics Shape Education Governance in Latin America*. SUNY Press. <https://sunypress.edu/Books/T/Translating-Global-Ideas2>
- Diker, G., Briscioli, B., & Stoppani, N. (2023). La institucionalización de la evaluación educativa en Argentina: Disputas por sus sentidos y usos. *Jornal de Políticas Educacionais*, 17(2). <https://doi.org/10.5380/jpe.v17i2.92216>
- DINIECE. (2003). *Plan de evaluación de la calidad educativa 2003-2007* [Government document]. Dirección Nacional de Información y Evaluación de la Calidad Educativa. Ministerio de Educación de la Nación. República Argentina. <http://repositorio.educacion.gov.ar/dspace/handle/123456789/93137>
- Esper, T. (2024). Changing paradigms: A historical analysis of school autonomy and

- accountability policies in Colombia. *Nordic Journal of Studies in Educational Policy*, 0(0), 1–15. <https://doi.org/10.1080/20020317.2024.2396555>
- Falabella, A., & Ramos Zincke, C. (2019). La larga historia de las evaluaciones nacionales a nivel escolar en Chile. *Cuadernos Chilenos de Historia de la Educación; Núm. 11* (2019); 66-98. <http://revistaschilenas.uchile.cl/handle/2250/154892>
- Falleti, T. G. (2010). *Decentralization and Subnational Politics in Latin America*. Cambridge University Press.
- Feldfeber, M., Duhalde, M., Puigrós, A., & Robertson, S. (2018). *La privatización educativa en Argentina*. (1ra Edición). Education International - CTERA.
- Furuta, J. (2022). The Rationalization of “Education for All”: The Worldwide Rise of National Assessments, 1960–2011. *Comparative Education Review*, 66(2), 228–252. <https://doi.org/10.1086/718829>
- Ganimian, A., J. (2015). *El termómetro educativo: Informe sobre el desempeño de Argentina en los Operativos Nacionales de Evaluación (ONE) 2005-2013*. Proyecto Educar 2050.
- Gobierno de Córdoba. (2013). *Documento para la construcción de la Autoevaluación Institucional*. Dirección General de Planeamiento, Información y Evaluación Educativa. Ministerio de Educación. <https://www.igualdadycalidadcba.gov.ar/SIPEC-CBA/publicaciones/PNFP/Auto%20evaluacion%20Institucional.pdf>
- Grek, S., Maroy, C., & Verger, A. (2020). Introduction: Accountability and datafication in education: Historical, transnational and conceptual perspectives. In S. Grek, C. Maroy, & A. Verger (Eds.), *World Yearbook of Education 2021: Accountability and Datafication in the Governance of Education*. Routledge. <https://doi.org/10.4324/9781003014164>
- Hall, P. A. (2013). Tracing the Progress of Process Tracing. *European Political Science*, 12(1),

- 20–30. <https://doi.org/10.1057/eps.2012.6>
- Henry, A. D., Ingold, K., Nohrstedt, D., & Weible, C. M. (2014). Policy Change in Comparative Contexts: Applying the Advocacy Coalition Framework Outside of Western Europe and North America. *Journal of Comparative Policy Analysis: Research and Practice*, *16*(4), 299–312. <https://doi.org/10.1080/13876988.2014.941200>
- Henry, A. D., Ingold, K., Nohrstedt, D., & Weible, C. M. (2022). Advocacy Coalition Framework: Advice on Applications and Methods. In *Methods of the Policy Process*. Routledge.
- Hossain, M. (2023). Large-scale data gathering: Exploring World Bank’s influence on national learning assessments in LMICs. *International Journal of Educational Development*, *102*, 102877. <https://doi.org/10.1016/j.ijedudev.2023.102877>
- Instituto de Investigaciones Psicológicas. (2021, October 21). *Prueba Provincial de Logros de Aprendizaje PRISMA*. Universidad Nacional de Córdoba.
<https://iipsi.psicologia.unc.edu.ar/prueba-provincial-de-logros-de-aprendizaje-prisma/>
- Jenkins-Smith, H. C., Nohrstedt, D., & Weible, C. M. (2019). *The Advocacy Coalition Framework: An Overview of the Research Program* (pp. 92–114).
- Kamens, D. H., & Benavot, A. (2011). National, regional and international learning assessments: Trends among developing countries, 1960–2009. *Globalisation, Societies and Education*, *9*(2), 285–300. <https://doi.org/10.1080/14767724.2011.577337>
- Kingdon, J. W. (2014). Agendas, Alternatives, and Public Policy. In *Political Science Quarterly* (Second Edition, Vol. 100, Issue 1). <https://doi.org/10.2307/2150882>
- Llorente, A. (2016, February 22). ¿Por qué es tan difícil que Argentina vuelva a tener un índice de inflación confiable? *BBC News Mundo*.

- https://www.bbc.com/mundo/noticias/2016/02/160218_argentina_inflacion_indec_all
- Löblová, O. (2018). When Epistemic Communities Fail: Exploring the Mechanism of Policy Influence. *Policy Studies Journal*, 46(1), 160–189. <https://doi.org/10.1111/psj.12213>
- López Leavy. (2024). *Unpacking The Influence of Large-Scale Learning Assessments Data on Education Policy Formulation in Argentina and the City of Buenos Aires*. European Conference of Educational Research, Nicosia. <https://eera-ecer.de/ecer-programmes/conference/29/contribution/59026>
- Mahoney, J. (2010). After KKV: The New Methodology of Qualitative Research. *World Politics*, 62(1), 120–147. <https://doi.org/10.1017/S0043887109990220>
- Mahoney, J., & Thelen, K. (2010). A Theory of Gradual Institutional Change. In J. Mahoney & K. Thelen (Eds.), *Explaining Institutional Change: Ambiguity, Agency, and Power* (pp. 1–37). Cambridge University Press. <https://doi.org/10.1017/CBO9780511806414.003>
- Marsh, D., & Sharman, J. C. (2009). Policy diffusion and policy transfer. *Policy Studies*, 30(3), 269–288. <https://doi.org/10.1080/01442870902863851>
- Matovich, I., & Esper, T. (2023). Following New Philanthropy by Network Ethnography: How Did the Varkey Foundation Land and Expand in Latin America? *ECNU Review of Education*, 20965311231168422. <https://doi.org/10.1177/20965311231168422>
- McCann, E., & Ward, K. (2012). Policy Assemblages, Mobilities and Mutations: Toward a Multidisciplinary Conversation. *Political Studies Review*, 10(3), 325–332. <https://doi.org/10.1111/j.1478-9302.2012.00276.x>
- Ministerio de Educación de la Nación. (2023). *Dispositivos de evaluación de aprendizajes de estudiantes: Experiencias jurisdiccionales*. Ministerio de Educación de la Nación. https://www.argentina.gob.ar/sites/default/files/experiencias_jurisdiccionales_2.pdf

- Moreno Salto, I. (2023). *Harmony and Cacophony in Large-scale Assessments in Education: A Mexican Account*. Brill. <https://brill.com/display/title/69042>
- Moschetti, M. C., Nistal, M., & Sáenz Guillén, L. (2023). *Subvenciones a la educación privada: ¿para qué y para quiénes? Argentinos por la Educación*.
https://argentinosporlaeducacion.org/wp-content/uploads/2023/11/Subvenciones-a-la-educacion-privada_-%C2%BFpara-que-y-para-quienes_.pdf
- Moyano, I. R., & Rodrigo, L. (2022). La evaluación de la calidad de la educación en Argentina. Tendencias y trayectoria de la política educativa nacional (1990-2022). *Revista Latinoamericana de Políticas y Administración de la Educación*, 17, Article 17.
- Mukherjee, I., & Howlett, M. (2015). Who Is a Stream? Epistemic Communities, Instrumental Constituencies and Advocacy Coalitions in Public Policy-Making. *Politics and Governance*, 3(2), 65–75.
- Narodowski, M., Gottau, V., & Moschetti, M. (2016). Quasi-State monopoly of the education system and socio-economic segregation in Argentina. *Policy Futures in Education*, 14(6), 687–700. <https://doi.org/10.1177/1478210316645016>
- Nullmeier, F., & Kuhlmann, J. (2022). Introduction: A Mechanism-Based Approach to Social Policy Research. In J. Kuhlmann & F. Nullmeier (Eds.), *Causal Mechanisms in the Global Development of Social Policies* (pp. 3–29). Springer International Publishing. https://doi.org/10.1007/978-3-030-91088-4_1
- Obarrio, M. (2003, May 28). *Kirchner destrabó el conflicto docente*.
<https://www.lanacion.com.ar/politica/kirchner-destrabo-el-conflicto-docente-nid499419/>
- Oliveira, D. A., & Feldfeber, M. (2024). Unionists, teachers, and academics resisting neoliberal reforms: Building a network as a third space. *Globalisation, Societies and Education*,

- 0(0), 1–13. <https://doi.org/10.1080/14767724.2024.2396294>
- Parcerisa, L., Verger, A., Martín, M. P., & Browes, N. (2022). Teacher autonomy in the age of performance-based accountability: A review based on teaching profession regulatory models (2017-2020). *Education Policy Analysis Archives*, 30, (100)-(100).
<https://doi.org/10.14507/epaa.30.6204>
- Pinkasz, D. (2022). *The use of learning assessment test results in education policy planning* [Foro Regional de Política Educativa, 5th, online, 2021]. UNESCO IIEP Buenos Aires. Oficina para América Latina. https://unesdoc.unesco.org/ark:/48223/pf0000382824_eng
- Porto de Oliveira, O., & Osorio Gonnet, C. (2023). Comparative Transfer Analysis: Policy Instruments, Space and Time. *Journal of Comparative Policy Analysis: Research and Practice*, 25(4), 367–384. <https://doi.org/10.1080/13876988.2023.2235308>
- Rhoten, D. (2000). Education decentralization in Argentina: A ‘global-local conditions of possibility’ approach to state, market, and society change. *Journal of Education Policy*, 15(6), 593–619. <https://doi.org/10.1080/02680930010000218>
- Rivas, A. (Ed.). (2021). *Examining Educational Policy in Latin America: Comprehensive Insights into Contemporary Reform*. Routledge.
- Rivas, A., & Sanchez, B. (2022). Race to the classroom: The governance turn in Latin American education. The emerging era of accountability, control and prescribed curriculum. *Compare: A Journal of Comparative and International Education*, 52(2), 250–268.
<https://doi.org/10.1080/03057925.2020.1756745>
- Rodríguez, L. R., Vior, S. E., & Más Rocha, S. M. (2018). Las Políticas de Evaluación de la Calidad Educativa en Argentina (2016-2018). *Educação & Realidade*, 43, 1405–1428.
<https://doi.org/10.1590/2175-623684907>

- Sabatier, P. A., & Weible, C. M. (2007). The Advocacy Coalition Framework: Innovations and Clarifications. In *Theories of the Policy Process, Second Edition* (2nd ed.). Routledge.
- Saldaña, J. (2021). *The Coding Manual for Qualitative Researchers* (Fourth Edition). SAGE Publications.
- Smith, W. C. (2016). *The Global testing culture: Shaping education policy, perceptions, and practice*. Symposium Books.
- Steiner-Khamsi, G. (2021). Externalisation and structural coupling: Applications in comparative policy studies in education. *European Educational Research Journal*, 20(6), 806–820. <https://doi.org/10.1177/1474904120988394>
- Steiner-Khamsi, G., Appius, S., & Nägeli, A. (2024). Comparing two transfer spaces over time and against a global script: The case of school-autonomy-with-accountability. *Journal of Education Policy*, 0(0), 1–21. <https://doi.org/10.1080/02680939.2024.2394594>
- Termes, A., & Pagès, M. (2024). *School accountability policies moving across scales: A comparative case study in decentralized educational systems*. <https://bristoluniversitypressdigital.com/edcollchap/book/9781447368045/ch003.xml>
- The Guardia. (2014, May 6). OECD and Pisa tests are damaging education worldwide—Academics. *The Guardian*. <https://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics>
- Vallejos, S. (2021, June 16). Pruebas Aprender: El operativo no se hará este año y se desplegarán “otras evaluaciones estandarizadas más eficientes”, dijo Trotta. *La Nación*. <https://www.lanacion.com.ar/sociedad/pruebas-aprender-el-operativo-no-se-realizara-este-ano-y-se-desplegaran-otras-evaluaciones-nid16062021/>
- Verger, A., Ferrer-Esteban, G., & Fontdevila, C. (2024). Not everything that moves must

- converge: Evidence from global policy and practice on performance-based accountability. In D. Brent Edwards, A. Verger, M. McKenzie, & K. Takayama (Eds.), *Researching Global Education Policy: Diverse Approaches to Policy Movement* (pp. 39–69). Policy Press. <https://doi.org/10.56687/9781447368045-006>
- Verger, A., Fontdevila, C., & Parcerisa, L. (2019). Reforming governance through policy instruments: How and to what extent standards, tests and accountability in education spread worldwide. *Discourse: Studies in the Cultural Politics of Education*, 40(2), 1–23. <https://doi.org/10.1080/01596306.2019.1569882>
- Verger, A., Novelli, M., & Altinyelken, H. K. (2018). Global Education Policy and International Development: A Revisited Introduction. In M. Novelli & H. K. Altinyelken (Eds.), *Global Education Policy and International Development: New Agendas, Issues and Policies* (2nd Edition, pp. 1–34). Bloomsbury. https://www.academia.edu/30394205/Global_Education_Policy_and_International_Development_New_Agendas_Issues_and_Policies_COMPLETE_BOOK_2012_version_
- Weible, C. M. (2018). Instrument Constituencies and the Advocacy Coalition Framework: An essay on the comparisons, opportunities, and intersections. *Policy and Society*, 37(1), 59–73. <https://doi.org/10.1080/14494035.2018.1417705>
- Wimmer, A. (2021). Domains of diffusion: How culture and institutions travel around the world and with what consequences. *American Journal of Sociology*, 126(6), 1389–1438. <https://doi.org/10.1086/714273>
- Wimmer, A., & Glick Schiller, N. (2003). Methodological nationalism, the social sciences, and the study of migration: An essay in historical epistemology. *International Migration Review*, 37(3), 576–610. <https://doi.org/10.1111/j.1747-7379.2003.tb00151.x>

Ydesen, C., Milner, A. L., Aderet-German, T., Gomez Caride, E., & Ruan, Y. (2023).

Educational Assessment and Inclusive Education: Paradoxes, Perspectives and Potentialities. Palgrave Macmillan.

5. Conclusion and agenda for further research

In light of the three articles, this conclusion section seeks to connect and jointly discuss their main findings, highlighting their combined theoretical and empirical contributions to educational research. I start by briefly discussing how countries' institutional frameworks, the dominant ideas in their educational space, and the leverage of political coalitions and interest groups have determined and mediated the adoption and trajectory of SAWA reforms. Following that, I explain the divergent influence of global factors and actors in spreading SAWA reforms in Argentina and Colombia. Next, I look into the four elements presented in the introduction –autonomy, standardization & accountability, competition & marketization, and performance incentives– and their associated policies to account for which ones were adopted, institutionalized, abandoned, or rejected in each country. Finally, I outline this dissertation's contribution to the study of policy transfer and policy change and introduce my future research agenda.

5.1. Comparing across key mediating factors: institutions, ideas, and coalitions

Throughout the articles, I examined policy transfer and policy change considering three key elements: ideas, institutions, and interest groups organized into policy coalitions. These three elements are privileged mediators of policy change ahead of influential factors, such as external pressures from IOs, shocks derived from economic or political crises or new societal problems. Taken together, ideas, institutions, and interest groups are potent drivers of political action and policy change (Béland, 2010; Hall, 2010). Starting with ideas, the study sheds light on the underpinning value systems and the theory of change of this global education reform. On the one hand, I fleshed out SAWA theory of change to achieve its promised educational improvement goals by dissecting its composing elements and interlocking policies (see Figure 1 in Introduction).

On the other hand, I demonstrated who embodied SAWA tenets, how they affect power relationships among different actors and interest groups, and in what ways ideas were re-interpreted across different localities.

On the institutional front, my analysis focused on three aspects. Initially, I mapped the institutional landscape of Argentina and Colombia, outlining both regulations and legislation, as well as the unwritten norms and traditions that shape these systems. A notable example was Argentina's institutional ethos around universal and public education, epitomized in its past teaching state. This historical legacy performs as a barricade to neoliberal reforms (Ydesen et al., 2023). Secondly, I examined the constraining or enabling effects of institutional norms on policy agents. This included identifying the beneficiaries and disadvantaged groups of existing policies, crucial for understanding how actors align to uphold or challenge the status quo, or countries' institutional ties with IOs as potential avenues for policy influence (Díaz Ríos, 2024; Meyer et al., 1997). Thirdly, I explored the pathways towards institutional change through the mechanisms of policy layering, drift, and conversion as more subtle and slower ways yet largely influential ways of producing significant change (Mahoney & Thelen, 2010).

Finally, I investigated the role of agents, assembled across policy coalitions, as vehicles to advance their interests and produce change. Building off the Advocacy Coalitions Framework (Jenkins-Smith et al., 2019; Sabatier & Weible, 2007), the coalition approach sheds light on the power and economic and political disputes among interest groups influencing policy trajectories. Furthermore, building off the coalition lens, I identified some of the more salient personages who championed SAWA reforms across different cases. Therefore, the following sections recapitulate how institutional, ideational, and coalitional forces have structured SAWA policies in Argentina and Colombia.

5.2. Global determinants of education reform: the pressure and influence of International Organizations

Since the Bretton Woods era, the expanding network of IOs has shaped a cohesive global policy agenda, wielding considerable influence over national policies. While the main drivers and mechanisms of IOs influence have been addressed earlier on, their uneven effects across countries required deeper analysis. When are IO mechanisms more likely to produce policy transfer? How do IOs determine their influencing strategy in each context? To what extent do countries debt levels, institutional ties to IOs, or ideological alignment affect their susceptibility to external pressures? These questions often mirror the chicken-and-egg dilemma: do countries turn to IOs purely out of necessity, or do politicians leverage this relationship to push through otherwise contested reforms? The answer likely lies in a mix of these factors, having different emphases depending on each case.

Hence, I contend that the World Bank and the OECD ability to promote SAWA reforms in Argentina was more limited than in Colombia for two reasons. First, since the 2000s, Argentina has been governed mainly by coalitions ideologically averse to SAWA tenets. Second, Argentina faced weaker structural and normative pressures to borrow them given its limited IO exposure compared to Colombia. When selecting Argentina and Colombia, I departed from two structurally similar cases that diverged in the adoption and trajectory of the SAWA reform package. Nonetheless, present similarities do not mean historical ones. As an upper middle-income country, Argentina has always received minuscule official development assistance money.⁶⁵ Among other

⁶⁵ However, it has had more than 21 agreements with the International Monetary Fund. See (BCRA, 2023). *History of Argentine Agreements with the IMF*. Available at: https://www.bcra.gob.ar/Institucional/Historial_Acuerdos_FMI.asp

consequences, IOs like the World Bank have had a minor influence in shaping Argentina's education sector, except for concrete time periods, such as the 1990s. In contrast, Colombia faced slower development during the 20th century, marked by significant social disparities and a 50-year armed conflict. Colombia has a tight relationship with the World Bank, key in swaying secondary and vocational education policies since the 1960s (Díaz Ríos & Urbano-Canal, 2023).

As discussed earlier, Argentina's proximity to IOs' policy recipes, especially neoliberal ones, was shaped by its government's ideological orientation. However, on top of ideas, countries face varying degrees of freedom to follow or dismiss IOs policies depending on several institutional factors, such as debt levels or institutional membership (Hossain, 2022, Pal, 2014). Member-states commit to adhere to several regulations, not only during their accession process but also afterward. Let's now circle back to our two cases to discuss how the government's ideology, political interests, and institutional constraints explain the heterogeneity in SAWA adoption and evolution.

5.2.1. Argentina's pendular stance ahead International Organizations

During the 1990s, Argentina closely followed the Washington Consensus policies, leading to governance decentralization, expenditure cuts, and, in education, the adoption of large-scale evaluation systems (Burki & Perry, 1998). In particular, the World Bank funded Argentina's NLSAs. Nonetheless, Argentina's ideological proximity to IOs' agenda was inseparable from harder determinants or conditions in the so-called 'convertibility plan.' In 1991, Argentina enforced a radical fiscal and economic reform around exchange-rate parity between US dollars and pesos to cut down hyperinflation. The convertibility plan, as known, involved privatizing public companies –water, electricity, oil, airline, communications, and so forth–, massive borrowing from international lenders, and adherence to neoliberal principles. In 2000, Argentina

began participating in PISA, becoming one of the earliest Latin American countries to do so. Aligned with its predecessor ideas, Fernando de la Rúa (1999-2001) remained close to neoliberal policies, both on the economic and educational front. During de la Rúa's presidency, the Minister of Education intended to use the national assessment with certification purposes at the end of secondary education and create an independent quality evaluation agency like in Chile (Benveniste, 2002; Díaz Ríos, 2024). However, these reforms sank with the president's desertion after the December 2001 crisis.

Following the crisis, Nestor (2003-2007) and Cristina Kirchner's left-leaning administrations (2007-2015) performed a symbolic rivalry with IOs recipes. Argentinians associated IOs policies as partially responsible for the 2001 crisis, making them an easy target of criticism. More so, Kirchners international agenda aimed to re-ignite Argentina's third-stance ahead of Western and Communist poles, now rehearsing an equilibrium between them and China. As 2001 got further away, the country's poor educational performance in PISA fed domestic scandals and criticism from opposing groups, putting pressure for educational reforms. Nonetheless, Argentina's government kept refusing external influences advocating for performance-based accountability reforms in the education sector.

In the following years, Mauricio Macri's administration (2015-2019) would change, once again, Argentina's stance ahead of IOs. Perhaps the most relevant sign was the government's failed attempt to join the OECD. Mauricio Macri aimed to distance its administration from the past Peronist governments, claiming Argentina's re-entrance into the global arena. During this time, ideas of educational accountability were brought to the forefront of the public agenda along with other OECD economic requirements. Nonetheless, Argentina was rapidly dismissed as a member due to its ongoing economic turmoil (Carrió, 2019), while most accountability reforms were

unsuccessful. Interestingly, after the Peronist return in 2019, the country officially started a new accession process with the OECD in 2022 (OECD, 2024). This shift in the country's stance opens new questions about future policy pathways once the membership process has advanced.

Regarding the second driver of IO influence, institutional constraints, I argue that Argentina conceded limited options for the OECD or the World Bank to push for SAWA policies. Most World Bank projects in Argentina focus on infrastructure, such as building schools or developing Education Management Information Systems (EMIS). Additionally, the OECD's early rejection of Argentina's membership bid curtailed further avenues for influence. Despite Argentina's early PISA participation, most national administrations have tended to ignore it due to poor results (Steiner-Khamsi, 2003). In sum, whereas performance-based accountability ideas have not been entirely disregarded, their impact over Argentina's education policy has remained limited.

5.2.2. The rising role of the OECD in Colombia

Throughout history, Colombia's relationship with IOs, notably the World Bank, has differed from Argentina's. The World Bank played a key part in Colombia via loans and technical assistance programs for vocational education and secondary education reforms (Caviedes, 2018; Cruz, 2012; Díaz Ríos & Urbano-Canal, 2023). The ties are so close that a specialized Education Minister unit manages the country's projects with the Bank. Nonetheless, rather than the World Bank, Colombia's ties with the OECD have been central to understanding the evolution of SAWA policies.

Throughout the 1990s, both countries adhered to the Washington consensus policies, including decentralization, public spending cuts to education, and establishing an 'evaluatory

state.’ Exemplars of these policies in Colombia have been César Gaviria’s (1990-1994) *‘Plan de Apertura Educativa’* (1991), fostering state restructuring along decentralization, privatization, and contract-out of educational services (Edwards et al., 2023). Interestingly, although included in the 1994 General Education Law, the consolidation of a national testing system began almost a decade later than in Argentina. This is surprising, as most South American countries developed their testing systems in the mid-1990s (shown in Figure 2 in Introduction). If Buenos Aires pioneered accountability and testing reforms in Argentina in the 2010s, the same happened in the late 1990s with Bogotá for Colombia. During Cecilia María Vélez tenure as Bogotá’s education secretary, the city endorsed managerial ideas of performance management, standardization, and incentives to reform its education system. Vélez scaled up her agenda as National Minister of Education (2002-2010). During her tenure, Colombia received substantial financial support from the World Bank and the Inter-American Development Bank to materialize the required EMIS and national testing system for distance monitoring (Díaz Ríos & Urbano-Canal, 2023).

Undoubtedly, the process of Colombia’s accession to the OECD marked a historical turning point. As discussed in Article 2, OECD membership signified a symbolic closure for decades of violence and economic instability and the start of a new era for the country. Colombia’s accession (2011-2018) was signed by President Santos’ (2010-2018) political ambition and the peace negotiation agreement with armed groups, rubricated in 2018. However, entering the OECD also had an impact on its education system. Unlike other policy sectors, the OECD does not mandate significant reforms in the field of education for its new members. In fact, Colombia became the first country whose educational system was evaluated as part of its admission process. OECD’s rise as a key reference in Colombia’s education was evidenced by its exponential growth in policy documents’ citations, confirmed by high-ranking officials’ views during interviews.

Likewise, Colombia's participation in PISA and its use as a national benchmark of education quality further underscore the normative and symbolic value that the country places on the OECD.

Beyond a normative role, Colombia's ties to the OECD offer insight into how lesson drawing and policy brokerage operates. Contrary to purely rationalist perspectives on lesson-drawing, the study reveals that decision-makers are influenced not only by their beliefs and preferences but also by their context (Meseguer, 2004). For instance, in its analysis of Colombia's accountability system, the OECD discouraged officials from increasing evaluation stakes (OECD, 2016). In its recommendation, it used Mexico as an example, highlighting potential undesirable consequences that could arise from the assessment. OECD's advice aligned Colombian decision makers preferences and also of other influential actors, such as FECODE, favoring its adoption.

In contrast, the OECD urged Colombia to introduce a national curriculum, arguing that its absence exacerbated inequities in access to knowledge across the country (OECD, 2016). Here, despite an alignment of ideology and preferences between OECD members and their Colombian counterparts, the national curriculum remained a highly contentious issue in Colombia. A national curriculum is resisted by FECODE due to its constraints on teacher autonomy, or because of the tensions around the recent history of the country's armed conflict. Thus, institutional and political constraints hindered the implementation of the OECD's advice. In summary, the ideological and institutional proximity between Colombia and the OECD is pivotal for understanding the trajectory of autonomy and accountability reforms in the country since the 2000s, as well as for elucidating its divergences compared to Argentina.

5.2.3. Looking beyond the nation-state: cross-scalar dynamics of global policies circulation.

The Argentine case has shown us the importance of analyzing global dynamics from a cross-scalar perspective. While IOs' influence diluted in the national space, it thrived sub-nationally. Within sub-national spaces, IO membership is 'off the table.' However, politically and administrative autonomous states can engage with IOs in multiple ways, from borrowing funds and requiring technical assistance to participating in global initiatives. As presented earlier, Buenos Aires politicians used global reform ideas ahead of domestic constituents to present themselves antagonistically to the national Peronist administration. At the time (2007-2015), Mauricio Macri was Buenos Aires' mayor. During his tenure, Buenos Aires implemented most of the performance-based accountability policies and joined PISA as an independent jurisdiction in 2015. Beyond Macri's political aspirations, other government members, such as Silvia Montoya, also pursued personal interests in their advocacy for performance-based accountability policies. Montoya used her Buenos Aires role as a platform for reaching global positions, becoming head of the UNESCO Institute for Statistics in 2015.

Later, the provinces of Córdoba and Mendoza emulated Buenos Aires and joined PISA in the 2018 and 2022 cycles, respectively. Undoubtedly, none of the three jurisdictions was subject to institutional pressures to participate in PISA nor to adopt standardized assessments as done. Arguably, the ideological proximity to global testing ideas was crucial in explaining sub-national adherence. Domestic politicians sought to gain legitimacy from participating in PISA, which can act as a projection screen for gaining visibility in front of domestic constituents.

In the case of Mendoza, PISA participation also fed a comparison appetite with cities rendered more culturally, contextually, and economically similar. For Mendoza's officials,

comparing their performance with that of Santiago de Chile was more relevant than to Buenos Aires. Likewise, Mendoza's provincial assessment initiatives received the World Bank's endorsement (Banco Mundial & Gobierno de Mendoza, 2023). In this case, the alignment between the province and the World Bank's agenda rendered a close collaboration in testing development, including external financing. Therefore, as argued, local echo acted as a trigger for IOs' engagement in policy transfer dynamics. In sum, Argentina's sub-national ties with global reform ideas confirm that IOs' influence is not limited to the national space and that policy transfer can be 'bounded blind', not distinguishing between administrative or territorial delimitations.

5.3. The characteristics and effects of governance decentralization in education

Education decentralization has been at the core of policy reforms in Latin America since the 1990s. Among its multiple consequences, the most relevant is power redistribution across different government levels — the nation-state, provincial governors, and municipalities. However, as Falletti (2010) has shown, decentralization does not always render stronger sub-national units ahead of national ones. Scholars have emphasized the significance of the sequence in which fiscal, administrative, and political decentralization occurs, as it shapes power hierarchy across governmental scales (Falletti, 2010; Edwards et al., 2023). To summarize their argument, when administrative decentralization happens first, the national government will withhold larger political authority in subsequent waves – as occurred in Argentina (Falletti, 2010; Rothen, 2000). On the contrary, if political decentralization occurs first, this will be followed by fiscal and administrative, resulting in increased power at local levels, as in Colombia (Azóca & Giraldo, 1994; Edwards et al., 2023). Building on this argument, I argue that the initial effect on

intergovernmental power balance can be substantially reshaped by later institutional changes through layering, conversion, and drift mechanisms.

5.3.1. The shifting dynamics of educational decentralization in Argentina

In Argentina, decentralization influence on school governance has been far from straightforward. After the 1990s reforms, the national government retained substantial decision-making authority because political decentralization lagged behind administrative and fiscal decentralization (Falleti, 2010; Beach & Barrenechea, 2011). During the 1970s, the military government transferred primary school administration to provinces without fiscal counterparts. Later, the 1992-1993 cycle consolidated decentralized governance by transferring secondary and higher-education institutions –except universities–, this time including fiscal resources although in times of fiscal austerity. Nonetheless, as the 1970s’ administrative decentralization corroborated the provincial capacity to administer schools without fiscal transfer, national authorities had little interest in withdrawing their political authority (Rhoten, 2000). This phenomenon, called the reactive mechanism, describes how the delayed political decentralization allowed the national state to maintain leverage over subnational entities (Falleti, 2010).

During the presidencies of Néstor and Cristina Kirchner (2003–2015), Argentina experienced a re-centralization of education policies. This process was institutionalized through the Federal Education Council, which shifted from issuing voluntary resolutions during the 1990s to enforcing mandatory directives. This transformation granted the national government greater authority in shaping the education sector and constraining provincial autonomy, such as shifts in curriculum guidelines or assessment procedures. The influence of partisan and territorial political interests, particularly Kirchner’s close ties with CTERA, played a pivotal role in sidelining efforts to intensify test-based accountability mechanisms in the education sector. Instead, political

priorities aligned more closely with maintaining union support and ensuring stability within the system (Chambers-Ju, 2024).

Utilizing institutional conversion mechanisms, Macri's administration (2016 to 2019) marked a significant shift in the role of the national government, its approach to federalism, its focus on provincial autonomy and accountability policies. Leveraging the existing laws, Macri's administration enacted a state retrenchment by liberalizing fiscal transfers to the provinces, decoupling them from specific reforms or programs, and allowing greater discretion in allocating funds. Simultaneously, the Macri administration embraced test-based accountability policies in education, signaling a new dynamic in federal-provincial relations characterized by wider local autonomy alongside heightened demands for educational results. In other words, as the national government stepped back, it pressured provinces to take responsibility for educational outcomes (Benveniste, 2002).

At the same time, the consolidation of provincial authority in administering the education sector alongside its political and fiscal competencies resulted in greater political autonomy and decoupling from national guidelines. Starting in 2016, as provinces asserted their independence, the federal government's role blurred due to its successive failures to reach country-wide agreements on education. This erosion of federal leadership allowed provincial actors to leverage education for internal partisan-electoral differentiation. The education sector became a crucial arena for policy divergence and partisan contestation between provincial and national governments, exemplified by their approach and the use of standardized learning assessments or participation in international testing programs. Provinces found political incentives to take distance or align with national policies based on ideological affinity or partisan strategies.

This dynamic not only intensified partisan and territorial disputes but also contributed to the fragmentation of the national education landscape. As provinces tailored their educational agendas to local political considerations, the national space displayed the political cleavages between provincial and federal authorities. In sum, if in recent years after 1990s decentralization the national government withheld political authority over provinces (Falleti, 2010), the combination of policy alterations and their feedback effects upon relevant constituents resulted in wider provincial autonomy, at least from most resourceful ones.

5.3.2. Decentralization in Colombia: controlling schools at a distance

In Colombia, the decentralization process was politically driven, starting with the direct election of mayors in 1986 and departmental governors in 1991, followed by fiscal and administrative decentralization in 1993-1994 (Azóca & Giraldo, 1994). Colombia's decentralization was quite different from Argentina's. If Argentina's provinces had to handle school administration without funds from the 1970s until the 1990s, Colombia's fiscal and administrative decentralization came after political transfer. During the 1991-1994 reform cycle, Colombia national government handover school administration to Certified Territorial Entities, which were responsible for their education system management, including curricula, education quality, and staff management. However, administrative and fiscal decentralization of school administration has been marked by significant resource disparities (Edwards et al., 2023).

On paper, each of the 97 Certified Territorial Entities has political, administrative, and fiscal authority over its school system. The national government transfers funds to Certified Territorial Entities based on enrollment numbers and teaching staff. Still, a crucial component of financial support depends on the resources available to each department. In 2001, the counter-

reform triggered after the crisis became a turning point in Colombia's fiscal decentralization. The national government seized momentum to advance a fiscal reform fostering per-capita financing and a new resource allocation formula, constraining territorial autonomy (Rodriguez Guarín, 2023). This has widened departments and territorial entities' disparities in governing their education systems, with Bogotá being an outlier due to its economic resources (Esper, 2024).

Furthermore, in 2002 the Colombian government initiated a re-centralization strategy aligned with New Public Management principles. The new policy directions were grounded in the 2001 fiscal reform as well as the existing guidelines in the 1994 General Education Law. First, the government established national curricular standards in 2002, as mandated by the 1994 law. To further consolidate central authority, it launched the national evaluation system "SABER", whose intensity grew over time, alongside a centralized EMIS, occupying a crucial role in fiscal oversight. The new EMIS allowed the national government to control resource allocation based on the number of students per school and teachers, imposing stricter guidelines on territorial entities and shifting decision-making authority away from local control.

At the same time, the 2001 fiscal reform was a critical tool to dispute FECODE's power, as it enabled salary caps and discretionary limitations to teachers' promotion set by the Ministry of Education. Since teacher salaries remained a national responsibility during the 1990s reform, it became a crucial tool in political negotiations with the unions. Undoubtedly, central control over financial resources and teachers' compensation meant significant leverage in curtailing union influence (Chambers-Ju, 2024; Schneider, 2024). In conclusion, Colombia's recentralizing efforts did not aim to reclaim school administration. Instead, it sought to establish a centralized system of educational control through performance and fiscal oversight.

5.4. Antagonistic ideas in constant dispute

The power of ideas in explaining reforms, institutional changes, or political developments is undeniable. Policy paradigms operate as normative frameworks that guide human actions and decision-making toward specific ideals and encapsulate particular interests and value systems for organizing the social world (Béland, 2009). Paradigms do not merely guide decisions in abstract terms, they mirror and embody power relations: ideas entail whose interests are prioritized. One of the main challenges in working with ideas is operationalizing them to trace their effects and influence on the social phenomena we observe. I approached this task by identifying the actors and organizations championing policy ideas to disentangle the interests they pursue both discursively and in material terms.

In the two cases, I exposed an ongoing dispute between two worldviews and their advocates that have shaped the country's education sectors over the past decades. I referred to one of them as the neoliberal-managerial paradigm. In education, this paradigm advocates for a withdrawal of the state in favor of increased private sector involvement, the imposition of market-based competition to regulate social interactions, and the suppression of collective class action through labor unions (Harvey, 2005). In Colombia, these ideas were evident in three concrete moments: (1) the 1991 "*Plan de Apertura Educativa*," which liberalized private sector participation (Edwards et al., 2023), (2) the funding cuts occurred during the 2001 counter-reform Rodríguez Guarín, 2023), and (3) the adoption of Concession Schools in 1999 (Edwards et al., 2017). However, this paradigm would be incomplete without acknowledging its managerial component. While the neoliberal dimension emphasizes privatization and market logic, the managerial part focuses on centralized oversight, procedural standardization and benchmarking, and performance-based incentives (Gunter et al., 2016). Managerialism examples include the "teacher-proof" curriculum

reform of the 1970s, preceding the start of the ‘*Movimiento Pedagógico*,’ or the 2000s “Educational Revolution,” whose theory of change centered on standards, data, and outcome evaluation (see Figure 1 in Article 1).

On the opposite side of the ideological spectrum lies the democratic-inclusive paradigm. In stark contrast to neoliberal ideas, this paradigm advocates for a welfare state model in which the state acts as the primary guarantor of education for all, limiting private sector participation. At its core, this paradigm views education as a fundamental human right and a public good, opposing the pro-market neoliberal ideas. Hence, performance-based accountability, standardized evaluations, performance-based pay, and constraints on professional autonomy are met with skepticism and seen as undermining mechanisms of the intrinsic social value of education (Oliveira & Feldfeber, 2024). Teacher unions frequently rely on the democratic-inclusive paradigm to mobilize actors and resources in the political arena, framing their opposition to reforms as a defense of public education and professional autonomy.

What differentiates Argentina and Colombia is the length and extent of neoliberal-managerial dominance in education. In Colombia, champions of neoliberal-managerial ideas have ruled the political scene from the mid-1990s until 2023, directly affecting the country’s education sector. In contrast, while neoliberal-managerial ideas reigned Argentina during the 1990s, opposing ideas prevailed from the early 2000s onwards. The varying success of the political groups representing these paradigms largely explains cross-country differences in the adoption and institutionalization of SAWA policies.

5.5. Coalitions and interest groups: victories, defeats, and unfinished businesses

In the adopted historical and comparative approach, I employed a mid-range theoretical framework to analyze how the sequence of events and the actions of political coalitions have shaped the varied trajectories of the SAWA reform. This mid-range approach combines large-scale institutional dynamics with individual actors' agency, allowing for a more nuanced understanding of the reform process often overlooked in macro-sociological historical approaches (see Falleti and Thelen, 2015). Using the Advocacy Coalition Framework as a theoretical foundation, I identified coalitions of supporters and opponents of SAWA reforms over time and across different cases.

In the Advocacy Coalition Framework, the ideas, values, and beliefs around which coalitions are organized are stable over time and grant coalitions of temporal endurance (Jenkins-Smith et al., 2019; Sabatier & Weible, 2007). For example, we observed that coalitions formed by teacher unions, leftist intellectuals, and left-leaning political parties frequently resisted reforms, focused on managerial autonomy and accountability. However, this opposition was not always homogeneous. For instance, during the last Peronist coalition government in Argentina (2019–2023), we observed greater acceptance of standardized evaluation policies influenced by the preceding cycle of reforms. The success of Mauricio Macri's administration (2015–2019) in institutionalizing an “evaluation culture” (Diker et al., 2023), coupled with the urgency created by Argentina's poor learning outcomes in national and international assessments, facilitated a discursive shift favoring standardized testing spread at both national and sub-national levels.

Among the different interest groups pertaining to policy coalitions, teacher unions have been pivotal to our analysis. The two have national representation and local imbrications with sizable mobilization capacity, granting them strong veto power over policy changes. However, there are sensitive differences between FECODE in Colombia and CTERA in Argentina. For

instance, FECODE has greater internal cohesion, partly due to its success in blocking the municipalization of education in the 1990s. As a result, all unionized teachers in state schools are FECODE's members. In contrast, Argentina's decentralization fragmented the union landscape, creating multiple factions with diverse political alignments (Chambers-Ju, 2024). Additionally, FECODE and CTERA faced divergent opportunities for influencing the government's agenda. While FECODE was an unconditional opponent of right-wing governments in Colombia from the 1990s until 2023, CTERA was a strategic ally of the Peronist governments during 16 years of rule from the 2000s onwards. This partnership enabled CTERA to secure sustained salary increases (Claus & Bucciarelli, 2020) and substantially sway national education policy.

Hence, I argue that the adoption, retention, institutionalization, or rejection of SAWA policies hinges on coalitions' ability to assert their agendas within the political arena. Figure 1 in Paper 3 illustrates this dynamic and outlines how coalitions shape educational policy. Different coalitions seek to access decision-making positions after elections or in times of crisis, such as the Ministry of Education. Once in office, they push to define policy problems, promote their preferred solutions, and mobilize support to drive reforms (Jenkins-Smith et al., 2019; Mukherjee & Howlett, 2015). When capable of coupling the streams to place reforms into the public agenda, policies often follow two routes. When proposed alternatives 'align' with the existing institutional framework or create minor changes to established power relations, opposing groups have limited room for action. However, these policies often produce changes to existing rules. For instance, the intensification of testing frequency without altering their stakes, as occurred in Argentina (2016-2019) and Colombia (2012-2014), was achieved with minimal resistance.

However, opposition intensifies when reforms challenge the 'status quo' or alter power dynamics. Reform success depends on overcoming opposition and securing legislative support for

normative changes. These reforms often represent what Haas (1993) labeled second-order changes. A case in point is Argentina’s failed attempt (2015-2019) to amend Article 97 of the Education Law to publish standardized test results, which lacked legislative backing. Similarly, Colombia’s project to implement a national curriculum (2014-2016) faced strong resistance from teachers, who viewed it as threatening school autonomy.

In conclusion, the main driver explaining the differences in SAWA policies in Argentina and Colombia has been the role of policy coalitions. As argued, the ability of coalitions in favor or against SAWA reforms to leverage the effects of governance decentralization, the shift in political mood or the pressures from IOs underline the adoption or resistance of SAWA reforms in both countries. Table 6 summarizes the outcome of this historical analysis, expressed in five different categories: adoption (A), partial adoption (P), failed adoption (F), suspension after adoption (S). For each SAWA policy, coalitions have faced political or institutional resistance as changes affected the existing power structures and benefits among stakeholders. Hence, whether reform advocates in Argentina and Colombia succeed or fail in overcoming such resistance lies the answer to SAWA policies trajectories over the last two decades.

Table 6. A Comparison of SAWA policies in Argentina and Colombia (1990-2023)

SAWA Elements	Policy	Argentina	Colombia
A. School Autonomy & Decentralization	1. Decentralization	A (1993)	A (1988-1993)
	2. Curriculum Autonomy	A (1997)	A (1994)
	3. Budget Autonomy	-	P (1998)
	4. Staffing Autonomy	-	P (1998)
B. Accountability & Standardization	5. National learning standards/curriculum	A (1993)	S (1994); A (2002/3)
	6. Standardized national exams in primary education (ISCED 1)	A (1993)	A (2002/3)
	7. Standardized national exams in secondary education (ISCED 2)	A (1993)	A (2002/3)
	10. Standardized exams for certification	-	

SAWA Elements	Policy	Argentina	Colombia
	purposes		
	11. Exam results are publicly available at the school level	F (2017)	
	11. School choice	A (1962)*	-
	12. Per-capita funding	-	A (2001)
C. Competition (with pro-market elements)	13. School rankings or quality indexes	F (2017)	A (2015)
	14. Private management of public schools (i.e. Charter schools)	-	A (1998)
	15. State subsidies to private schools (demand or supply-side)	A (1947)	12. Per-capita funding
D. Performance-based incentives	16. Teacher bonuses linked to students' performance on standardized assessments	-	S (2015-2017)
	17. School bonuses linked to students' performance in standardized assessments	-	-

Note: Adoption (A), Partial Adoption (P), Failed Adoption (F), Suspension after adoption (S), and (-) Not Applicable when it was never seriously considered.

** In cases of policies were adopted prior to the prior of study*

5.6. Intended research contributions

This dissertation was guided by two central questions: (1) to what extent have autonomy and accountability policies spread throughout Latin America? and (2) what factors explain the variations in the adoption and institutionalization of these reforms in the cases of Argentina and Colombia? Having dedicated this conclusion to summarize the salient point that answered them, I aim now to highlight this study's intended contributions to policy studies and policy transfer research.

Firstly, as shown in the introduction, the uneven adoption of autonomy and accountability policies across 23 Latin American countries reaffirms that theories of convergence towards a common model of educational governance overlook country-specific variations, leading to the often-inaccurate conclusion of a homogenizing world. The detailed examination of the various policies that operationalize these reforms has allowed us to observe more precisely which elements

have gained greater prominence and which have encountered more resistance or been less compatible with the specific characteristics of each country. In a forthcoming publication, I delve further into these findings, analyzing how global drivers of economic coercion from IOs, competition among countries in the educational arena, and normative pressures to follow international norms or scripts can explain the adoption of such reforms. Of course, on top of global drivers, I also consider whether the countries' characteristics, such as government ideologies, the presence of unions, or levels of centralization or decentralization of the educational system, have more prevalence in explaining the observed variation across these policies. Despite this ongoing work, the findings presented here indicate that elements closely associated with centralized control over education, such as standardized testing and national curricula, have been more widely disseminated than other aspects of this reform, such as competition mechanisms, school-level decision-making regarding teaching staff or budgets, and financial incentives tied to standardized test performance.

Using historical institutionalism as an analytical framework, a second contribution of this thesis pertains to the processes of policy change driven by layering, conversion, and drift mechanisms. For instance, in the case of Colombia, the progressive accumulation of reforms through successive layers, combined with the strategic exploitation of contextual drifts (Hacker et al., 2015), has gradually consolidated a new logic of educational system governance rooted in the principles of new public management. In Argentina, what initially emerged as a decentralization that limited the provinces' political autonomy in education (Falleti, 2010), has progressively evolved into a scenario where the provinces increasingly disregard the national role. In other words, over time, shifts in the balance of power among actors in the system have led to changes in institutional arrangements without requiring abrupt breaks from the established order.

A third contribution relates to the effort to explain and unravel how policy coalitions operate to drive policy change, whether through adopting new reforms or modifying existing ones. Our perspective on the mechanisms at play seeks to explain the regularities in the actions of different actors and their effects as part of a coherent process that recurs in a consistent manner (Kuhlmann & Nullmeier, 2021). In this regard, the proposed mechanism that explains how policy coalitions work operated through a regular pattern. Coalitions' nature lies in their shared beliefs and interests in imposing their policy preferences, for which they attempt to influence the public agenda and secure the necessary support for their reforms. At the same time, this mechanism highlights the challenges that arise during the reform process. The more a reform deviates from existing logic and shakes existing power relations among interest groups, the greater the resistance and, consequently, the need to secure broad cross-sectoral agreements to avoid vetoes.

The final contribution to be highlighted relates to the view of the policy transfer process as a phenomenon encompassing multiple stages beyond mere adoption, including the subsequent trajectory of the policy. Focusing on the evolution of the policy as part of the transfer process, emphasizes the analytical relevance of observing persistent factors and actors that advocated for its adoption or, conversely, sought to oppose it. However, this temporal evolution reflects shifts in the balance of power among actors and changes in the social, political, and institutional contexts in which these reforms are embedded. By paying attention to these dynamics, we can better understand why specific policies are abandoned while others maintain their relevance over time. Similarly, the multi-scalar lens allowed us to overcome biases restricting the analysis to the nation-state level, enabling us to explore the diverse effects occurring across transnational, national, and local levels. This approach has revealed how Bogotá played a crucial role in shaping the trajectory

of reforms in Colombia and how Buenos Aires succeeded in transferring its pro-evaluation agenda to the national level, which was subsequently replicated by other jurisdictions.

5.7. A developing research agenda

Ahead of a rapidly changing global landscape, some of the certainties that we have withheld until today seem to fall apart. Western hegemony is clearly in decline, along with its values and institutions – democracy, public schooling, peace, scientific prevalence, etc. The illiberal upheaval, the rise of corporations wealthier than most nation-states, and the economic success of authoritarian regimes cast doubts over future pathways in global governance: how will IOs respond and adapt to this changing environment? What agendas will thrive, and which ones will be buried? How will the still-developing new world order reshape education policy? These are intriguing questions lying ahead of us. As social scientists, these puzzles challenge our established viewpoint, asking for new frames and lenses to face pressing and unsettling changes.

Firstly, concerning policy diffusion and transfer research, there remain unanswered questions about the unfolding of global mechanisms in explaining the spread of reforms. To what extent can the re-emerged anti-liberal values or the counter 2030-Agenda –which denies climate change, attacks minorities, and denies social justice values– influence countries and bring about previously unthinkable reforms? We have suggested here that specific characteristics of policies, more closely aligned with neoliberal principles of competition and marketization, could explain why they have been less widely adopted. Is this truly the case? What specific attributes of policy templates make them more universally applicable or less so? Beyond the usual suspects in reform diffusion, what other actors play a significant role in this process? How do transnational corporations seek to impose their policy agendas and values worldwide?

The second focus of this research agenda concerns the evolving influence of multilateral organizations and their future agendas for education governance in response to shifts in the global political-economic order. As previously noted, Latin America has emerged as a region under the growing influence of the OECD. This is evidenced by the increasing participation of countries in PISA assessments and the growing number of member states or applicants (Chile in 2011, Colombia in 2020, Costa Rica in 2021, with Peru, Brazil, and Argentina formally seeking admission). What drives these countries to join the OECD, and conversely, what interests and benefits does the organization gain from incorporating more developing nations like those in Latin America? The accession process is particularly significant due to the intense exchange of ideas, interests, and personnel, which can facilitate policy transfer and shape new agendas within both the member states and the organization itself. Additionally, a less-explored aspect is the formation of regional alliances between regional multilateral bodies and global organizations like the OECD or the World Bank, which play pivotal roles in shaping policy directions.

Lastly, despite the apparent exhaustion of the SAWA agenda, many of the proposed policies remain in effect, as do the underlying ideas. School curricula is yet a battleground for educational and societal control. Educational testing and accountability continue to evolve in light of the growing datafication of public administration. Education privatization is still on the rise. Undoubtedly, existing policies will be adjusted in light of new prevailing agendas, as with SAWA. Furthermore, teachers and teachers' policies are a privileged target, underscored in our analysis, but directly linked to any educational reform the effects of autonomy and accountability policies. In what direction will these policies move? What interests will prevail, and who will endorse them? These are some of the open questions that call for answers in the years to come.

References

- Acosta, F. (2019). OECD, PISA and the Educationalization of the World: The Case of the Southern Cone Countries. In C. Ydesen (Ed.), *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex* (pp. 175–196). Springer International Publishing. https://doi.org/10.1007/978-3-030-33799-5_9
- Addey, C., Sellar, S., Steiner-Khamsi, G., Lingard, B., & Verger, A. (2017). The rise of international large-scale assessments and rationales for participation. *Compare: A Journal of Comparative and International Education*, 47(3), 434–452. <https://doi.org/10.1080/03057925.2017.1301399>
- Au, W. (2007). High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis. *Educational Researcher*, 36(5), 258–267. <https://doi.org/10.3102/0013189X07306523>
- Avelar, M., & Ball, S. J. (2019). Mapping new philanthropy and the heterarchical state: The Mobilization for the National Learning Standards in Brazil. *International Journal of Educational Development*, 64, 65–73. <https://doi.org/10.1016/j.ijedudev.2017.09.007>
- Azóca, J. C. T., & Giraldo, H. D. (1994). El proceso de descentralización educativa en Colombia. *Revista Colombiana de Educación*, 29, Article 29. <https://doi.org/10.17227/01203916.5364>
- Baek, C. (2022). Internalising externalisation: Utilisation of international knowledge in education policymaking. *British Journal of Sociology of Education*, 43(1), 159–176. <https://doi.org/10.1080/01425692.2021.2007357>
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215–228. <https://doi.org/10.1080/0268093022000043065>
- Ball, S. J. (2009). New states, new governance and new education policy. In S. J. Ball, M. Apple,

- & L. A. Gandin (Eds.), *The Routledge International Handbook of the Sociology of Education*. Routledge Handbooks Online. <https://doi.org/10.4324/9780203863701.ch14>
- Banco Mundial, & Gobierno de Mendoza. (2023). *De vuelta a lo básico: Evaluación de fluidez lectora y uso pedagógico de resultados para la recuperación de los aprendizajes. La Experiencia de la Provincia de Mendoza, Argentina*.
- Barrera-Orsorio, F., Fasih, T., Patrinos, H. A., & Santibáñez, L. (2009). *Decentralized Decision-making in Schools: The Theory and Evidence on School-based Management*. World Bank. <https://openknowledge.worldbank.org/handle/10986/2632>
- Bartlett, L., & Vavrus, F. (2016). *Rethinking Case Study Research*. Routledge.
- Baxter, J. G., & León Cadavid, M. C. (2021). From Political Intentions to Structural Interventions A Review of Two Decades of Education Policy Reforms in Colombia. In A. Rivas (Ed.), *Examining Educational Policy in Latin America* (pp. 96–108). Routledge.
- Beach, D., & Pedersen, B. (2016). *Causal Case Study Methods: Foundations and Guidelines for Comparing, Matching, and Tracing*. University of Michigan Press. <https://doi.org/10.3998/mpub.6576809>
- Beach, D., & Pedersen, R. B. (2018). Selecting Appropriate Cases When Tracing Causal Mechanisms. *Sociological Methods & Research*, 47(4), 837–871. <https://doi.org/10.1177/0049124115622510>
- Beach, D., & Pedersen, R. B. (2019). *Process-tracing methods: Foundations and guidelines* (Second Edition). University of Michigan Press.
- Beech, J. (2019). The Long and Winding Road to Inclusion: Educational Policies in Argentina (2003–2015). In C. Ornelas (Ed.), *Politics of Education in Latin America* (Vol. 49, pp. 17–42). Brill Sense.

- Beech, J., & Barrenechea, I. (2011). Pro-market educational governance: Is Argentina a black swan? *Critical Studies in Education*, 52(3), 279–293.
<https://doi.org/10.1080/17508487.2011.604077>
- Béland, D. (2009). Ideas, institutions, and policy change. *Journal of European Public Policy*, 16(5), 701–718. <https://doi.org/10.1080/13501760902983382>
- Béland, D. (2010). The Idea of Power and the Role of Ideas. *Political Studies Review*.
<https://journals.sagepub.com/doi/10.1111/j.1478-9302.2009.00199.x>
- Béland, D., & Powell, M. (2016). Continuity and Change in Social Policy. *Social Policy & Administration*, 50(2), 129–147. <https://doi.org/10.1111/spol.12207>
- Berliner, D. C., & Glass, G. V. (2022). The effects of high-stakes testing on students and schools. *Review of Educational Research*, 92(3), 445–477.
[doi:10.3102/00346543211070058](https://doi.org/10.3102/00346543211070058)
- Benveniste, L. (2002). The Political Structuration of Assessment: Negotiating State Power and Legitimacy. *Comparative Education Review*, 46, 89–118.
- Blatter, J., Portmann, L., & Rausis, F. (2022). Theorizing policy diffusion: From a patchy set of mechanisms to a paradigmatic typology. *Journal of European Public Policy*, 29(6), 805–825. <https://doi.org/10.1080/13501763.2021.1892801>
- Boli, J., Ramirez, F. O., & Meyer, J. W. (1985). Explaining the Origins and Expansion of Mass Education. *Comparative Education Review*, 29(2), 145–170.
- Bonal, X. (2002). Plus ca change ...the world bank global education policy and the post-washington consensus. *International Studies in Sociology of Education*, 12(1), 3–22.
<https://doi.org/10.1080/09620210200200080>
- Bonal, X., Pagès, M., Verger, A., & Zancajo, A. (2023). Regional policy trajectories in the

- Spanish education system: Different uses of relative autonomy. *Education Policy Analysis Archives*, 31. <https://doi.org/10.14507/epaa.31.8031>
- Borges, F. A. (2018). Neoliberalism with a Human Face? Ideology and the Diffusion of Latin America's Conditional Cash Transfers. *Comparative Politics*, 50(2), 147–167.
- Braginski, R. (2018, March 2). Macri pidió un “ranking” de escuelas y desató una polémica. *Clarín*. https://www.clarin.com/sociedad/macri-pidio-ranking-escuelas-desato-polemica_0_BJVXhM8_G.html
- Bromley, P., Furuta, J., Kijima, R., Overbey, L., Choi, M., & Santos, H. (2023). Global Determinants of Education Reform, 1960 to 2017. *Sociology of Education*, 96(2), 149–167. <https://doi.org/10.1177/00380407221146773>
- Bromley, P., Rie, K., Overbey, L., Furuta, J., Choi, M., & Santos, H. (2022). *World Education Reform Database (WERD): A Global Dataset on Education Reforms** (World Education Reform Database). <https://static1.squarespace.com/static/62e92b5b7adb4e2530864f96/t/6329e9f41e963f6871367fd1/1663691254768/WERD+methods+Sept+22.pdf>
- Buckner, E., & Russell, S. G. (2013). Portraying the Global: Cross-national Trends in Textbooks' Portrayal of Globalization and Global Citizenship. *International Studies Quarterly*, 57(4), 738–750. <https://doi.org/10.1111/isqu.12078>
- Bunar, N., & Ambrose, A. (2016). Schools, choice and reputation: Local school markets and the distribution of symbolic capital in segregated cities. *Research in Comparative and International Education*, 11(1), 34–51. <https://doi.org/10.1177/1745499916631064>
- Burki, S. J., & Perry, G. E. (1998). *Beyond the Washington consensus*. The World Bank. <https://doi.org/10.1596/0-8213-4282-7>

- Camphuijsen, M. K., & Parcerisa, L. (2023). Teachers' beliefs about standardised testing and test-based accountability: Comparing the perceptions and experiences of teachers in Chile and Norway. *European Journal of Education*, 58(1), 67–82.
<https://doi.org/10.1111/ejed.12540>
- Capano, G., & Howlett, M. (2020). The Knowns and Unknowns of Policy Instrument Analysis: Policy Tools and the Current Research Agenda on Policy Mixes. *SAGE Open*, 10(1), 215824401990056. <https://doi.org/10.1177/2158244019900568>
- Cardenas, S. (2009). *School Based Management in Latin America*.
<https://doi.org/10.13140/2.1.1985.3443>
- Carrió, T. (2019, December 1). Claves para entender los esfuerzos de la Argentina por ingresar a la OCDE. *Apertura Negocios*. <https://www.apertura.com/economia/Claves-para-entender-los-esfuerzos-de-la-Argentina-por-ingresar-a-la-OCDE-20171201-0002.html>
- Caviedes, J. F. R. (2018). El Banco Mundial y la política educativa para Colombia y Brasil. *Revista Latinoamericana de Estudios Educativos (México)*, XLVIII(1), 53–72.
- Chambers-Ju, C. (2006). *Después de la Autonomía: Programas para Mejorar la Gestión Escolar en El Salvador, Colombia, Chile y Brasil*. PREAL.
- Chambers-Ju, C. (2021). Adjustment Policies, Union Structures, and Strategies of Mobilization: Teacher Politics in Mexico and Argentina. *Comparative Politics*, 53(2), 185–207.
<https://doi.org/10.5129/001041521X15918883398085>
- Chambers-Ju, C. (2024). *Mobilizing Teachers: Education Politics and the New Labor Movement in Latin America*. Cambridge University Press.
<https://www.cambridge.org/core/books/mobilizing-teachers/3FE82A72A4EB7D78B9B60BF6940D074E>

- Claus, A., & Bucciarelli, M. E. (2020). *El financiamiento educativo en Argentina. El desafío de la equidad bajo un esquema federal*. CIPPEC.
- Collier, D. (2011). Understanding Process Tracing. *PS: Political Science & Politics*, 44(4), 823–830. <https://doi.org/10.1017/S1049096511001429>
- Cowen, R. (2009). The transfer, translation and transformation of educational processes: And their shape-shifting? *Comparative Education*, 45(3), 315–327. <https://doi.org/10.1080/03050060903184916>
- Cox, C. (2012). Política Y Políticas Educativas En Chile 1990-2010. *Revista Uruguaya de Ciencia Política*, 21(1), 13–42.
- Cruz, A. L. A. (2012). La influencia del Banco Mundial en las políticas educativas de Colombia. *Pedagogía y Saberes*, 36, Article 36. <https://doi.org/10.17227/01212494.36pys69.79>
- De Grauwe, A. (2005). Improving the Quality of Education through School-Based Management: Learning from International Experiences. *International Review of Education / Internationale Zeitschrift Für Erziehungswissenschaft / Revue Internationale de l'Education*, 51(4), 269–287.
- del Río, A., Knutsen, C. H., & Lutscher, P. M. (2024). Education Policies and Systems Across Modern History: A Global Dataset. *Comparative Political Studies*, 00104140241252075. <https://doi.org/10.1177/00104140241252075>
- Demas, A., & Arcia, G. (2015). *What Matters Most for School Autonomy and Accountability: A Framework Paper* [Working Paper]. World Bank. <https://openknowledge.worldbank.org/handle/10986/22086>
- Derqui, J. M. (2001). Educational decentralization policies in Argentina and Brazil: Exploring the new trends. *Journal of Education Policy*, 16(6), 561–583.

<https://doi.org/10.1080/02680930110087825>

Di Gropello, E. (2006). *A Comparative Analysis of School-based Management in Central*

America. Washington, DC: World Bank. <https://doi.org/10.1596/978-0-8213-6525-0>

Díaz Ríos, C. M. (2020). The Role of Policy Legacies in the Alternative Trajectories of Test-Based Accountability. *Comparative Education Review*, 64(4), 619–641.

<https://doi.org/10.1086/710767>

Díaz Ríos, C. M. (2024). *Translating Global Ideas: How Policy Legacies and Domestic Politics Shape Education Governance in Latin America*. SUNY Press.

<https://sunypress.edu/Books/T/Translating-Global-Ideas2>

Díaz Ríos, C., & Urbano-Canal, N. (2023). The World Bank and Education Policy in Colombia:

A Comparative Analysis of the Effects of International Organizations' Learning on

Domestic Policy. *Journal of Comparative Policy Analysis: Research and Practice*, 25(1),

101–117. <https://doi.org/10.1080/13876988.2021.1991796>

Diker, G., Briscioli, B., & Stoppani, N. (2023). La institucionalización de la evaluación

educativa en Argentina: Disputas por sus sentidos y usos. *Jornal de Políticas*

Educacionais, 17(2). <https://doi.org/10.5380/jpe.v17i2.92216>

DINIECE. (2003). *Plan de evaluación de la calidad educativa 2003-2007* [Government

document]. Dirección Nacional de Información y Evaluación de la Calidad Educativa.

Ministerio de Educación de la Nación. República Argentina.

<http://repositorio.educacion.gov.ar/dspace/handle/123456789/93137>

Dolowitz, D. P., & Marsh, D. (2000). Learning from Abroad: The Role of Policy Transfer in

Contemporary Policy-Making. *Governance*, 13(1), 5–23. [https://doi.org/10.1111/0952-](https://doi.org/10.1111/0952-1895.00121)

1895.00121

- Edwards, D. B., Caravaca, A., Rappeport, A., & Sperduti, V. R. (2023). World Bank Influence on Policy Formation in Education: A Systematic Review of the Literature. *Review of Educational Research*, 00346543231194725.
<https://doi.org/10.3102/00346543231194725>
- Edwards, D. B., DeMatthews, D., & Hartley, H. (2017). Public-private partnerships, accountability, and competition: Theory versus reality in the charter schools of Bogotá, Colombia. *Education Policy Analysis Archives*, 25(10).
<https://doi.org/10.14507/epaa.25.2556>
- Edwards, D. B., Moschetti, M. C., Grillet, P., Quilabert, E., & Caravaca, A. (2023). *Descentralización y privatización educativa en Colombia: Facilitadores y tensiones*. Education International.
- Edwards Jr, D. B., Rappeport, A., Sperduti, V., & Caravaca, A. (2021). *The influence of the World Bank on policy formation, policy implementation, and private education: A systematic review of the literature*. <https://doi.org/10.13140/RG.2.2.31264.71688>
- Esper, T. (2023). Quasi-markets, accountability, and innovation: Analyzing the case of Free Schools in England through teachers' perspectives. *Revista Española de Educación Comparada*, 42, 193–220. <https://doi.org/10.5944/reec.42.2023.34408>
- Esper, T. (2024). Shifting Paradigms: The Trajectory of School Autonomy and Accountability Policies in Colombia. *Nordic Journal of Studies in Educational Policy*.
- Falabella, A., & Ramos Zincke, C. (2019). La larga historia de las evaluaciones nacionales a nivel escolar en Chile. *Cuadernos Chilenos de Historia de la Educación; Núm. 11 (2019)*; 66-98. <http://revistaschilenas.uchile.cl/handle/2250/154892>
- Falleti, T. G. (2010). *Decentralization and Subnational Politics in Latin America*. Cambridge

University Press.

- Falleti, T. G., & Mahoney, J. (2015). The comparative sequential method. In J. Mahoney & K. Thelen (Eds.), *Advances in Comparative-Historical Analysis* (pp. 211–239). Cambridge University Press. <https://doi.org/10.1017/CBO9781316273104.009>
- Feldfeber, M., Duhalde, M., Puigrós, A., & Robertson, S. (2018). *La privatización educativa en Argentina*. (1ra Edición). Education International - CTERA.
- Fernández, A. E. (2007). América Latina: El fin del “Consenso de Washington”, el estado de la democracia y el debate sobre las dos izquierdas. *European Review of Latin American and Caribbean Studies*, 82, 129–139.
- Frostenson, M. (2015). Three forms of professional autonomy: De-professionalisation of teachers in a new light. *Nordic Journal of Studies in Educational Policy*, 2015(2), 28464. <https://doi.org/10.3402/nstep.v1.28464>
- Furuta, J. (2022). The Rationalization of “Education for All”: The Worldwide Rise of National Assessments, 1960–2011. *Comparative Education Review*, 66(2), 228–252. <https://doi.org/10.1086/718829>
- Ganimian, A. J. (2016). Why do some school-based management reforms survive while others are reversed? The cases of Honduras and Guatemala. *International Journal of Educational Development*, 47, 33–46. <https://doi.org/10.1016/j.ijedudev.2015.12.001>
- Gilardi, F. (2010). Who Learns from What in Policy Diffusion Processes? *American Journal of Political Science*, 54(3), 650–666.
- Greany, T., & Higham, R. (2018). *Hierarchy, markets and networks: Analysing the “self-improving school-led system” agenda in England and the implications for schools*. UCL Institute of Education Press.

- Gunter, H. M., Grimaldi, E., Hall, D., & Serpieri, R. (Eds.). (2016). *New Public Management and the Reform of Education: European lessons for policy and practice*. Routledge.
<https://doi.org/10.4324/9781315735245>
- Hacker, J. S., Pierson, P., & Thelen, K. (2015). Drift and conversion: Hidden faces of institutional change. In J. Mahoney & K. Thelen (Eds.), *Advances in Comparative-Historical Analysis* (pp. 180–208). Cambridge University Press.
<https://doi.org/10.1017/CBO9781316273104.008>
- Hall, P. A. (1993). Policy Paradigms, Social Learning, and the State: The Case of Economic Policymaking in Britain. *Comparative Politics*, 25(3), 275–296.
<https://doi.org/10.2307/422246>
- Hall, P. A. (2010). Historical Institutionalism in Rationalist and Sociological Perspective. In J. Mahoney & K. Thelen (Eds.), *Explaining Institutional Change: Ambiguity, Agency, and Power* (pp. 204–224). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511806414.009>
- Hallett, T. (2010). The myth incarnate: Recoupling processes, turmoil, and inhabited institutions in an urban elementary school. *American Sociological Review*, 75(1), 52–74.
<https://doi.org/10.1177/0003122409357044>
- Hanushek, E. A., Link, S., & Woessmann, L. (2013). Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics*, 104, 212–232. <https://doi.org/10.1016/j.jdeveco.2012.08.002>
- Harvey, D. (2005). *A Brief History of Neoliberalism*. Oxford University Press.
- Henisz, W. J., Zelner, B. A., & Guillén, M. F. (2005). The Worldwide Diffusion of Market-Oriented Infrastructure Reform, 1977–1999. *American Sociological Review*, 70(6), 871–

897. <https://doi.org/10.1177/000312240507000601>

Henry, A. D., Ingold, K., Nohrstedt, D., & Weible, C. M. (2014). Policy Change in Comparative Contexts: Applying the Advocacy Coalition Framework Outside of Western Europe and North America. *Journal of Comparative Policy Analysis: Research and Practice*, 16(4), 299–312. <https://doi.org/10.1080/13876988.2014.941200>

Heyneman, S. P. (2003). The history and problems in the making of education policy at the World Bank 1960–2000. *International Journal of Educational Development*, 23(3), 315–337. [https://doi.org/10.1016/S0738-0593\(02\)00053-6](https://doi.org/10.1016/S0738-0593(02)00053-6)

Hogan, J., & Howlett, M. (2015). Reflections on Our Understanding of Policy Paradigms and Policy Change. In J. Hogan & M. Howlett (Eds.), *Policy Paradigms in Theory and Practice: Discourses, Ideas and Anomalies in Public Policy Dynamics* (pp. 3–18). Palgrave Macmillan UK. https://doi.org/10.1057/9781137434043_1

Högberg, B., & Lindgren, J. (2021). Outcome-based accountability regimes in OECD countries: A global policy model? *Comparative Education*, 57(3), 301–321. <https://doi.org/10.1080/03050068.2020.1849614>

Hood, C. (1991). A Public Management for All Seasons? *Public Administration*, 69(1), 3–19. <https://doi.org/10.1111/j.1467-9299.1991.tb00779.x>

Hossain, M. (2022a). *Diffusing “Destandardization” Reforms across Educational Systems in Low- and Middle- Income Countries: The Case of the World Bank , 1965 to 2020*. <https://doi.org/10.1177/00380407221109209>

Hossain, M. (2022b). Diffusing “Destandardization” Reforms across Educational Systems in Low- and Middle-Income Countries: The Case of the World Bank, 1965 to 2020. *Sociology of Education*, 95(4), 320–339. <https://doi.org/10.1177/00380407221109209>

- Hossain, M. (2023). Large-scale data gathering: Exploring World Bank's influence on national learning assessments in LMICs. *International Journal of Educational Development, 102*, 102877. <https://doi.org/10.1016/j.ijedudev.2023.102877>
- Jenkins-Smith, H. C., Nohrstedt, D., & Weible, C. M. (2019). *The Advocacy Coalition Framework: An Overview of the Research Program* (pp. 92–114).
- Jerrim, J., & Sims, S. (2022). School accountability and teacher stress: International evidence from the OECD TALIS study. *Educational Assessment, Evaluation and Accountability, 34*(1), 5–32. <https://doi.org/10.1007/s11092-021-09360-0>
- Kamens, D. H., & Benavot, A. (2011). National, regional and international learning assessments: Trends among developing countries, 1960-2009. *Globalisation, Societies and Education, 9*(2), 285–300. <https://doi.org/10.1080/14767724.2011.577337>
- Lascoumes, P., & Le Galès, P. (2007). Introduction: Understanding Public Policy through Its Instruments? From the Nature of Instruments to the Sociology of Public Policy Instrumentation. *Governance, 20*(1), 1–21. <https://doi.org/10.1111/j.1468-0491.2007.00342.x>
- Lastra-Anadón, C. X., & Mukherjee, S. (2019). *Cross-country evidence on the impact of decentralisation and school autonomy on educational performance*. OECD. <https://doi.org/10.1787/c3d9b314-en>
- Le Galès, P. (2016). Performance measurement as a policy instrument. *Policy Studies, 37*(6), 508–520. <https://doi.org/10.1080/01442872.2016.1213803>
- Lundström, U. (2015). Teacher autonomy in the era of New Public Management. *Nordic Journal of Studies in Educational Policy, 2015*(2), 28144. <https://doi.org/10.3402/nstep.v1.28144>
- Mahoney, J., & Goertz, G. (2004). The Possibility Principle: Choosing Negative Cases in

- Comparative Research. *American Political Science Review*, 98(4), 653–669.
<https://doi.org/10.1017/S0003055404041401>
- Mahoney, J., & Rueschemeyer, D. (Eds.). (2003). *Comparative Historical Analysis in the Social Sciences* (1st ed.). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511803963>
- Mahoney, J., & Thelen, K. (2010). A Theory of Gradual Institutional Change. In J. Mahoney & K. Thelen (Eds.), *Explaining Institutional Change: Ambiguity, Agency, and Power* (pp. 1–37). Cambridge University Press. <https://doi.org/10.1017/CBO9780511806414.003>
- Mahoney, J., & Thelen, K. (Eds.). (2015). *Advances in Comparative-Historical Analysis*. Cambridge University Press. <https://doi.org/10.1017/CBO9781316273104>
- Maroy, C., & Pons, X. (Eds.). (2019). *Accountability Policies in Education: A Comparative and Multilevel Analysis in France and Quebec*. Springer Cham.
<https://link.springer.com/book/10.1007/978-3-030-01285-4>
- Marsh, D., & Sharman, J. C. (2009). Policy diffusion and policy transfer. *Policy Studies*, 30(3), 269–288. <https://doi.org/10.1080/01442870902863851>
- Matovich, I., & Esper, T. (2023). Following New Philanthropy by Network Ethnography: How Did the Varkey Foundation Land and Expand in Latin America? *ECNU Review of Education*, 20965311231168422. <https://doi.org/10.1177/20965311231168422>
- Mentini, L. (2023). Enacting accountability reforms in Brazil: The mediating role of school leadership. *Práxis Educativa*, 18, 1–21. <https://doi.org/10.5212/PraxEduc.v.18.20774.006>
- Mentini, L., & Levatino, A. (2023). A “three-legged model”: (De)constructing school autonomy, accountability, and innovation in the Italian National Evaluation System. *European Educational Research Journal*, 14749041221148280.

<https://doi.org/10.1177/14749041221148280>

Meseguer, C. (2004). What Role for Learning? The Diffusion of Privatisation in OECD and Latin American Countries. *Journal of Public Policy*, 24(3), 299–325.

Meyer, J. W., Boli, J., Thomas, G. M., & Ramirez, F. O. (1997). World Society and the Nation-State. *American Journal of Sociology*, 103(1), 144–181. <https://doi.org/10.1086/231174>

Meyer, J. W., & Rowan, B. (1977). Institutionalized Organizations: Formal Structure as Myth and Ceremony. *American Journal of Sociology*, 83(2), 340–363.

Milner, A. L., & Ydesen, C. (2024). Producing accountability with autonomy in Denmark and England: The analytical potential of assemblage theory with critical realism in comparative education. *Nordic Journal of Studies in Educational Policy*, 0(0), 1–14. <https://doi.org/10.1080/20020317.2024.2392892>

Ministerio de Educación Nacional. (2010). *Revolución Educativa 2002-2010 Acciones y lecciones*. Ministerio de Educación Nacional. <https://dds.cepal.org/redesoc/publicacion?id=1132>

Morais De Sá e Silva, M. (2017). *Poverty Reduction, Education, and the Global Diffusion of Conditional Cash Transfers*. Palgrave Macmillan Cham. <https://link.springer.com/book/10.1007/978-3-319-53094-9>

Morais de Sá e Silva, M., & Porto de Oliveira, O. (2023). Incorporating Time into Policy Transfer Studies: A Comparative Analysis of the Transnational Policy Process of Conditional Cash Transfers and Participatory Budgeting. *Journal of Comparative Policy Analysis: Research and Practice*, 25(4), 418–438. <https://doi.org/10.1080/13876988.2023.2193961>

Moreno Salto, I. (2023). *Harmony and Cacophony in Large-scale Assessments in Education: A*

- Mexican Account*. Brill. <https://brill.com/display/title/69042>
- Morgan, C. (2007). *The OECD Programme for International Student Assessment: Unraveling a Knowledge Network*.
https://www.academia.edu/4907060/The_OECD_Programme_for_International_Student_Assessment_Unraveling_a_Knowledge_Network
- Moschetti, M. C., & Gottau, V. (2016). De Las Rocallosas a Los Andes: El Experimento De Escuelas Charter En El Contexto Político Argentino. *Educação & Sociedade*, 37(134), 55–72.
- Moschetti, M. C., Nistal, M., & Sáenz Guillén, L. (2023). *Subvenciones a la educación privada: ¿para qué y para quiénes? Argentinos por la Educación*.
https://argentinosporlaeducacion.org/wp-content/uploads/2023/11/Subvenciones-a-la-educacion-privada_-%C2%BFpara-que-y-para-quienes_.pdf
- Mukherjee, I., & Howlett, M. (2015). Who Is a Stream? Epistemic Communities, Instrumental Constituencies and Advocacy Coalitions in Public Policy-Making. *Politics and Governance*, 3(2), 65–75.
- Mundy, K., Green, A., Lingard, B., & Verger, A. (2016). Introduction: The Globalization of Education Policy – Key Approaches and Debates. In K. Mundy, A. Green, B. Lingard, & A. Verger (Eds.), *Handbook of Global Education Policy* (pp. 1–20). Wiley-Blackwell.
- Narodowski, M., Gottau, V., & Moschetti, M. (2016). Quasi-State monopoly of the education system and socio-economic segregation in Argentina. *Policy Futures in Education*, 14(6), 687–700. <https://doi.org/10.1177/1478210316645016>
- Narodowski, M., & Moschetti, M. (2015). The growth of private education in Argentina: Evidence and explanations. *Compare: A Journal of Comparative and International*

- Education*, 45(1), 47–69. <https://doi.org/10.1080/03057925.2013.829348>
- Narodowski, M., Moschetti, M., & Alegre, S. (2016). Teacher Strikes and Private Education in Argentina. *Revista Colombiana de Educación*, 70, 27–45.
- Niemann, D., & Martens, K. (2018). Soft governance by hard fact? The OECD as a knowledge broker in education policy. *Global Social Policy*, 18(3), 267–283.
<https://doi.org/10.1177/1468018118794076>
- OECD. (2004). *Reviews of National Policies for Education: Chile 2004*. OECD.
<https://doi.org/10.1787/9789264106352-en>
- OECD. (2011). *School Autonomy and Accountability: Are They Related to Student Performance?* OECD. <https://doi.org/10.1787/5k9h362kcx9w-en>
- OECD. (2016). *Education in Colombia*. OECD Publishing. <https://www.oecd-ilibrary.org/docserver/9789264250604-en.pdf?expires=1663001166&id=id&accname=ocid177456&checksum=199C36D991FF309DE8425BE098B830B9>
- OECD. (2024). *Accession Argentina*. OECD. <https://www.oecd.org/en/countries/argentina.html>
- Oliveira, D. A., & Feldfeber, M. (2024). Unionists, teachers, and academics resisting neoliberal reforms: Building a network as a third space. *Globalisation, Societies and Education*, 0(0), 1–13. <https://doi.org/10.1080/14767724.2024.2396294>
- Ozga, J. (2009). Governing education through data in England: From regulation to self-evaluation. *Journal of Education Policy*, 24(2), 149–162.
<https://doi.org/10.1080/02680930902733121>
- Pagès, M. (2021). Enacting performance-based accountability in a Southern European school system: Between administrative and market logics. *Educational Assessment, Evaluation*

- and Accountability*, 33(3), 535–561. <https://doi.org/10.1007/s11092-021-09359-7>
- Paglayan, A. (2024). *Raised to Obey: The Rise and Spread of Mass Education: 134*.
- Pal, L. A. (2014). Introduction: The OECD and Policy Transfer: Comparative Case Studies. *Journal of Comparative Policy Analysis: Research and Practice*, 16(3), 195–200. <https://doi.org/10.1080/13876988.2014.910910>
- Parcerisa, L., Verger, A., Martín, M. P., & Browes, N. (2022). Teacher autonomy in the age of performance-based accountability: A review based on teaching profession regulatory models (2017-2020). *Education Policy Analysis Archives*, 30, (100)-(100). <https://doi.org/10.14507/epaa.30.6204>
- Patrinós, H. A., Arcia, G., & Macdonald, K. (2015). School autonomy and accountability in Thailand: Does the gap between policy intent and implementation matter? *PROSPECTS*, 45(4), 429–445. <https://doi.org/10.1007/s11125-015-9368-8>
- Pawson, R. (2006). *Evidence-based Policy*. SAGE Publications Ltd. <https://doi.org/10.4135/9781849209120>
- Piattoeva, N., & Vasileva, N. (2023). Taming the time zone: National large-scale assessments as instruments of time in the Russian Federation. *Education Policy Analysis Archives*, 31. <https://doi.org/10.14507/epaa.31.7323>
- Pierson, P., & Skocpol, T. (2002). Historical Institutionalism in Contemporary Political Science. In I. Katznelson & H. V. Milner (Eds.), *Political Science: State of the Discipline* (pp. 693–721). W. W. Norton & Co. <http://www.gbv.de/dms/sub-hamburg/347252621.pdf>
- Pizmony-Levy, O. (2011). Bridging the Global and Local in Understanding Curricula Scripts: The Case of Environmental Education. *Comparative Education Review*, 55(4), 600–633. <https://doi.org/10.1086/661632>

- Porto de Oliveira, O., & Osorio Gonnet, C. (2023). Comparative Transfer Analysis: Policy Instruments, Space and Time. *Journal of Comparative Policy Analysis: Research and Practice*, 25(4), 367–384. <https://doi.org/10.1080/13876988.2023.2235308>
- Ramirez, F. O. (2012). The world society perspective: Concepts, assumptions, and strategies. *Comparative Education*, 48(4), 423–439. <https://doi.org/10.1080/03050068.2012.693374>
- Rhoten, D. (2000). Education decentralization in Argentina: A ‘global-local conditions of possibility’ approach to state, market, and society change. *Journal of Education Policy*, 15(6), 593–619. <https://doi.org/10.1080/02680930010000218>
- Rivas, A. (Ed.). (2021). *Examining Educational Policy in Latin America: Comprehensive Insights into Contemporary Reform*. Routledge.
- Rivas, A., & Sanchez, B. (2022). Race to the classroom: The governance turn in Latin American education. The emerging era of accountability, control and prescribed curriculum. *Compare: A Journal of Comparative and International Education*, 52(2), 250–268. <https://doi.org/10.1080/03057925.2020.1756745>
- Rivas, A., & Scasso, M. G. (2019). Low stakes, high risks: The problem of intertemporal validity of PISA in Latin America. *Journal of Education Policy*. <https://doi.org/10.1080/02680939.2019.1696987>
- Rodrigo, L. (2020). Medición, rendimiento y calidad. Las políticas de evaluación del sistema educativo en la Argentina reciente (2015-2019). *Educere et Educare*, 10.17648/educare.v15i35.23893-10.17648/educare.v15i35.23893. <https://doi.org/10.17648/educare.v15i35.23893>
- Rodriguez Guarín, L. A. (2023). *La contrarreforma educativa neoliberal en Colombia*. LETRAME.

- Rodríguez, L. R., Vior, S. E., & Más Rocha, S. M. (2018). Las Políticas de Evaluación de la Calidad Educativa en Argentina (2016-2018). *Educação & Realidade*, 43, 1405–1428. <https://doi.org/10.1590/2175-623684907>
- Sabatier, C. M. W., Paul A. (Ed.). (2018). *Theories of the Policy Process* (4th ed.). Routledge. <https://doi.org/10.4324/9780429494284>
- Sabatier, P. A., & Jenkins-smith, H. C. (1993). *Policy change and learning: An advocacy coalition approach*. Avalon Publishing.
- Sabatier, P. A., & Weible, C. M. (2007). The Advocacy Coalition Framework: Innovations and Clarifications. In *Theories of the Policy Process, Second Edition* (2nd ed.). Routledge.
- Saldaña, J. (2021). *The Coding Manual for Qualitative Researchers* (Fourth Edition). SAGE Publications.
- Santos, Í., & Centeno, V. G. (2023). Inspirations from abroad: The impact of PISA on countries' choice of reference societies in education. *Compare: A Journal of Comparative and International Education*, 53(2), 269–286. <https://doi.org/10.1080/03057925.2021.1906206>
- Schneider, B. R. (2024). *Routes to Reform: Education Politics in Latin America*. Oxford University Press.
- Seawright, J., & Gerring, J. (2008). Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options. *Political Research Quarterly*, 61(2), 294–308.
- Seitzer, H., Baek, C., & Steiner-Khamsi, G. (2023). Instruments of lesson-drawing: Comparing the knowledge brokerage of the OECD and the World Bank. *Policy Studies*, 0(0), 1–21. <https://doi.org/10.1080/01442872.2023.2220282>
- Simmons, B. A., & Elkins, Z. (2004). The Globalization of Liberalization: Policy Diffusion in

- the International Political Economy. *American Political Science Review*, 98(1), 171–189.
<https://doi.org/10.1017/S0003055404001078>
- Simmons, E. S., & Rush Smith, N. (Eds.). (2021). *Rethinking Comparison: Innovative Methods for Qualitative Political Inquiry*. Cambridge University Press.
<https://doi.org/10.1017/9781108966009>
- Simons, A., & Voß, J.-P. (2018). The concept of instrument constituencies: Accounting for dynamics and practices of knowing governance. *Policy and Society*, 37(1), 14–35.
<https://doi.org/10.1080/14494035.2017.1375248>
- Stefes, C. H. (2019). Historical Institutionalism and Societal Transformations. In W. Merkel, R. Kollmorgen, & H.-J. Wagener (Eds.), *The Handbook of Political, Social, and Economic Transformation* (pp. 95–105). Oxford University Press.
<https://doi.org/10.1093/oso/9780198829911.003.0009>
- Steiner-Khamsi, G. (2021). Externalisation and structural coupling: Applications in comparative policy studies in education. *European Educational Research Journal*, 20(6), 806–820.
<https://doi.org/10.1177/1474904120988394>
- Steiner-Khamsi, G., Appius, S., & Nägeli, A. (2024). Comparing two transfer spaces over time and against a global script: The case of school-autonomy-with-accountability. *Journal of Education Policy*, 0(0), 1–21. <https://doi.org/10.1080/02680939.2024.2394594>
- Steiner-Khamsi, G., Martens, K., & Ydesen, C. (2024). Governance by numbers 2.0: Policy brokerage as an instrument of global governance in the era of information overload. *Comparative Education*, 0(0), 1–18. <https://doi.org/10.1080/03050068.2024.2308348>
- Steiner-Khamsi, G., & Morais de Sa e Silva, M. (2024). Qualitative Comparative Policy Studies: An Introduction from the Special Section Editors. *Journal of Comparative Policy*

- Analysis: Research and Practice*, 0(0), 1–11.
<https://doi.org/10.1080/13876988.2024.2383061>
- Suárez. (2023). *40 años del Movimiento Pedagógico 1982-2022. Los nuevos desafíos de la educación en Colombia hoy*. Magisterio Editorial.
- Tedesco, J. C. (2020). *Educación y sociedad en la Argentina (1880-1955)*. UNIFE: Editorial Universitaria.
https://www.academia.edu/44615073/Educaci%C3%B3n_y_sociedad_en_la_Argentina_1880_1955_Autor_Juan_Carlos_Tedesco_Compilador_Dar%C3%ADo_Pulfer_UNIFE_Editorial_2020
- Teng, S. S., Abu Bakar, M., & Layne, H. (2020). Education reforms within neoliberal paradigms: A comparative look at the Singaporean and Finnish education systems. *Asia Pacific Journal of Education*, 40(4), 458–471. Scopus.
<https://doi.org/10.1080/02188791.2020.1838884>
- Thelen, K. (2004). *How Institutions Evolve: The political economy of skills in Germany, Britain, the United States and Japan*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511790997>
- Varoufakis, Y. (2024). *Technofeudalism*.
<https://www.penguin.co.uk/books/451795/technofeudalism-by-varoufakis-yanis/9781529926095>
- Verger, A. (2012). Framing and selling global education policy: The promotion of public–private partnerships for education in low-income contexts. *Journal of Education Policy*, 27(1), 109–130. <https://doi.org/10.1080/02680939.2011.623242>
- Verger, A., Ferrer-Esteban, G., & Fontdevila, C. (2024). Not everything that moves must

- converge: Evidence from global policy and practice on performance-based accountability. In D. Brent Edwards, A. Verger, M. McKenzie, & K. Takayama (Eds.), *Researching Global Education Policy: Diverse Approaches to Policy Movement* (pp. 39–69). Policy Press. <https://doi.org/10.56687/9781447368045-006>
- Verger, A., Ferrer-Esteban, G., & Parcerisa, L. (2020). In and out of the ‘pressure cooker’: Schools’ varying responses to accountability and datafication. In *World Yearbook of Education 2021*. Routledge.
- Verger, A., Fontdevila, C., & Parcerisa, L. (2019). Constructing School Autonomy with Accountability as a Global Policy Model: A Focus on OECD’s Governance Mechanisms. In C. Ydesen (Ed.), *The OECD’s Historical Rise in Education: The Formation of a Global Governing Complex* (pp. 219–243). Springer International Publishing. https://doi.org/10.1007/978-3-030-33799-5_11
- Verger, A., Fontdevila, C., & Zancajo, A. (2016). *The privatisation of education: A political economy of global education reform* (G. Steiner-Khamsi, Ed.; Issue 2). Teachers College Press. <https://doi.org/10.1080/02680939.2017.1389557>
- Verger, A., Moschetti, M. C., & Fontdevila, C. (2020). How and why policy design matters: Understanding the diverging effects of public-private partnerships in education. *Comparative Education*, 0068. <https://doi.org/10.1080/03050068.2020.1744239>
- 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), 5–30.
- Watson, S., & Supovitz, J. (2001). Autonomy and Accountability in Standards-Based Reform. *Education Policy Analysis Archives*, 9, 32–32. <https://doi.org/10.14507/epaa.v9n32.2001>

- Weyland, K. G. (2007). *Bounded rationality and policy diffusion: Social sector reform in Latin America* [Electronic resource]. Princeton University Press. <http://www.columbia.edu/cgi-bin/cul/resolve?clio13206175>
- Wiborg, S. (2015). Privatizing Education: Free School Policy in Sweden and England. *Comparative Education Review*, 59(3), 473–497. <https://doi.org/10.1086/681928>
- Wimmer, A. (2021). Domains of diffusion: How culture and institutions travel around the world and with what consequences. *American Journal of Sociology*, 126(6), 1389–1438. <https://doi.org/10.1086/714273>
- Wimmer, A. (2023). Layered Legacies. How Multiple Histories Shaped the Attitudes of Contemporary Europeans. *Sociological Science*, 10, 1–46. <https://doi.org/10.15195/v10.a1>
- Woessmann, L., Luedemann, E., Schuetz, G., & West, M. R. (2009). *School Accountability, Autonomy and Choice Around the World*. Edward Elgar Publishing. <https://ideas.repec.org/b/elg/eebook/13540.html>
- Ydesen, C., Milner, A. L., Aderet-German, T., Gomez Caride, E., & Ruan, Y. (2023). *Educational Assessment and Inclusive Education: Paradoxes, Perspectives and Potentialities*. Palgrave Macmillan.
- Ydesen, C., Vidal, D., Paulo, S., & Allender, T. (2019). *The OECD 's Historical Rise in Education*. Palgrave Macmillan Cham. <https://doi.org/10.1007/978-3-030-33799-5>
- Zancajo, A. (2019). Schools in the Marketplace: Analysis of School Supply Responses in the Chilean Education Market. *Educational Policy*. <https://doi.org/10.1177/0895904819881781>
- Zapp, M. (2017). *The World Bank and Education: Governing (through) knowledge*.

International Journal of Educational Development, 53, 1–11.

<https://doi.org/10.1016/j.ijedudev.2016.11.007>

Appendix A. Country Coding Sheet for the SAWA Reform Database

This is the coding protocol for the dissertation project titled “Schools Autonomy with Accountability in Latin America and the Caribbean: studying the diffusion of the new governance of education.” The purpose of this protocol is to identify the adoption of the different policy instruments that operationalized the four elements of the SAWA agenda, namely: (1) school autonomy, (2) accountability and standardization, (3) competition and marketization, and (4) performance-based incentives (see **Appendix A for a detailed description**). The coding protocol contains a number of questions that target each of the 14 instruments of interest. The protocol is organized as a survey containing **57 questions** in total. Data for answering these questions comes from different publications from International Organizations or relevant policy authorities of the country of interest. All sources should be listed in the first sub-section.

The project's interest is to identify the country-level adoption at the regulatory level and the year of adoption of each of the four elements. By policy adoption, we understand the moment when a given policy was sanctioned either as a law, into a policy decree or under a different regulatory instrument. The main focus is the adoption at a national level, yet there are questions about whether the instruments have been implemented at a sub-national level (i.e., state). There are also additional questions that can help gain relevant contextual information as well as other questions in case of doubts or need for further evidence upon the presence of the instrument. If any question cannot be answered using the data from the document, please record: “N/A” (Not Available).

A necessary clarification is that our main focus are state-controlled schools (public schools). When answering your questions, please remember that we are interested in changes regarding how the state governs its schools, as private schools are often run under different rules.

How to save/store the policy documents

- Start with the country code (you can find it here)
- Add the year (i.e., 2012)
- Author (i.e., OECD; World Bank)
- Document title

Example: CHL_2017_OECD Review of National Education Policies

Coding guidance and FAQs:

How to start and make progress?

Begin by reading and coding the documents we have identified from international organizations. After, follow with any document published by the country of interest. Finally, search for the remaining questions on Google or other search engines, but always rely on academic sources or government publications. When using academic sources (i.e. papers or gray literature) try to find the official publication of the policy of interests. For instance, if an author mentions that Chile adopted school autonomy under the LOCE in 1990 (Education Law from 1990), search for the law text to double-check the information.

How much to quote?

The quotes in this coding sheet fulfill two purposes. First, to support our answers with reliable data. Second, to gain contextual information on the policy characteristics in each country. This will help the reader/user to unpack the nuances behind each policy adoption. Thus, in sum, do not hesitate to add multiple quotations that highlight different policy aspects, even when prior quotes would have sufficed to answer the preceding Yes/No question.

What do we understand by ‘policy adoption’?

By policy adoption, we understand the moment when a given policy was sanctioned either as a law, into a policy decree, or under a different regulatory instrument or policy program. We are not focused on when the policy started being applied but rather when it was passed into law or approved. For instance, a law might establish the need for per-capita financing in 1995 but the system would not start operating until 2001. In that case, we will record the year of adoption (1995) and not implementation (2001).

Where to look for NLSAs?

Reformed Project Atlas: <https://reformedproject.eu/sawa-atlas/>

What is per-capita financing?

Per-capita financing means that the state (either national or sub-national) determines funds allocation based on the number of students enrolled in each school. Therefore, changes in the number of enrolled students would impact the funding received by the school or by the entity administering the school system (i.e. municipalities).

What do we mean by parental or school choice?

The terms parental or school choice” mean parents have the freedom to choose the schools their child will attend for primary and secondary education. Traditionally children are assigned to a public school according to where they live and a child's school is assigned by the district.

Country/education system	
UN Country Code and number?	
Reviewer	
Papers and Documents included	
Please include all relevant policy documents and/or academic papers you used to complete this country sheet.	
BACKGROUND INFORMATION	
1. Does the country have a federal or unitary organization?	<ul style="list-style-type: none"> ● Federal system ● Unitary system
1.1 Please add a short description of how the system is organized	
2. How is the school system organized? (Follow ISCED classification) <i>Check on SITEAL Website and UNESCO</i>	Primary education: <i>i.e. 6 years</i> Lower secondary education: <i>3 years</i> Upper secondary education: <i>3 years</i>
3. How many years of schooling are compulsory (at the present time)?	<ul style="list-style-type: none"> ● 9 ● 10 ● 11 ● 12 ● 13 ● If other, specify: _____
4. Which government level is responsible for state school administration?	<ul style="list-style-type: none"> ● The National Government; ● The State or Provincial governments; ● Municipalities or Local Education Authorities;

	<i>If relevant, add a short description.</i>
4.1 In the case of decentralized governance of education, when possible , please include the year of decentralization reforms.	
4.2. Year of Decentralization Reforms	
SAWA Element 1: School Autonomy	
5. Did school principals in state-controlled schools ever have the autonomy to make budgetary/financial decisions? (<i>i.e. How to allocate school funds</i>).	<ul style="list-style-type: none"> ● Yes ● No
Please insert the direct quote with the page number that helped you answer and continue to Q4-Q5.	
6. In what year was budgetary or funds-allocation autonomy adopted?	<i>i.e., 1994</i>
7. What is the law, regulation, policy or program through which this instrument was adopted? Please include all relevant details (type of regulation/policy, year, number, author/publisher or source if available) <i>Feel free to add a short description.</i>	
8. If Q5 is YES, please specify to the best of your ability the degree of budgetary autonomy the school has.	<ul style="list-style-type: none"> ● Total (or 100%) control of the school budget (<i>i.e.</i>, it is allocated to a school bank account); ● Partial control of the school budget (<i>i.e.</i>, the principal can make certain decisions); ● Vague definition (it is not specified); ● Unsure;
Curriculum Autonomy	

9. Did school principals in state-controlled schools ever have curriculum autonomy?	<ul style="list-style-type: none"> ● Yes ● No
Please insert the direct quote with the page number that helped you answer.	
10. In what year was the curriculum autonomy adopted?	
11. What is the law, regulation, or policy through which this instrument was adopted? Please include all relevant details (type of regulation/policy, year, number, author/publisher or source if available) <i>Feel free to add a short description.</i>	
Staffing Autonomy	
12. Did school principals or boards in state-controlled schools ever have staffing autonomy?	<ul style="list-style-type: none"> ● Yes ● No ● Partial (i.e., the principal participates in teachers' evaluation policies that directly impact the candidate's position in a future roster; or the principal is one of the different stakeholders making the hiring decision)
Please insert the direct quote with the page number that helped you answer.	
13. In what year was the staffing autonomy adopted?	
14. What is the law, regulation, or policy through which this instrument was adopted? Please include all relevant	

<p>details (type of regulation/policy, year, number, author/publisher or source if available) <i>Feel free to add a short description.</i></p>	
<p>SAWA Element 2: Accountability and Standardization</p>	
<p>School curricula & learning standards</p>	
<p>15. Has the country ever adopted a national curriculum or academic learning standards?</p>	<ul style="list-style-type: none"> ● Yes ● No
<p>Please insert the direct quote with the page number that helped you answer.</p>	
<p>16. What year was the national curriculum or academic learning standards adopted?</p>	
<p>17. What is the law, regulation, or policy through which this instrument was adopted? Please include all relevant details (type of regulation/policy, year, number, author/publisher or source if available) <i>Feel free to add a short description</i></p>	
<p>Standardized Learning Assessments (NLSAs)</p>	
<p>18. Has the country ever adopted a national a large-scale assessment (NLSAs) for primary education (ISCED level 1)?</p>	<ul style="list-style-type: none"> ● Yes ● No
<p>Please insert the direct quote with the page number that helped you answer.</p>	

<p>19. What is the law, regulation, or policy through which this instrument was adopted? Please include all relevant details (type of regulation/policy, year, number, author/publisher or source if available) <i>Feel free to add a short description</i></p>	
<p>20. In what years has the assessment been adopted?</p>	
<p>21. What is/was the assessment name? Please indicate if the assessment has changed its name over the years</p>	
<p>National Large Scale Assessments (NLSAs) for Lower Secondary Education (ISCED Level 2)</p>	
<p>22. Has “the country” ever adopted NLSAs for lower secondary education (ISCED level 2)?</p>	<ul style="list-style-type: none"> ● Yes ● No
<p>Please insert the direct quote with the page number that helped you answer.</p>	
<p>23. What year has the assessment been adopted?</p>	
<p>24. What is the assessment name? Please indicate if the assessment has changed its name over the years. <i>[Might be the same as in Q22]</i></p>	

<p>25. What is the law, regulation, or policy through which this instrument was adopted? Please include all relevant details (type of regulation/policy, year, number, author/publisher or source if available)</p> <p><i>Feel free to add a short description</i></p>	
National Large Scale Assessments (NLSAs) for Upper Secondary Education (ISCED Level 3)	
<p>26. Has “the country” ever adopted NLSAs for higher secondary education (ISCED level 3)?</p>	<ul style="list-style-type: none"> ● Yes ● No
<p>Please insert the direct quote with the page number that helped you answer.</p>	
<p>27. What year has the assessment been adopted?</p>	
<p>28. What is the assessment name? Please indicate if the assessment has changed its name over the years. <i>[Might be the same as in Q22 / Q25]</i></p>	
<p>29. What is the law, regulation, or policy through which this instrument was adopted? Please include all relevant details (type of regulation/policy, year, number, author/publisher or source if available)</p>	
<p>30. What are the uses or consequences of the NLSAs in upper secondary education?</p>	<ul style="list-style-type: none"> ● The exam is a graduation requirement for upper-secondary ● The exam is an entry requirement for higher education ● There are no direct consequences attached to the exam

<p>If Q30 is YES, please insert the direct quote with the page number that helped you answer.</p>	
<p>NLSAs uses and characteristics</p>	
<p>31. Considering NLSAs at any school level (primary, low_secondary, or upp_secondary):</p> <p>Has “the country” ever made school’s assessment results publicly available?</p> <p><i>This is frequently expressed in school rankings.</i></p>	<ul style="list-style-type: none"> ● Yes, results were publicly available at the school level on the Ministry of Education or specialized agency website. ● Results are publicly available but at an aggregate level (not individualized schools) ● No, results are not publicly available
<p>Please insert the direct quote with the page number that helped you answer.</p>	
<p>32. Considering national assessments in every school level, please add a brief description including:</p> <ul style="list-style-type: none"> ● Why was the assessment adopted? ● What are the main uses of the NLSAs? ● Any other relevant contextual information <p>Differentiate among levels (primary, low_sec, upper_sec) when possible.</p> <p>Feel free to use direct quotes and/or cite what helped you answer this question.</p>	

<p>33. To the best of your ability, in what ways do NLSAs results are used? Select all that apply</p>	<p>Results are barely used in practice. Results are used for evaluating curriculum achievement Educational authorities use results for policy planning or program development. Educational authorities use results to penalize underperforming schools Educational authorities use results to try to improve/support underperforming schools Educational authorities use results to reward top-performing schools To rank schools according to their performance Families use results for selecting schools (school choice) Results are used for teachers' appraisal</p>
<p>34. Does the country have any standardized assessments at the sub-national level?</p>	<p>Yes No</p>
<p>Please insert the direct quote with the page number that helped you answer.</p>	
<p>35. Are standardized assessments at the sub-national level required by the national/federal level?</p>	<ul style="list-style-type: none"> ● Yes ● No ● Unsure ● Not Applicable (N/A)
<p>Please insert the direct quote with the page number that helped you answer.</p>	
<p>SAWA Element 3: Competition and Marketization</p> <p>Parental Choice</p>	

37. Have parents/guardians ever been allowed to choose which state-controlled school their children attend?	<ul style="list-style-type: none"> ● Yes ● No
If Q37 is YES, please insert the direct quote with the page number that helped you answer.	
38. What year was parental choice adopted?	
39. What is the law, regulation, or policy through which this instrument was adopted? Please include all relevant details (type of regulation/policy, year, number, author/publisher or source if available)	
Per-capita funding/financing	
40. Has the country ever established school funding (not to the province or state) linked to the number of students they have (per-capita financing)?	<ul style="list-style-type: none"> ● Yes ● No
Please insert the direct quote with the page number that helped you answer.	
41. In what year was per-capita financing adopted?	
42. What is the law, regulation, or policy through which per-capita financing was adopted? Please include all relevant details (type of regulation/policy, year, number, author/publisher or source if available)	

Deregulation of provision	
43. Does the national government allow private operators to manage state schools? (i.e. Charter schools in the U.S., Colegios en Concesión in Colombia, or Academies and Free Schools in England)	<ul style="list-style-type: none"> ● Yes ● No
Please insert the direct quote with the page number that helped you answer.	
44. What year was the private operation of state schools adopted?	
45. What is the law, regulation, or policy through which private operation of state schools was adopted? Please include all relevant details (type of regulation/policy, year, number, author/publisher or source if available)	
46. Does the national government subsidize (totally or partially) private schools for their operations? (i.e., to cover teachers' salaries; due to the lack of state schools in a given area)	<ul style="list-style-type: none"> ● Yes ● No
Please insert the direct quote with the page number that helped you answer.	
47. In what year were state subsidies to private schools adopted?	

48. What is the law, regulation, or policy through which state subsidies to private schools were adopted? Please include all relevant details (type of regulation/policy, year, number, author/publisher or source if available)	
SAWA Element 4: Performance-based incentives	
49. Has the country ever established school funding linked to performance in standardized assessments or any other school performance metrics (e.g., absenteeism, grade repetition, student survival rates, etc.)? <i>Useful keywords: school bonus; school incentives; school grants</i>	<ul style="list-style-type: none"> ● Yes ● No
Please insert the direct quote with the page number that helped you answer.	
50. What year was the performance-based financing policy adopted?	
51. What is the law, regulation, or policy through which performance-based financing policy was adopted? Please include all relevant details (type of regulation/policy, year, number, author/publisher or source if available)	
School Rankings	
52. Has the country ever used school rankings based on national assessment results?	<ul style="list-style-type: none"> ● Yes ● No
Please insert the direct quote with the page number that helped you answer.	
53. In what year were school rankings adopted?	

<p>54. What is the law, regulation, or policy through which school rankings were adopted? Please include all relevant details (type of regulation/policy, year, number, author/publisher or source if available)</p>	
<p>Teachers' bonuses</p>	
<p>55. Is students' performance in standardized assessments used to determine teachers' or principals' salaries (performance-based payments)? <i>Either as part of teachers' wages or as a performance bonus.</i></p>	<p>Yes</p>
<p>Please insert the direct quote with the page number that helped you answer.</p>	
<p>56. What years were teachers' or principals' performance-based payments adopted?</p>	
<p>57. What is the law, regulation, or policy through which this instrument was adopted?</p>	
<p>Notes Please, write down any comments, questions or remaining doubts after coding the present document.</p>	

Table 1. The different dimensions and instruments of the SAWA reform

Reform Elements	Policies	Definition
A. Autonomy: decision-making power is decentralized to the school level	1. Schools make decisions about curriculum	Principals or school boards are granted decision-making power over what content is taught. This includes the ability to disregard or not follow the national / state curriculum or learning standards. Curriculum autonomy also means the principal or school board has power over the organization of curricula (i.e., decisions over what subjects are taught or the number of hours assigned to different courses, such as more instructional time dedicated to maths over art education).
	2. Schools make decisions about staffing	Schools' principals or boards are granted decision-making power over teacher hiring or firing policies. This could mean total control over hiring or firing or participation in collegial bodies in charge of teachers' allocation decisions.
	3. School makes decisions about budget allocation	Principals or school boards are granted decision-making power over funds allocation. Funds can be allocated to school infrastructure purposes (i.e. improvement of certain areas), purchase of equipment (i.e., computers, pedagogical resources, books, etc), or participation in extra-curricular activities or programs. This ranges from total control over the school budget –being allocated directly to a school bank account– or partial control, with decision-making constrained to certain areas.
	4. School makes teaching-style-related decisions	School principals or teachers can freely decide on the teaching / pedagogical approach in the school (e.g., project-based learning, student-center pedagogy, traditional lecturing/teacher-center pedagogy, etc)
B. Accountability & Standardization	4. Standardized national exams or sub-national when education governance is decentralized (i.e. US) in primary education (UNESCO ISCED 2011 level 1)	The country has a national standardized large-scale assessment in primary education on any subject (mathematics, reading, science, etc). Assessments can be sample-based or census-based. The periodicity of the assessment can vary, being yearly, every two or three years.
	5. Standardized national exams or sub-national when education governance is decentralized (i.e. US) in secondary education (UNESCO ISCED 2011 level 2 or 3)	The country has a national standardized large-scale assessment in secondary education on any subject (mathematics, reading, science, etc). Assessments can be sample-based or census-based. The periodicity of the assessment can vary, being yearly, every two or three years.

Reform Elements	Policies	Definition
	6. Standardized exams to exit secondary education level (accreditation function) (ISCED 2011 level 2 or 3).	Students must sit and approve a standardized assessment in specific subjects to obtain their degree at the respective level (different from exams to enter higher education).
	7. Exams results are publicly available at the school level	The government makes standardized results for primary or secondary publicly available through online platforms. Important clarification: it is not enough if school members can access results, it has to be publicly available.
	8. National learning standards/curriculum for students	The national government establishes a centralized national curriculum or academic learning standards, defined as minimum required content/skills expected for all students to learn.
C. Marketization & Competition	9. Charter schools / fully-subsidized privately managed schools	The state—either at the national or sub-national level—fully funds private providers to manage schools (e.g., the Charter model in the U.S., Academies in the U.K., Colegios en Conseción in Colombia).
	10. Per-capita funding	School funding in allocated on a student per-capita basis. This is, the national government transfers funds to education administrative bodies (i.e., states, municipalities, boards or schools) according to the reported number of students. This method contrasts with other schemes where funds correspond to the number of teachers in schools or population in school-age estimations.
	11. State subsidies to private schools	The national government gives supply-side subsidies to private schools, reducing operating costs. These schools can charge fees to families. Often, these subsidies are argued on the basis of a lack of public offering in a given area or due to larger demand than available seats in state-managed schools.
	12. School Choice	Families can choose what school to send their child either in the public sector (normally a given for private schools). This means that the government does not impose restrictions on what schools families can choose.
D. Performance-based incentives*	13. Merit-based payment depends on students' performance on national or regional assessments	Teachers' and/or principals' salaries are subject to student performance on national assessments. This could mean just a small portion of their salary (e.g., 30% of their total compensation is variable depending on results), or an additional payment/bonus will be

Reform Elements	Policies	Definition
		disbursed based on the achievement of certain pre-determined results. Improvement can be measured as a percentage of variation from prior year (s) or another baseline metric or as the achievement of a minimum performance threshold.
	14. Economic rewards to the best performing schools.	Schools showing good results in national standardized assessments obtain financial rewards from the national government. Economic rewards can range from salary bonuses for teachers/principals; or any type of additional funds for the school. Improvement can be measured as a percentage of variation from prior year (s) or another baseline metric or as the achievement of a minimum performance threshold.