INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Ciências
ULisboa

Structural inequality in education: a network science approach

Harvey Spencer Sánchez Restrepo

PhD in Complexity Sciences

Supervisor:
Doctor Jorge Manuel Anacleto Louçã, Full Professor,
Iscte – Instituto Universitário de Lisboa

December, 2025

Department of Information Science and Technology

Structural inequality in education: a network science approach

Harvey Spencer Sánchez Restrepo

PhD in Complexity Sciences

Jury:

Doctor Luís Eduardo de Pinho Ducla Soares, Full Professor at Iscte - Instituto Universitário de Lisboa (President)
Doctor Jeffrey Horace Johnson, Full Professor at The Open University (England)
Doctor Pablo Padilla Longoria, Full Professor at Universidad Autónoma de México
Doctor Rui Jorge Lopes, Associate Professor at Iscte - Instituto Universitário de Lisboa
Doctor Jorge Manuel Anacleto Louçã, Full Professor at Iscte – Instituto Universitário de Lisboa

December, 2025

A mis padres, Rocío y Arturo, por el talento y el amor con que sembraron y llenaron de flores ambos lados del camino. A mi abuela Irma, por enseñarme a persistir, a avanzar con brújula propia y que el tiempo para sanar no se escatima. A mis hermanos, Cris y Alexie, por recordarme que la guadaña y la razón se sujetan siempre con dos manos. A MG, por recordarme que el sendero siempre tiene un lado terso.

Acknowledgments

This thesis marks the end of a cycle that began as an anticipated adventure, one of those that inflates to recall its fictions. and deflates to not forget its edges; imagined like an albatross at sea, with broad wings to fly out the window if the doors conspired to corner dreams, or if the clouds signalled the time still left to sleep.

I am grateful to life for its unexpected turns, because during the pandemic, it allowed me, once again, to design a large-scale, unfiltered evaluation system to closely study the phenomena that have guided my steps throughout my whole life.

I would like to express my special gratitude to my supervisor, Jorge Louçã, for his generous support, patience, and friendship, which sustained my commitment through to the completion of this process, and with whom I trust I will develop a lifelong friendship.

I also express my gratitude to Portuguese society, the ISCTE-IUL, and the Information Sciences, Technologies and Architecture Research Center (ISTAR), for the support received through the *Bolsa de Investigação da FCT* (reference: 4466), under the *Doctoral Program in Complexity Sciences*. Thanks to this opportunity, I was able to delve deeper into a project that transcends the technical and theoretical aspects of education and go a little further, touching on various aspects of the human condition.

I know it is necessary to continue countering the socioeconomic determinism that hits children born in latitudes historically plundered or, in modern times, devoid of everything—even of meaning. For this reason, I also thank all those who fight for a public education that promotes awareness and freedom to combat oppression, as it is one of the few ways to alleviate the burden borne by those who lack opportunities in their childhood. From my own experience, I am convinced that Education is a crack through which it is possible to escape, to dream, and to thrive in freedom.

Finally, I would like to thank Tere Bracho, Rodrigo Jiménez, Luis Chias, David Gualpa, Belén Carmona, Adriana López, Elizabeth Alcantar, Luz Macías, and all my friends from Aleph. They remind me that it is always possible to bring together a considerable number of people to embark on adventures with a collective identity, because they know that an education that enables lifelong learning is the only way to avoid becoming modernly disposable, and because they have confirmed that, together, we do it better!

Thank you all for helping to keep this little flame alive.

Resumo

Compreender os determinantes do direito humano à educação exige analisar relações complexas, multiescalares e não lineares. Para a formulação de políticas baseadas em evidências, os governos e as partes interessadas analisam conjuntos de dados maciços com milhões de medições provenientes de avaliações internacionais e nacionais de grande escala, o que demonstra claramente os desafios desta nova era de precisão na investigação educacional. No entanto, a antiga tradição de utilizar modelos lineares convencionais, embora útil, também tem ignorado retroalimentações sistémicas e interdependências entre fatores sociocognitivos, socioemocionais e contextuais. Por estas razões, tal como na maioria das disciplinas sociais, os decisores políticos na área da educação começaram a exigir ferramentas e abordagens mais sofisticadas para obter uma compreensão mais profunda e detalhada do impacto e da variabilidade dos processos subjacentes aos fenómenos educacionais complexos. Para enfrentar estes desafios, esta investigação integrou equações de processos estocásticos com modelos de redes, permitindo uma medição mais abrangente das lacunas socioeconómicas, de género e de desigualdade, utilizando ferramentas não lineares para incorporar estimativas estatísticas isoladas em modelos mais avançados e informativos.

Com esta abordagem inovadora, estabelecemos três marcos para a investigação educacional. Em primeiro lugar, desenvolvemos um quadro analítico não determinístico para reformular e estimar a métrica internacional da pobreza de aprendizagem, principal indicador do Objetivo de Desenvolvimento Sustentável 4. Ao atualizar fórmulas determinísticas para um conjunto de equações estocásticas inter-relacionadas, definimos e desenvolvemos processos estocásticos específicos para medir o índice absoluto, relativo e composto de privação de aprendizagem, combinando a profundidade qualitativa e o rigor quantitativo. Isto representa não apenas um refinamento técnico, mas uma mudança epistemológica essencial para a implementação oficial a nível mundial. Em segundo lugar, esta tese oferece várias vantagens na modelação de redes académicas baseadas na correlação de fatores, revelando algumas estruturas da privação e permitindo estimar os efeitos variacionais dos parâmetros da rede e as suas relações com o estatuto socioeconómico em diferentes escalas. Finalmente, esta investigação identificou e mediu fenómenos emergentes resultantes das interações, como resiliência académica, lacunas de desigualdade e distribuição assimétrica dos resultados de aprendizagem. A gama de características topológicas encontradas possui um enorme potencial, pois estabelece pontes que conectam diretamente conceitos educacionais e pedagógicos com resultados quantitativos, fornecendo uma estrutura robusta para a construção de cenários e simulações dinâmicas através da calibração e ajuste de parâmetros de ordem e controlo. Esta ferramenta conceptual e instrumental poderá ajudar decisores políticos e partes interessadas a enfrentar muitos desafios em larga escala para garantir uma educação inclusiva e equitativa para grupos marginalizados, fornecendo informações acionáveis para a conceção de políticas inclusivas e responsáveis.

Abstract

Understanding the determinants of the human right to education requires analyzing complex, multiscale, nonlinear relationships. For evidence-based policy, governments and stakeholders analyze massive datasets with millions of measurements from international and national large-scale assessments, highlighting the challenges of this new era of precision in educational research. However, the tradition of using conventional linear models, while useful, has also overlooked systemic feedback and interdependencies among socio-cognitive, socio-emotional, and contextual factors. Consequently, as in most social disciplines, education policymakers have begun to demand more sophisticated tools and approaches to gain deeper, more detailed insights into the impact and variability of the underlying processes of educational phenomena. To address these challenges, this research integrates stochastic process equations with network models to enable more comprehensive measurement of socioeconomic, gender, and inequality gaps by using non-linear tools to incorporate isolated statistical estimates into more informative models.

Through this novel approach, we set three milestones for educational research. First, we developed an analytical, non-deterministic framework to reformulate and estimate the international metric for learning poverty, the main indicator for education and the Sustainable Development Goal 4. By upgrading deterministic formulas to a set of interrelated stochastic equations, we defined and developed specific stochastic processes for absolute, relative, and compound indices of learning deprivation, combining a global perspective with both qualitative depth and quantitative rigor. This represents not just a technical refinement but an essential epistemological shift for official implementation worldwide. Second, this thesis offers several advantages in modeling the correlation-based academic networks of factors, unveiling structural drivers of deprivation while estimating the variational effects of network parameters and their relationships with socioeconomic status across scales. Finally, this research identified and measured emerging phenomena arising from interactions, such as academic resilience, inequality gaps, and asymmetrical distribution of learning outcomes. The wide range of topological features identified has enormous potential, as it establishes bridges that directly connect educational and pedagogical concepts with quantitative results, providing a robust framework for building dynamic scenarios. This valuable conceptual and instrumental tool might help policymakers to address many large-scale challenges to ensure inclusive and equitable education for historically marginalized groups, providing actionable insights for inclusive and accountable policy design.

Table of Contents

Acknowledgments ..... **iii**

Resumo ..... **v**

Abstract ..... **vii**

List of Abbreviations ..... **xi**

Introduction **1**

Chapter 1. Literature review and context **5**

1.1. Sustainable Development Goals 5

1.2. Education as a Complex System 7

1.3. Measuring Learning Outcomes 12

1.4. Network Analysis of Academic Inequality 18

1.5. From Nodes and Edges to Metrics and Topology 21

Chapter 2. Topological Properties of Inequality and Deprivation in an Educational System: Unveiling the Key-Drivers Through Complex Network Analysis **29**

Chapter 3. Inequality in Learning Outcomes: Unveiling Educational Deprivation Through Complex Network Analysis **37**

Chapter 4. A Complex Network Approach to Structural Inequality of Educational Deprivation in a Latin American country **51**

Chapter 5. Educational Deprivation in Latin America: Structural Inequality Beyond Borders **63**

Chapter 6. Educational Gaps in Latin American childhood: Regional Inequality in Learning Deprivation with Local Modulations **71**

Chapter 7. Topological Substitution: Mapping the Structural Mechanisms of Academic Resilience via Complex Network Analysis **79**

Chapter 8. Topological Patterns of Academic Success: Insights from Complex Network Analysis **103**

Chapter 9. Conclusion and Perspectives **119**

References **123**

List of Abbreviations

COVID-19 – Coronavirus disease caused by the SARS-CoV-2 virus

EBP – Evidence-Based Policy

ET – Ethnic self-identification

GAML – Global Alliance to Monitor Learning

GE – Gender

HLM – Hierarchical Linear Models

LD – Learning Deprived

LDI – Learning Deprivation Index

LP – Learning Poverty

LSA – Large-Scale Assessments

OECD – Organisation for Economic Co-operation and Development

PISA – Programme for International Student Assessment

RA – Rurality

RAD – Rate of Absolute deprivation

ReD – Relative Deprivation

RV – Random variable

SD – Schooling Deprived

SDG – United Nations Sustainable Development Goal

SDG 4 – United Nations Sustainable Development Goal 4

SES – Socioeconomic Status

SOA – Society of Actuaries

SP – Stochastic process

TIMSS – Trends in International Mathematics and Science Study

TS – Type of school

UNESCO – United Nations Educational, Scientific and Cultural Organization

UNICEF – United Nations Children's Fund

Introduction

The probability that each person will achieve a minimum level of dignity and a sufficient level of influence in the decisions they make throughout their lives, as well as the accumulated prosperity they manage to satisfy their needs and guide their efforts towards fulfilling their dreams and desires, depends, essentially, on two colossal forces: *chance* and *determinism*. The first begins to operate from the conception of a new human being and actively participates in the definition of the psycho-biological characteristics, phylogenetic expression, zip code, and circumstances of birth of each of us [Sánchez-Restrepo, H. et al, 2018]. Given the initial conditions, the second force is detonated with profound intensity across a wide range of scenarios, generating an unequal distribution of probabilities for keeping alive, growing up healthy, or developing capacities, abilities, or freedom to learn and choice, among other living conditions, remaining conditioned by the initial endowment and context for the whole life. This simplified representation of the intricate social system that humanity has constructed over its evolutionary history is grounded in the establishment of hierarchies through cumulative and feedback processes, which often lead to multi-order phase transitions [Solé, R., 2011], the essence of what we usually call "complex systems," in which the evolution of the "measurable results" almost always refers to the observable parameters of the system at its macroscopic scale, produced mainly by the emergence of collective properties from nonlinear interactions at lower scales [Morin, E., 1992; Barabási, A. L., 2003; Davis, B., & Sumara, D., 2014; Liu, Y. Y., & Barabási, A. L., 2016; Hébert-Dufresne, L., et al, 2024].

Education can also be conceptualized as a complex adaptive system, characterized by multiple interacting components operating across different levels—individual, institutional, and societal, most of them exhibiting non-linear dynamics, where small changes in one component—e.g., socioeconomic status, study of habits or metacognition—can produce incommensurable effects on sociocognitive or socioemotional outcomes [Morin, E., 2007; Mason, M., 2008]. Furthermore, educational inequality, learning poverty, relative deprivation gaps, and resilience can be seen as emergent properties arising from the dynamics of many agents that receive, promote, and participate in cultural, social, economic, and pedagogical subsystems and related dynamics.

Most of the literature shows that an educational system is a highly interlinked convergence of cultural, social, economic, and pedagogical systems. Modelling different layers of educational systems as networks—where nodes correspond to entities such as students, schools, or policies, and edges represent relationships or dependencies—researchers can uncover hidden

structures and patterns related to educational and learning outcomes [Pósfai, M., & Barabási, A. L., 2016], providing several advantages over classical statistical approaches. First, it enables the identification of topological properties such as clustering, centrality, and modularity, which reveal how educational opportunities and resources are distributed. Second, it opens a new set of heuristic tools for detecting non-trivial relationships among social determinants—e.g., ethnicity, socioeconomic status, rurality—on academic performance, offering insights into systemic vulnerabilities, propensity to learning deprivation, and resilience [Newman, M., 2018; Gao, J., Barzel, B., & Barabási, A. L., 2016].

This research departs from an increasingly general idea: classic econometrical tools are both, useful and insufficient [Johnson, J. H., 2016; Brinkhof, L. P. et al, 2025], a statement the probably comes from its almost exclusive reliance on linear regression models, which automatic use has been leading to a fragmented view of the highly interrelated educational ecosystem, and its wide range of complex educational phenomena [van der Maas, H., 2017]. As simplified econometric models often fail to capture the deep-seated, structural nature of inequality in learning outcomes and their inherent relationship with socioeconomic and educational deprivation, all its methodological limitation—the inability to model complexity, non-linearity, and interactions— appears as the primary reason for the observed lack of explanatory power in many educational policy models [Angela Espinosa & Jon Walker, 2017; Johnson, J. H., Fortune, J., & Bromley, J., 2017].

The central motivation for this project arose from the persistent difficulty policymakers face in recognizing the determinants of learning poverty and their low capacity for identifying both main issues: (1) the structural factors for learning deprivation, and (2) the key drivers of learning in adverse conditions; losing the opportunity for provoking academic resilience in socio-economically disadvantaged students, the crucial challenge for low- and middle-income countries to make the United Nations Sustainable Development Goal 4 (SDG 4): *"Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all"* [UNESCO, 2020], a reasonable utopia.

After years of interdisciplinary research, bringing together experts from diverse fields, including education science, pedagogy, economics, sociology, geography, statistics, applied mathematics, informatics, and data science, among others, this project now stands at the intersection of educational management, stochastic processes, and network science, offering a novel foundation for developing evidence-based policy through the lens of complex systems to provide actionable insights that might have the potential to transform educational policy design, reduce learning poverty, and promote equity in education systems worldwide.

Research Objectives

The primary objective of this thesis is to model the structure of the asymmetric distribution of learning deprivation among students across socioeconomic and cultural levels—also known as structural inequality in education—using complex network analysis. The proposed approach encompasses a broad conceptual and instrumental framework for analyzing topological and statistical properties of networks representing a selection of educational systems with elevated levels of inequality. This analysis identifies some of the numerous underlying mechanisms and thresholds for deprivation, segregation, and resilience at multiple scales within Latin American educational systems—one of the most unequal regions on the planet. To achieve this, the research was structured around three specific objectives:

1. To model learning deprivation gaps and the structure of inequality as a network. (Chapters 2 and 3). This objective focuses on validating the use of network science to represent educational systems, demonstrating how inequality behaves as a system property rather than an individual attribute.
2. To analyze the structural inequality of educational deprivation. (Chapters 4 and 5). This objective expands the analysis beyond national borders, using UNESCO's large-scale assessment data to map the non-linear interdependence among learning, cultural capital, and the socioeconomic status of students' families across all Latin American countries.
3. To identify the structural mechanisms of academic resilience. (Chapters 6 and 7). Moving from the macro to the micro scale. This objective seeks to understand how successful students in high-poverty contexts navigate the system, hypothesizing that resilience is a strategy of "topological substitution."

Thesis Structure

To present advances and results in sequence, the thesis is organized as a compendium of seven peer-reviewed articles, progressing from general to more advanced and realistic models. Chapter 1 presents a brief literature review, a few crucial concepts, and some notation details to facilitate better reading and understanding of the published articles. The following chapters (2–4) develop the core of the initial steps, focusing on models applied to real data, and establishing the fundamental topological characteristics of educational deprivation, using some network metrics to reveal the specific factors that highlight inequality at different scales. These

analyses primarily focus on three aspects that often determine poverty and learning deprivation: *socioeconomic status* (SES), student gender—male or female—, school financing type, and self-identification with an ethnic group. In the two subsequent chapters (5–6), this analysis is extended to comparative regional contexts to draw inferences for a set of Latin American countries, examining how local modulations affect national rates of absolute and relative learning deprivation.

Chapters 7 and 8 present the latest research, where we conducted a systematic analysis of interactions among sociocognitive, socioemotional, and 140 additional factors to identify a specific topological mechanism that distinguishes academically resilient and high-achieving students from their more privileged peers. The first one is a chapter of the book edited by *Springer Nature* named *Socioeconomic Studies of Today's Mexico: A Mathematical Overview*, which is already accepted and in its camera-ready version. The last one is an article entitled *Topological Patterns of Academic Success: Insights from Complex Network Analysis*, in its final version. These results were accepted to be published by *Springer Nature Complexity Conference Series* and presented at the *6th World Conference on Complex Systems*. This last part of the research is particularly relevant because it led to a novel networks' definition of resilience suggesting that it "should not be conceived as a stable personal attribute, but as a dynamic and adaptive process that arises from the continuous interaction of individual, family, and institutional factors [Wang, F., 2025; Sánchez-Restrepo, H., 2026; Sánchez-Restrepo, H. et al, 2026].

Finally, we offer a *Conclusion and Perspectives*, integrating the progressive and cumulative findings an expansion of this three-phase project reviewing methodological limitations and future research need, and a discussion about stakeholders perspectives and practical application of network science in education policy effectiveness, bringing arguments to confirm that the tools of complexity science are essential for reframing the analysis of educational inequality, aa well as its implications for quality and equity, especially in Latin American countries, which the political culture is a deep-lack in using evidences, and a marked dislike for the word "evaluation". In our opinion, this thesis structure ensures that each chapter contributes to improving the use of essential quantitative evidence to support the central argument: *educational policy must be quantified, and accountability should not be an option*.

CHAPTER 1

Literature review and context

1.1. Sustainable Development Goals

The exacerbated situation of both new and historical conflicts, combined with the almost invisible link between public policy definition and an evidence-based institutional framework for accountability, has consolidated an alarming situation across the planet: poverty has continued to increase dramatically in recent years [UNESCO, 2015; 2022; 2024; 2026; UNESCO & OEI, 2025]. Currently, more than 838 million people live in extreme poverty [World Bank, 2024], which means that at least one in ten individuals cannot afford even the minimum cost of food required for survival. For many, daily life is marked by uncertainty—about where they will sleep, whether they will eat, or if a common illness could become fatal within days. To address this challenging situation, as part of global agreements in the last decade, the United Nations (UN) declared the *Sustainable Development Goals* (SDG) as

> "An urgent call for action by all countries—developed and developing—in a global partnership. They recognise that ending poverty and other deprivations must go hand-in-hand with strategies that improve health and education, reduce inequality, and spur economic growth" [UNESCO, 2015]

In education, the outlook is equally concerning. Current global estimates for the period 2019–2025 indicate that nearly 60% of all 10-year-old children worldwide—equivalent to more than 390 million—experience learning poverty [World Bank, UNESCO, & UNICEF, 2021, UNESCO, 2026]. The situation is significantly more severe in middle- and low-income countries, where over 70% and 90% of children, respectively, lack basic reading comprehension skills. All around the world, conditions for increasing learning poverty have been further exacerbated by prolonged school closures and learning disruptions caused by the COVID-19 pandemic [UNESCO Institute for Statistics, 2022; 2024; World Bank, UNESCO, & UNICEF, 2022; UNESCO & OEI, 2025], erasing two decades of educational progress. In light of this evidence, most countries began to reinforce programs, allocate untapped resources, and create or join to existing and new *Large-Scale Assessments* (LSA) to diagnose their educational systems and estimate gaps linked to socioeconomic, gender, ethnicity, or rural origin. Bad news

and reactions have been everywhere, and diverse social groups started to pressure governments to prioritize SDG 4, particularly Target 4.1: “By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes” [UNESCO, 2017]. Most countries publicly acknowledge the problem; their commitment has introduced a wide range of well-known and new challenges, from scaling educational systems' resources and increasing budgets for this sector and reaching universal access, to developing pedagogical strategies and didactic activities that intend to go beyond mere credentialing for truly improving learning outcomes. From our point of view, most of the countries are losing this battle.

1.1.1. The challenge of monitoring SDG 4

As a fundamental pillar of most modern societies; in all countries, Education shows significant disparities persisting across their territories and among its population groups, half of them with strong barriers in access to education and in the distribution of learning, increasing and deepening the intergenerational gaps between wealthier and less wealthy families, even in non-traditional academic dimensions of the students, as socioemotional skills, high expectations, joy, happiness, satisfaction with life or with the use of time [Hanushek, E. A., & Woessmann, L., 2020; Rudd, G., Meissel, K., & Meyer, F. ,2021; Marta, E. M., & Michelle, C., 2023]. Therefore, empirical research demonstrates that educational inequality is lifelong persistent, acting as a catalyst for a wide variety of social, cultural and economic inequities—shaping income, health, life satisfaction, and civic engagement [Blanden et al., 2023], becoming an important fuel for social and political conflict and for conditioning and predicting population health services and trajectories [Piketty, T., 2020], and producing a cascade of cumulative disadvantages [Hazy, J. K., & Ashley, A., 2011; Bloome, D., Dyer, S., & Zhou, X., 2018].

Despite being accompanied by a set of global indicators adopted by the UN General Assembly, most of which were developed after extensive consultations with various stakeholders worldwide through the *Global Alliance to Monitor Learning* (GAML) [UNESCO Institute for Statistics, 2022], official monitoring is not an easy process, as these indicators cover aspects such as primary and secondary education completion rates, early childhood development, and, in a fuzzier, not fully comparable way, the quality of education [Azevedo, J., 2020]. Although today, monitoring SDG 4 is a crucial bottleneck for breaking the cycle of poverty, reducing inequalities, and fostering sustainable development. Yet a relevant part of the discussion focuses on separating the percentage of children suffering learning poverty as a

simple dichotomic indicator for each child, showing a relevant lack of conceptual and empirical indicators about the magnitude of the students who were not learning deprived but became learning deprived, the depth of the already learning deprived, and those new students who will be. To implement and deploy actions in a timely manner, governments need valid, reliable, and timely data, information, and knowledge that allows them to diagnose, focus, and monitor the effects of their strategies and their progress in every school, not just "averages and long-term trends" [Sanchez-Restrepo, H., 2021]. This approach—mainly known as *Evidence-Based Policy* (EBP)—is supposed to be supported by the best available scientific evidence. However, implementing EBP is not a simple endeavor, among other reasons, because it requires educators to navigate a vast array of studies, discern high-quality research, and apply findings appropriately across a wide diversity of classroom settings. The objective of developing better pedagogical strategies and more effective and specific didactic practices is quite challenging because it often lies in bridging the gap between academic research and practical teaching methods, ensuring that evidence-based practices are significant, accessible, and adaptable to the learner's needs.

1.2. Education as a Complex System

There are several styles for managing education systems; the choice is not free and is frequently dictated by the executive branch. While some authorities strive to manage the system's resources and focus institutional capacities on maintaining order and hierarchy within the bureaucracy—where governance is structured vertically on the power—others are preoccupied with remaining in office as long as possible for controlling teachers' unions ant other organizations—governance distributed by political leadership. However, it is increasingly common to observe that the evaluation of the success of an education system's management is based on the analysis of improved indicators. In addition to SDG 4 and its target 4.1, the most relevant indicators revolve around students' socio-cognitive and emotional development. Authorities who adhere to this governance style develop most of their policies based on evidence and are quite inclined toward accountability; they seem convinced that big achievements will have to wait.

Education possesses prominent levels of inertia; all transformations processes require sustained, middle, and long-term efforts. For obvious reasons, modern educational research is dedicated to improving evidence about the system and developing EBP. Unfortunately, its effectiveness depends heavily on the scale, timeframe, and metrics adopted. While quantitative

data provides valuable information and addresses qualitative aspects at different scales, most successful outcomes will arrive in the future. And although policymakers know this, most believe that results must be delivered as quickly as possible, and therefore choose easy, low-impact solutions, keeping the system consistently below its potential. This subjective reality exemplifies what J. Johnson (2016) states:

> "One of the challenges of modelling complex systems is to integrate micro and macro theories of behavior. From a policy perspective, the target systems almost always include people as social and economic systems. Social systems have structures at many levels, from the individual at micro levels through organizations, and institutions at meso levels to nations and international structures at macro levels.

Obviously, educational systems have different legal frameworks coming from different economic, political, and functional agreements; their characteristics come from the social contract subscribed by all those different stakeholders who have the opportunity and enough force to push rules and demand information about results in distinct aspects. In this research, the units of study start from all educational actors—micro level—, passing to the schools as units with independent, aggregated and emergent properties—meso level—, to the aggregated scales as natural or artificial partitions of population or territory—macro level— [Tokuhama, T. 2012; Tracey, B., & Florian, K., 2016]. Following what Jofrrey (2016) refers to "the challenge of creating coherent multilevel theories to support policy", a general model was developed for representing a multilevel education system, synthetized in Figure 1. This model satisfies the BEP approach and the use of LSA results for analyzing interactions among agents or factors—that might be operationalized through out data or information— and that there are loops and feedback across this scales that hold entangled the multiple subsystems, which is why some authors also call it a "system of systems" [Johnson, J. H., 2016; Johnson, J. H., Fortune, J., & Bromley, J., 2017].

1.1.2. Micro-level: students, families, teachers, and principals

The foundation of the system lies at the micro level, which includes students, families, teachers, and principals as individual agents, the main educational actors for EBP. At this scale, dynamics operate at two layers: their individual properties—variables that can be observed and assessed directly, and the expressions of their behavior and actions—constructs or latent variables that should be inferred indirectly. Within this unit, agents and variables used to be the nodes, and

their edges represent an observation of the phenomena under study. As can be seen, students become the primary source of information, which is truly fortunate, since they are also the main objective of the entire system.

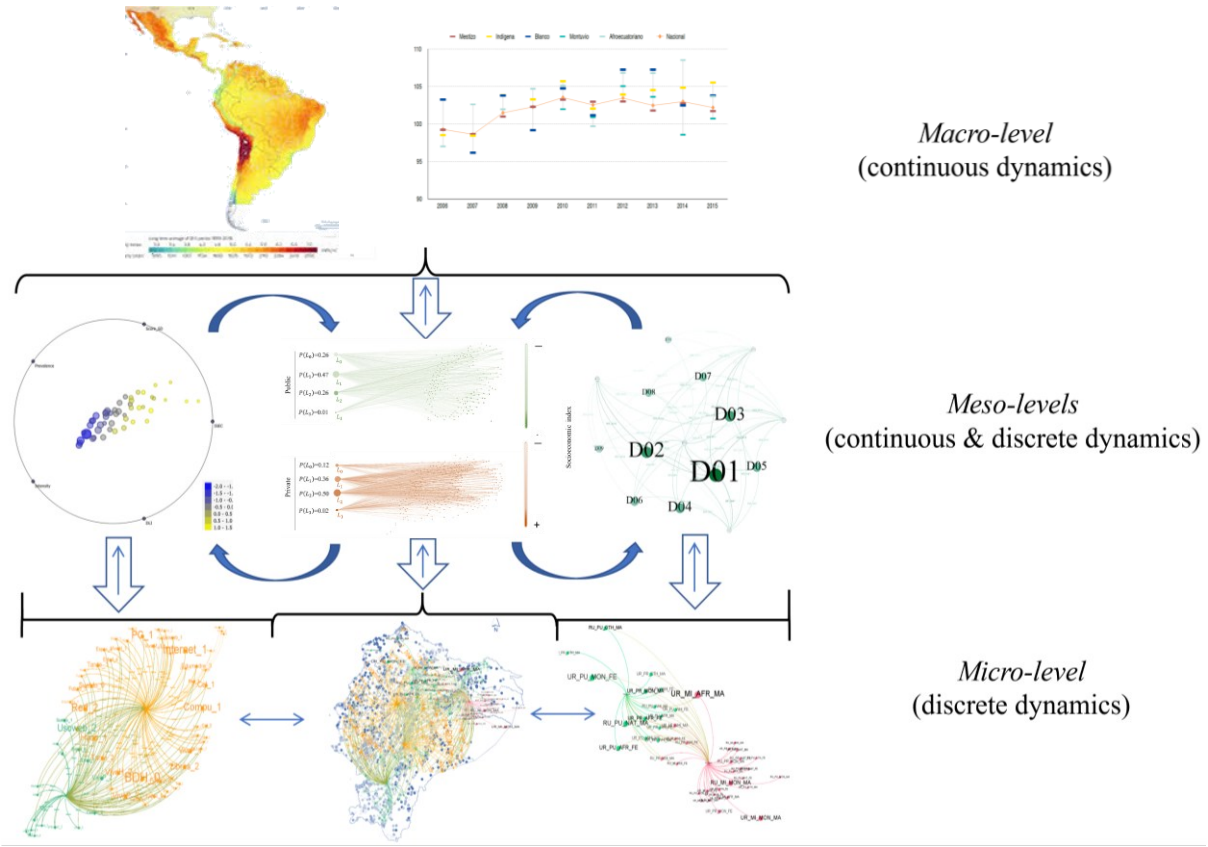


Fig 1. Multilevel educational system, Sánchez-Restrepo, H. (2020) based on Johnson, J. (2012).

For instance, some learners effortlessly absorb new material, while others wrestle with assignments or disengage. When looking for an explanation for high-academic achievers, exploring interactions and the structure of a wide set of variables for detecting if those associated with success are truly relevant and in which way is a much more dynamic and richer approach than just estimating low bivariate correlations among all pair of variables. Large-scale datasets—whether from PISA, TIMSS, or national assessments—attempt to map these behavioral differences using statistics, even identifying patterns such as learning deprivation or drop-out. Studying aggregated learning outcomes and student profiles is particularly insightful for carrying on network analysis. For instance, our three first publications were focused on these aspects—chapters 2, 3 and 4, especially when doing variational analysis, because small support groups form organically, or, conversely, peer networks unfortunately consolidate academic struggles. As in any model of social systems, in analyses at the micro scale, the

analysis of economic considerations must integrate individual constraints, costs and benefits, which often depend on the parameters or references observed in social aggregates. As dynamics arise from individual interactions, in which each agent has their own behaviors and degrees of freedom, the observable traits and their educational outcomes tend to be similar, with little intraclass variance, a condition largely motivated by the context and the social, economic, and cultural conditions that surround them.

1.1.3. The Meso-Level: Schools

In the meso level, the central units are schools but can integrate teachers and principals as decision makers or drivers of the system at that level. Their dynamics are both continuous and discrete, as many properties emerge from aggregated characteristics at the micro level and often depend on decisions made by those agents within that level. It is easy to see that schools are far more than administrative units because they function as complex adaptive systems shaped by culture, relationships, and routines with a small number of degrees of freedom [Nowosad, I., 2021].

Complexity often emerges from horizontal interactions—agents operating at the same level— for three mains sources: (1) *high density of connections*, due to the interaction among students, teachers and all other members of the scholar community; (2) *local decision-making*, given that agents at the same level often share similar goals but compete for resources or influence with different strategies and backgrounds; and (3) *distributed autonomy*: no single agent controls the system; instead, patterns arise from many small decisions interacting simultaneously. The convergence in time and space of these conditions create feedback loops that, in one hand, amplify variability and uncertainty, leading to nonlinear dynamics, and, in the other hand, produce emergent properties such as school culture, equity or inequality patterns, innovation diffusion or creativity—understood as the ability to think outside the box, make connections between seemingly disparate ideas, and question norms [Mueller, J. S. et al, 2014]—conditioned by micro-level collaboration and resistance of individuals and the capacities of the network for propagation. As can be seen, all these properties cannot be designed top-to-down, but they arise from bottom-to-up interactions and sustained by feedback across levels [Johnson, J. H., Fortune, J., & Bromley, J., 2017].

In chapters 3 and 4, some of these collective properties were analysed establishing that "when schools are studied as integrated units, the impact of socioeconomic status in education becomes even more evident, as a categorical class-property" —for a more clear and vast

explanation of the effect of peer socioeconomic status on student achievement, we suggest to review the Van Ewijk, R., & Sleegers, P. meta-analysis (2010).

For instance, the rate of self-adherence to ethnicity and the type of school financing emerged as key elements for segmenting student groups, highlighting that gaps between rural and urban areas are particularly pronounced—especially among the poorest students in rural regions, who experience the highest levels of deprivation. These findings helped to identify the key factors for developing policies targeted at these schools and the students they serve [Finnigan et al., 2018; Verger, 2018]. Therefore, any strategy aimed at meaningful change at the school level must consider the topology and structure of the local network when designing school profiles for intervention or ranking factors from maximum centrality to minimal relevance.

1.1.4. The Macro-Level: Strata and Territorial partitions

At the macro level, the system's observable properties are continuous and can be interpreted as order parameters that describe its overall state. This is at the broadest level of observation, and where the emergence of structures within systems is most clearly detected as one move from micro- and meso-scale to macro-scale. Since macro-level states are typically represented by measures derived from aggregating, weighting, or varying the Extract-Transform-Load (ETL) process for structured data from lower levels, they also serve as the substrate for producing indicators and modeling networks with nodes representing variables or states at this scale.

Again, this approach offers many advantages over classical linear models. First, it offers a general perspective for monitoring and analyzing sequences to show a longitudinal story, connecting in the wave past, present, and future, enabling stable references for variational analysis and simulation, with their subsequent comparison on the system's evolutionary trajectories. In this sense, it is valuable to hold the J. Johnson (2016) statement: “simulations do not offer 'predictions' of the future behavior of the system, but rather provide information about its possible behavior and, at best, estimate the probability that a policy will achieve the desired outcome at that level.” Secondly macro-level forces—driven by preassigned funding decisions, fluctuating political priorities, and local or global socioeconomic divisions—often exert profound effects on support opportunities, motivation, and learning and teaching capacities at the meson and micro levels [Gross, Meyer, and Hadjar, 2016], giving rise to feedback that often reparametrizes, and even reconfigures, the entire system.

One of the greatest contributions of network models is the identification of ordering parameters that can help policymakers take on greater responsibilities in more complex, higher-impact areas to improve educational equity and reduce learning deprivation. The core difficulty lies in anticipating and provoking variations within the system through intentional actions. Achieving genuine systemic improvement, as envisioned by SDG 4, requires confronting complex, uncomfortable structural realities that network analysis can reveal—such as the strong covariability between socioeconomic status (SES) and positive academic outcomes, or the stark differences in network topology between the poorest and wealthiest schools. In this sense, analyzing the topologies of educational networks can help unveil key drivers of the system or structural properties that reveal some of the system's failures, as demonstrated in the studies referenced in chapters 7 and 8.

1.3. Measuring Learning Outcomes

Measuring the quality of the system is equivalent to measuring at which level its outcomes meet the standards the system has planned for itself. Academic performance is widely seen as the primary indicator of educational success and has been widely used in the academic resilience literature [Martin, A. J., & Marsh, H. W., 2006; Rudd, K. L., & Yates, T. M., 2020; Wang, F., 2025]. As mentioned, efforts to measure compliance with SDG 4 only began at the end of the last decade. The overall indicator was defined as Learning Poverty [Azevedo, 2019] using a completely deterministic approach with significant theoretical and practical limitations. Given the global adoption of the Learning Poverty indicator (LPI), this research reformulates it in parallel using stochastic processes theory and then adds two highly informative indicators. In this way, the expected value of the process of counting students experiencing relative learning deprivation coincides exactly with the deterministic estimates. However, its generalization within a stochastic framework offers numerous statistical, simulation, and predictive advantages, as well as theoretical interpretation advantages, that the deterministic framework lacks.

1.3.1. Learning poverty

The construct *Learning Poverty* (LP) refers to “The ability to read with comprehension [as] a foundational skill that every education system around the world strives to impart by late in primary school […] as it's essential for all future learning and development” [World Bank, UNESCO & UNICEF, 2022]. The indicator is an additive measure combining the proportion

of primary-aged children who are out-of-school (*Schooling Deprived*, SD) with the proportion of students in school who perform below minimum proficiency levels in reading (*Learning Deprived*, LD). In the context of the SDG 4, according to GAML (2021), learning deprivation is defined as share of children at the end of primary who read at below the minimum proficiency level [Azevedo, J. P., 2020]. This combined measure effectively captures the dual imperative of educational policy: ensuring "more schooling" —access—and guaranteeing "better learning" —quality—, thus confirming that time spent in the classroom translates into essential skills and capabilities.

Following the definition of LP, students' literacy abilities are seen as the main students' learning outcomes. For this reason, they are analyzed as dependent variables of a wide variety of academic and non-academic variables collected for the LSA's. According to Bracho (2017), for measuring the level of learning deprivation it is necessary to measure it in two ways: absolute and relative.

The first concept proposes that 'there is an irreducible nucleus of needs that are common to every human being.' The distribution of students in learning deprived and non-learning deprived—expressed as rates—serves as a key diagnostic indicator of educational equity and helps to highlight gaps within the system [Sánchez-Restrepo, H. & Carmona-Soto, M.B., 2023]. Figure 2 shows the breakdown of the LP indicator into its two components, SD and LD. Children in the SD category are estimated by official student records and compared to the total population of children of the age corresponding to that grade level, according to established norms. LD, on the other hand, refers to students and depends directly on each student's skills, θ_i. To determine if a student suffers from learning deprivation, it is necessary to establish the cutoff point ψ_1 to define the skills required to progress beyond level L_1. All students below this cutoff point are in L_1, meaning they exhibit significant learning deficits that hinder continued educational growth.

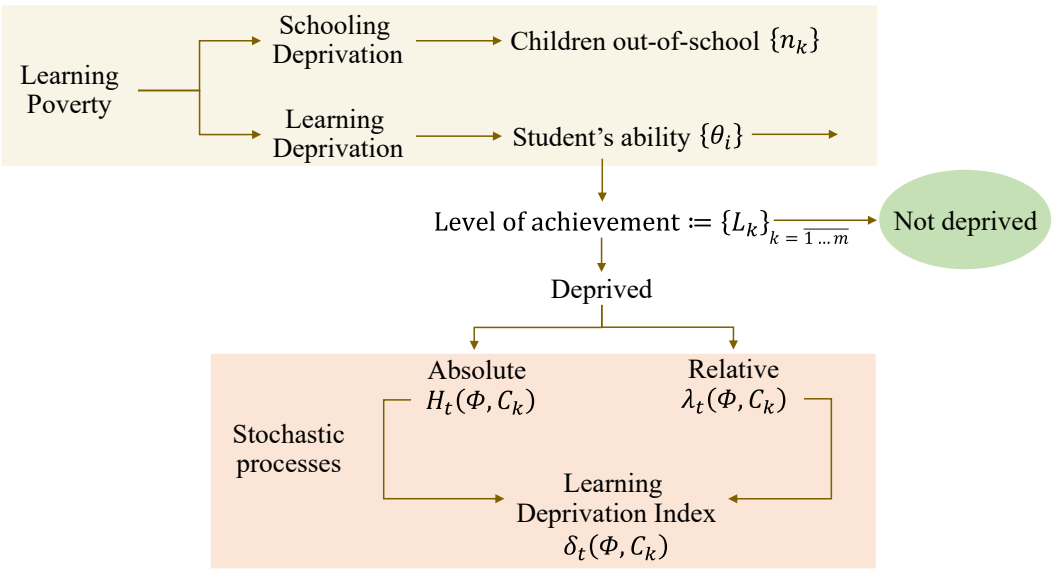


Fig. 2. Relationship between the indicators developed to measure learning deprivation, Sánchez-Restrepo, H., & Jorge Louçã, 2018; 2025), based on Bracho, T. (2017).

Once the ability level θ_i of all students has been estimated and the respective achievement level L_k has been assigned to each of them, there is a bifurcation that separates the not-learning deprived from the learning deprived. From the latter, two additional independent stochastic processes are triggered: (1) *Rate of Absolute deprivation* (RAD), and (2) *Relative Deprivation index* (ReD), which are combined to give rise to a composite stochastic process whose mathematical expectation is called *Learning Deprivation Index* (LDI).

1.3.2. Absolute deprivation

Initially, the key indicator for measuring learning poverty was conceived as a deterministic index coming directly from registers as the quotient of those students below the threshold of "enough learning" and total number of registered students, but there are, at least, two problems with this definition: (1) the number of students experiencing learning deprivation remains unknown until having LSA results, and (2) from a deterministic point of view, LSA just give a estimated number coming from a sample for one single point on the time. Given the theoretical and empirical limitations of this definition, since the beginning of this research we improved and generalized it as a stochastic process (SP), using the *Second Principle of Stochastic Modeling* of the *Society of Actuaries* (SOA), regarding statistical approximation models:

> "A phenomenon that exhibits statistical regularity can be described through a mathematical model that estimates, within any degree of accuracy, the probability of a given event in a

sufficiently large sequence of experiments. A deterministic model is a simplified stochastic model in which the proportion of occurrences of an event, estimated by the stochastic model, is assumed to have a probability of one." [Pinsky, M., & Karlin, S., 2010]

Thus, the prevalence of absolute deprivation was reformulated as the sum of random variables (RVs), where independence among students can or not be assumed since, in general, students' abilities are shaped [mostly] by nature, heredity, and individual context [Plomin & von Stumm, 2018; Deary, Cox, & Hill, 2021; Buser et al., 2023]. Then, for students suffering learning deprivation, the quantity $N_t(\Phi, L_1, C_k)$ is formulated in the same way for all chapters but is described in detail in chapter 7, where is properly defined as a sum of Bernoulli indicators $I_{i,t}(\theta_i; \Phi, L_1, C_k)$, estimating the occurrence of the deprivation event for the i-th student, for whom the single source of randomness is the ability θ_i, which distribution function is conditioned by its partition and boundaries given by (Φ, C_k) and where $\Omega(\Phi, C_k)$ refers to the universe of students bounded by the subset (Φ) and the community C_k [Sánchez-Restrepo, H., & Jorge Louçã, 2018; 2021; 2025]. In this sense, given that $\Omega(\Phi, C_k)$ is the total population of students in the community C_k and $n_t(\Phi, C_k)$ is the expectation of this RV, then $p_{L_1}(\Phi, C_k)$ is the probability of any student in the subset C_k might be learning deprived, following a general equation for estimating the function $H(\Phi, C_k)$, wich mathematical expectation is the unbiased estimative of RAD—the expected proportion of students in L_1—is given by the expression $H(\Phi, C_k) = \mathbb{E}[\mathcal{N}_t(\Phi, C_k)] \cdot \mathbb{E}[\mathbb{I}_t(\theta_i; \Phi, L_1, C_k)]$.

Figure 3 shows the estimates of RAD for each SES-decile in data coming from an official UNESCO's Latin American LSA named *Comparative and Explicative Regional Study* (ERCE, for its Spanish acronym), carried on with fifteen participating Latin American and Caribbean countries. In the left graph, it can be seen the profound inequality in the distribution of learning levels among socioeconomic groups, showing a *H*-function decreasing monotonously as the SES increases. In this comparison, numbers are really contrasting, while in the first decile, 63 of each 100 poor students do not learn the minimum ($H_{D01} = 0.631$), in the upper decile—the richest families— $H_{D10} = 0.099$, this means that, of every 100 students who do not develop elementary learning, 32 correspond to the first two deciles $(D01 + D02)$, while only 8 belong to the two highest deciles $(D09 + D10)$. On the contrary, of every 100 students in the outstanding group L_3, only 3 come from the deciles $D01$ and $D02$, while 55 come from the groups $D09$ and $D10$, 20 and 35, respectively, indicating the big gap between students coming from poor and rich families.

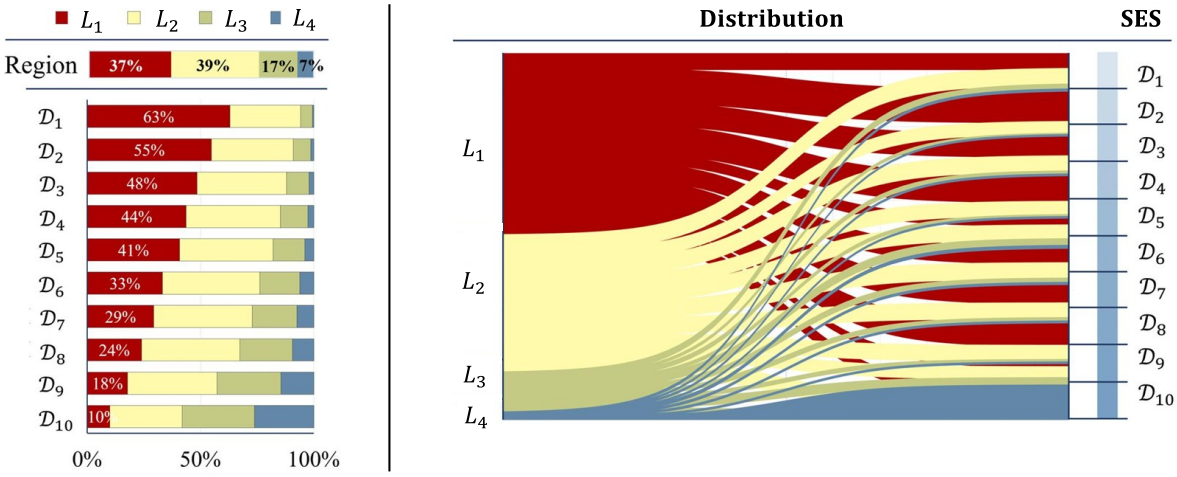


Fig. 3. Distribution of levels of achievement among SES deciles (left side). Relative flows of the distribution of levels of achievement among SES deciles.

1.3.3. Relative deprivation

After having a partition of a group of students learning deprived, a natural question was "how deep is the level of deprivation of each of them?", a concept linked with sociological proposal that states that "needs, thresholds and satisfactions are determined by each society" [Sen, 1982; Bracho, T., 2017]. Each deprived student, i.e., each student in the L_1-group, has a measure of the deepness of her/his deprivation given by the ReD index [Sánchez-Restrepo, H., & Jorge Louçã, 2018; 2021; 2025], which can be estimated through the distribution of the associated random variable (RV) measuring the intensity of the learning deprivation, a real value showing the gap between the ability' student and the first cut-off point ψ_1, a distance function of the ability θ_i considering the trajectory given by the Survival function (SF) $S_{\Lambda_i}(\theta_i; \Phi, C_k) = 1 - F_{\Lambda_i}(\theta_i; \Phi, C_k)$, also known as the complement of the probability distribution function (PDF), $F_{\Lambda_i}(s; \varepsilon_g, D_j, C_k) = \boldsymbol{P}\left[(\theta_i < s | L_1, \varepsilon_g, D_j, C_k)\right]$, where the inclusion of L_1 as a restriction of the measure space implies the boundary $\psi_0 < s < \psi_1$, bounded by the stratum ε_g, coming from the socioeconomic decile (SES-D_j), collecting all occurrences through the space of communities $k = 1,2,3 \ldots T$, where T is a finite number of communities among students suffering learning deprivation, all of them grouped by L_1.

1.3.4. Learning Deprivation Index

From the two measures explained in the previous sections, it was developed a synthetic index for representing learning deprivation of students from different strata, population groups, or territorial demarcations of interest. The LDI is a composition of *RAD* and *ReD* processes, i.e., a compound stochastic process (CSP) of both sources of randomness [Sanchez-Restrepo, H. & Jorge Louçã, 2019; 2025], which is equivalent to adding the intensity functions of all students in L_1, and weighting the result for all students grouped by $\mathcal{N}_t(\Phi; \varepsilon_g, D_j, C)$. The index $LDI_t(\Phi, c_k)$ is also denoted by $\delta_t(\Phi, c_k)$ for parametric analysis and is a synthetic measure of the total learning deprivation, observed in the student universe $\Omega(\Phi, c_k)$, at some point *t*, which can be estimated by aggregating the individual deprivation of each of the realizations, i.e., the expectation of the CSP. The LDI is a standardized and bounded metric, and it provides a consistent basis for comparing different subsets of students [Sanchez-Restrepo, H. & Jorge Louçã, 2019; 2025], its stability makes it particularly valuable for assessing the magnitude of learning gaps across time, territory, countries and regions.

For instance, if socioeconomic status is plotted on the horizontal axis—placing the poorest on the left—and LDI is plotted on the vertical axis as a function dependent on SES—placing those who do not suffer deprivation at the origin of the axe—, the dynamic relationship between these variables can be obtained to observe the composite phenomenon of inequality. As can be seen in Figure 4, the LDI could be extremely useful for detecting gaps between student groups following this procedure. First, calculate the average LDI for the entire population, the black horizontal line. Then, exclude the poorest 1% of the student population and recalculate the LDI using the remaining 99%, repeat the iteration excluding the next 1% and recalculate using the remaining 98%, and so on, until estimate the average LDI of the remaining 1% of wealthiest students. If the results are plotted with a circle proportional to the percentage of the population considered, the sequence of solid blue circles is obtained. If a similar procedure is followed, but now starting with the poorest 1%, and gradually including the wealthiest 1%, iterating until the wealthiest students are included, the sequence of white circles with a blue outline is obtained. As can be confirmed, the largest spheres in both series coincide with the population means, since the estimates in both cases include the entire population.

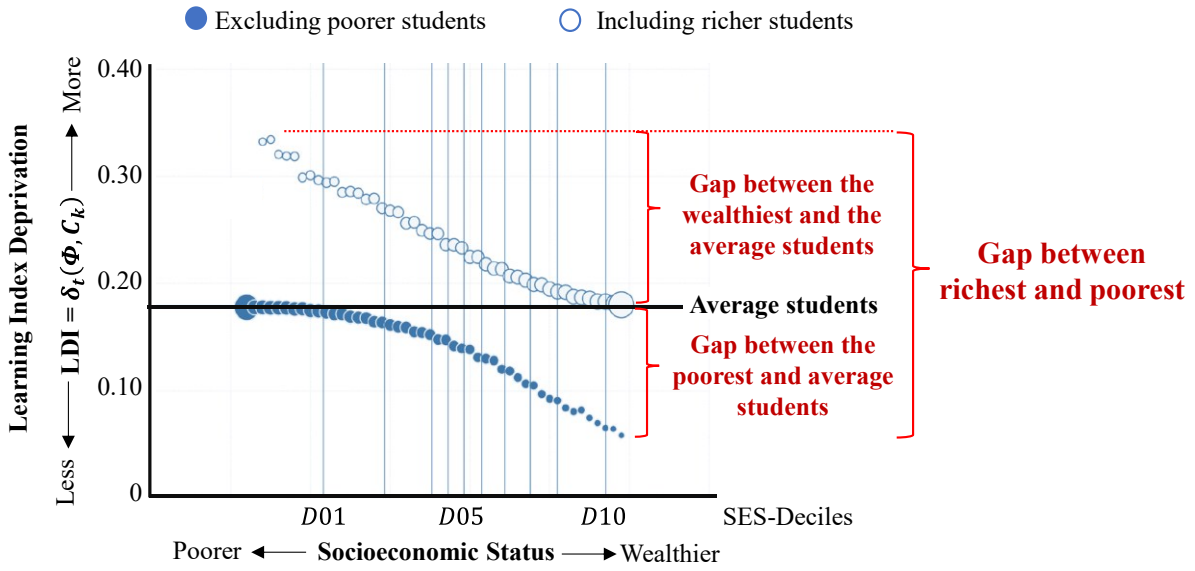


Fig. 4. Distribution of levels of achievement among SES deciles (left side). Relative flows of the distribution of levels of achievement among SES deciles.

Since the publication of the seminal *Coleman Report* (1966), empirical evidence has consistently demonstrated that family background frequently outweighs school resources in explaining variance in student outcomes [Reardon, 2011]. Results of our own analysis of multiple datasets align with a vast body of literature identifying SES as the strongest predictor of academic achievement [Chzhen & Bruckauf, 2016; Hanushek & Woessmann, 2020; Skopek & Passaretta, 2021]. At all scales, students, school, country, and region levels, we observe a pronounced and consistent gradient of educational deprivation across socioeconomic deciles. The previous result is surprising because its monotonicity shows clearly that learning deprivation is a function of SES, showing that the total gap between rich and poor students can be decomposed into the sum of gaps, providing evidence of how the LDI index—estimated by $\delta_t(\Phi, C_k)$—can detect small changes in the composition of the population and its an optimal candidate for variational analysis.

1.4. Network Analysis of Academic Inequality

From a complexity perspective, we might say that the structure of interactions determines the system's behavior. Whether analyzing the regional interdependence of learning outcomes or the structural mechanisms of resilience in high-poverty contexts, the challenge lies in mapping the potential relationships, unveiling hidden coherent interlinked latent variables, and understanding their underlying mechanisms within a plausible theoretical framework. This thesis adopts this relational stance, arguing that for dismantling the structural inequalities in

education, one must first render visible the complex network of dependencies that restricts the abilities of the most vulnerable. Unlike traditional econometric methods, which used to be limited to specific situations with many operational restrictions, modern network science provides both the conceptual and computational tools to analyze large-scale systems for uncovering "hidden structure of everything" [Barabási, 2016]. Network science also offers the methodological lexicon required to operationalize relational perspectives, but most LSA's — such as PISA or regional UNESCO studies—are still analyzed, almost exclusively, through Hierarchical Linear Models (HLM), these methods excel at ranking students but fail to map the structural interdependencies that produce those rankings.

Once the optimal functioning of the new indicators was proven and their measurement was confirmed using data provided by official LSAs, it became natural to use them to partition and model the networks associated with learning deprivation. A three-step sequential model was developed to analyze topological properties and identify the nodes holding the system out of statistical equilibrium. The first step is focused on disaggregating L_1-class by *SES*, each student is represented by a node and edges, weighted by $\lambda(LI_j)$, are directed to one of the SES-decile nodes $\{LI_j(\theta^j \rightarrow L_k^j \rightarrow (SES_d^j)\}\ \forall\, j$, a process which allows to analyze aggregated inequality at school level, as well as *In-degree distribution* for SES nodes. The second one is an extension for including rurality of the area (RA), type of school (TS), ethnic self-identification (ET), and gender (GE) to analyze their effects through the sequence $\{\theta^j \rightarrow L_k^j \rightarrow (SES_d^j) \rightarrow (RA_{C1}^j, TS_{C2}^j, ET_{C3}^j)\}\ \forall\, j$, where *C* denotes an index for each subcategory of the factors RA, TS and ET, and GE. Finally, the third step amplifies and strengthens the analysis through more than one hundred educational and non-educational factors associated with learning achievements trough the sequence $\{\theta^j \rightarrow L_k^j \rightarrow (SES_d^j) \rightarrow (RA_{C1}^j, TS_{C2}^j, ET_{C3}^j) \rightarrow FAL_{Cm}^j\}\ \forall\, j$, for integrating *m* different educational and non-educational factors[1], the result is shown in Figure 4, and in can be seen with more detail in chapter 4.

[1] Network analysis was carried out by Gephi 0.9.2 and statistical estimations and plots with R 3.5.0.

Learning deprivation: $\delta_t(\Phi, C_k) = H(\Phi, C_k) \cdot \mathbb{E}[\Lambda(\theta_i; \Phi, C_k)]$

Social determinants: $(SES_k^i) \rightarrow (RA_k^i, TS_k^i, ET_k^i, GE_k^i) \rightarrow \{AF_{m,n}^i\} \leftarrow$ **Associated factors**

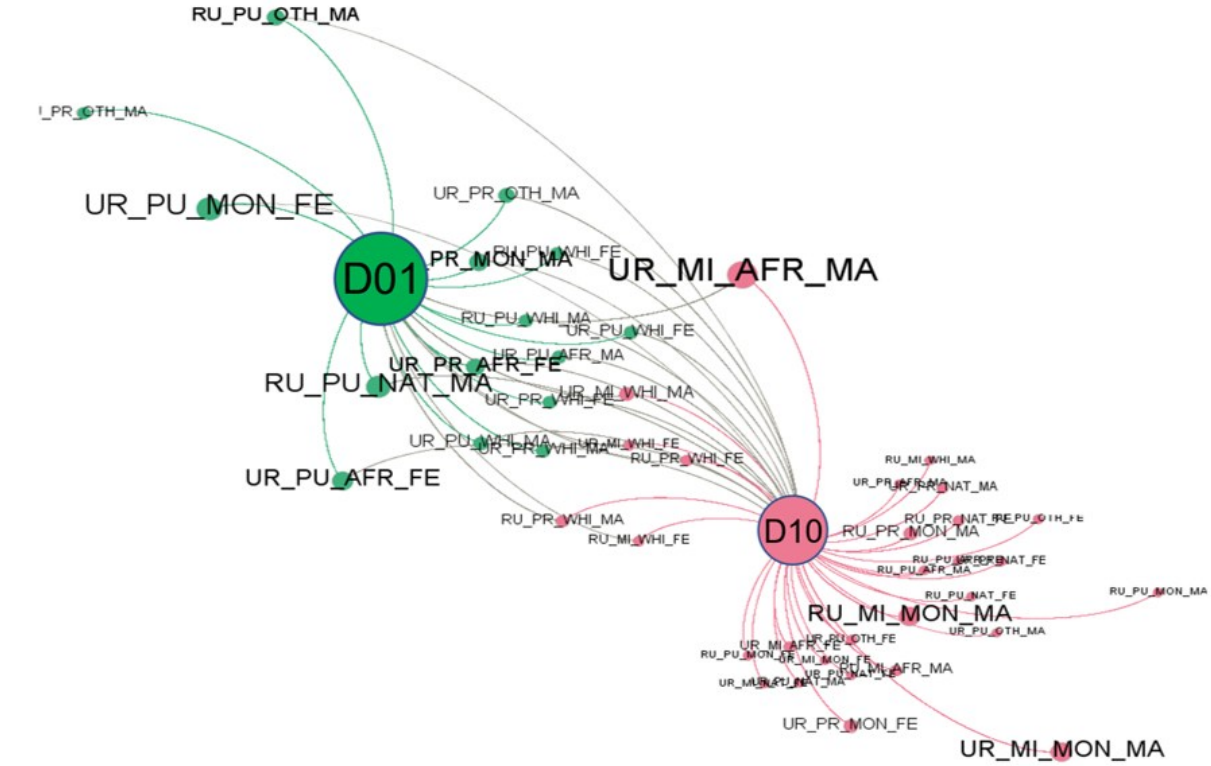


Fig. 5. Specification for the general network model using indicators for measuring learning deprivation and social determinants, ordered hierarchically for analysis by (1) Socioeconomic Status (SES), Rurality (RA), Type of school (TS), Ethnic self-identification (ET), and Gender (GE) [Sánchez-Restrepo, H., & Jorge Louçã, 2018; 2019].

These findings highlight how socioeconomic status (SES), often exacerbated by multiple associated disparities that persist throughout life, creates fertile ground for the vulnerability of those with less access to resources, following the line of thought that Charles Tilly (1998) deciphers in his work, Enduring Inequality: Exploitation and the Origins of Inequality, which proposes and explains that social inequalities can persist because "disparity is not an accidental consequence of individual differences, but a consequence of categorical pairs." In this regard, it is possible to show that systems with widespread inequality are sustained in that negative state because they remain institutionalized through power relations and privileges, claiming relations are systematically established to dominate the monopoly of opportunities, creating cycles of poverty based on exclusion mechanisms that operate in a coordinated manner to diminish the opportunities to the more disadvantaged or less connected persons.

This finding is undoubtedly profoundly relevant, since in classical models, SES is treated simply as a control variable or a static attribute, a decision that can amplify the risk of obscuring

or hindering our understanding of the mechanisms through which advantage is transmitted. Given that the sources of the relational nature of inequality imply that the achievement gap could, in reality, be a structural chasm provoked not only by barriers to access and retention, but also by barriers to the utilization and enjoyment of resources. In this regard, Bourdieu (1986) points out that “academic success is not simply a function of innate talent but depends on the possession of cultural codes and social connections validated by the school system”. Essentially, these other forms of capital are also inherently relational because they exist only in reference to a field of interaction. Thus, a student's cultural capital is not a static asset carried in isolation, but a "key" that works effectively only if it fits the “lock” of the institutional structure. Therefore, if the educational network is fragmented or segregated, the transmission of this capital is likely to be interrupted, creating what we could call topological barriers to social mobility.

1.5. From Nodes and Edges to Metrics and Topology

Despite the clear theoretical compatibility between network science and educational sociology, the empirical intersection of these fields remains underdeveloped, particularly at the meso and macro levels, due mainly to the fact that existing literature on networks in education has predominantly focused on the micro-level of organization. For instance, van der Maas and Siew (2020) had been developing network models for cognitive development and intelligence, as well as quantifying knowledge and the development of expertise, but the classical analysis continues being the mainstream—for an excellent metareview of research on educational inequality and socioeconomic disadvantage, we strongly suggest Sykes, C., et al (2025).

In most of the literature, network metrics might be classified into two categories based on their scale: (1) *local metrics*, which pertain to individual nodes, and (2) *global metrics*, whose purpose is to characterize the structural properties of the whole network, they can be grouped in three primary categories: Centrality measures, Clustering coefficients, and Assortativity indices, providing the grammar for translating mathematical formalism into sociological and concepts. In the context of modeling educational inequality and structural deprivation, these metrics cease to be mere abstract calculations; they become operational indicators of social capital, segregation, and resource flow within a complex network.

This division between scales examines the agency and position of individual nodes and the global scale. This strategy is focused on capturing the emergent properties of the entire structure. Within this multiscale framework, this taxonomy provides the necessary grammar to

interpret two distinct types of structures: (1) the sociological topology of educational systems, and (2) the psychometric architecture of learning outcomes, which later allowed the identification and subsequent characterization of academic resilience.

1.5.1. Local Metrics

At the local level, centrality measures allow us to identify the specific roles played by individual actors—whether they are schools in a social network or psychological factors in a resilience profile. In this sense, the Weighted degree serves as a primary measure of volume and influence. In a given social context, this metric captures the intensity of a student's interactions with a single factor, reflecting the intense use of a specific resource by a relevant number of students. Conversely, when applied to a psychometric network, it identifies the "system drivers"—those high-impact factors that hold the greatest predictive power over a student's trajectory. To understand the efficiency of these factors, *Closeness centrality* and *Harmonic closeness centrality* indicate how widely a node can access the network; in addition to their strong correlation, we prioritize Harmonic closeness for its ability to handle fragmentation. This makes it particularly valuable for analyzing educational systems, as it allows for the accurate quantification of accessibility, even for schools in isolated rural areas or segregated enclaves.

Moving beyond access, *Betweenness centrality* allows us to map the system's intermediaries. For instance, in our model, we develop an *Academic Correlation-based Factor Network*, where the adjacency matrix is based on a multivariate correlation matrix among demographic data and more than 140 factors associated representing Self-regulated learning, Learning practices and perceptions, Socio-emotional skills, Physical and mental self-care, Household structure, Educational support, School climate, Living conditions, Educational trajectory and Commitment with learning, reflecting accurately the underlying structure of the latent variables driving students' behavior, habits, emotions, and learning outcomes. In this scenario, high-betweenness nodes act as bridges—and sometimes as bottlenecks—controlling the flow of information in disparate social groups. When studying resilience, these nodes might represent mechanisms of transfer, linking contextual support—such as family—to behavioral outcomes oriented to success.

For analyzing hierarchy, *Eigenvector* and *Katz Centrality* are the main measures, which weigh a node's importance by the influence of its neighbors. These metrics help us formalize the concept of elite reproduction in social networks and identify points of "strategic leverage" in psychological profiles. Finally, the *Authority score*—from the HITS algorithm [Deguchi, T., et al, 2014]—can be used to distinguish those nodes that act as ultimate destinations for high-

quality resources, validating which factors serve as the system's primary reference points. Complementing these centrality measures, the *Clustering Coefficient* quantifies the degree of local cohesion. Sociologically, a high clustering coefficient suggests a dense, supportive peer group that generates bonding social capital—though we must be cautious, as excessive clustering can signal insularity or high availability of the factor at low cost. Psychometrically, this measure indicates stability; a group of students with a highly clustered set of factors possesses a redundant, mutually reinforcing system that is better equipped to buffer against external shocks.

1.5.2. Global Metrics

In this research, the emergent properties used to be related with the global scale because define the system's architecture, At this level, centrality distributions refers to hierarchy among nodes, while the assortativity indices and community structure metrics reveal the depth of inequality in learning outcomes, leaving to the Modularity class serving as the primary indicator of segregation. This measure optimizes the partition of the network into dense communities—distinct, non-overlapping groups that reveal the rigid topological boundaries of socioeconomic or ethnic segregation. When applied to resilience, global modularity maps functional maps of behavior, allowing the identification of latent domains of functioning guided by data and theoretical frameworks. Therefore, *Centrality* is crucial for identifying the most influential factors or agents within the complex network of a learning environment. By measuring how central certain traits, behaviors, or social connections are related, it is possible to point out which elements play a pivotal role in supporting high achievement facing adverse socioeconomic contexts—referred frequently as *academic resilience*. In this sense, nodes with high centrality often act as key facilitators of information flow, emotional support, or resource access, enabling resilient students to use internal and external resources in a more effective way.

By modelling an educational system as a network—where nodes represent agents and edges represent flows of a wide variety of academic and non-academic resources—we can move beyond descriptive statistics and focus on understanding processes and try to detect or infer relationships. Along this thesis, we identify and address this specific research gap in each opportunity for proposing that the tools of complex network analysis, typically reserved for biological or technological systems, can be repurposed to map the "topology of learning poverty", and other observable expressions of social inequality. Given that learning poverty started to be the global referent just a few years ago, there is a paucity of research that treats the entire educational system as a whole—even without including the interplay of socioeconomic factors, learning outcomes, and resource distribution, extending this view to the structural inequality of learning deprivation.

For deploying a comparative analysis of centrality metrics, it is mandatory to prepare the relational model and data—in this case, the challenge is located in the adjacency matrix—, after analysis, integrating the outcomes in tables and graphical models are very useful ways for starting to analyze and communicate the results in a clearly manner, as in Table 1 or Figure 6, where we present network parameters and representations for *Resilient* (Q1) and *Privileged* (Q5) students, analyzed on detail in chapters 7 and 8. In this case, the metrics of centrality Closeness, Harmonic closeness, and Betweenness are related with a fundamental divergence in students with high academic performance. In both networks, *Emotional Well-Being* acts as the single most critical bottleneck ($\mathrm{Q1}: 0.32$; $Q5: 0.33$), showing a role as the non-negotiable structural prerequisite for the flow of all other resources, regardless of socioeconomic status [Newman, M. E. J., 2018].

Given the values of the parameters—and its relevance for the network—*Well-being* appears like the most relevant factor for controlling the flow of all other academic and social resources into the high-performance system. However, in the resilient network are factors associated with internal effort, such as *Metacognition* ($Q1: 0.007$ vs. $Q5: 0.00$), and *Study habits* ($Q1: 0.06$ vs. $Q5: 0.00$), are also necessary bridges within the $Q1$ system, yet remain peripheral in the $Q5$ system.

Table 1. Network parameters for *Resilient* (Q1) and *Privileged* (Q5) students.

Factor	Weighted Degree		Closness centrality		Harmonic closness centrality		Betweeness centrality		Authority		Modularity class		Clustering		Eigen centrality		Katz Centrality	
	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1
Emotional Well-Being	4.74	4.65	0.70	0.72	0.78	0.82	0.32	0.33	0.24	0.31	0.24	0.31	0.03	0.04	1	3	0.40	0.41
Cooperation	5.87	4.92	0.70	0.70	0.82	0.81	0.21	0.23	0.29	0.32	0.29	0.32	0.04	0.04	2	2	0.43	0.45
Sense of Belonging	2.29	2.20	0.48	0.46	0.54	0.53	0.10	0.10	0.03	0.03	0.03	0.03	0.02	0.02	1	3	0.40	0.50
Bullying	2.85	2.15	0.52	0.51	0.59	0.57	0.09	0.06	0.07	0.07	0.07	0.07	0.02	0.02	2	4	0.50	0.60
Achievement Level	1.30	0.58	0.45	0.00	0.50	0.00	0.07	0.00	0.02	0.00	0.02	0.00	0.02	0.01	1	0	0.33	0.00
Study Habits	2.55	1.18	0.49	0.38	0.58	0.43	0.06	0.00	0.07	0.02	0.07	0.02	0.02	0.01	1	1	0.29	0.67
Self-Efficacy	5.04	3.66	0.61	0.59	0.74	0.71	0.04	0.01	0.28	0.28	0.28	0.28	0.03	0.03	0	2	0.63	0.77
Motivation	4.54	3.41	0.61	0.58	0.74	0.69	0.04	0.07	0.28	0.25	0.28	0.25	0.03	0.03	0	2	0.63	0.70
Metacognition	4.82	3.78	0.60	0.59	0.72	0.71	0.04	0.01	0.26	0.28	0.26	0.28	0.03	0.03	0	2	0.63	0.69
Family Support	2.22	1.80	0.45	0.46	0.51	0.51	0.03	0.03	0.02	0.03	0.02	0.03	0.01	0.01	1	3	0.60	0.67
Self-Control	3.25	2.35	0.60	0.57	0.70	0.67	0.02	0.07	0.24	0.22	0.24	0.22	0.03	0.02	0	2	0.85	0.71
Emotional Regulation	4.00	2.72	0.59	0.52	0.69	0.60	0.02	0.00	0.24	0.18	0.24	0.18	0.03	0.02	0	2	0.80	0.82
School Violence	2.29	1.70	0.50	0.36	0.56	0.42	0.02	0.00	0.05	0.01	0.05	0.01	0.02	0.01	2	4	0.67	1.00
School Climate	1.28	1.02	0.45	0.46	0.53	0.53	0.01	0.01	0.05	0.04	0.05	0.04	0.01	0.01	2	4	0.60	0.83
Vitality	2.21	1.83	0.49	0.44	0.54	0.48	0.01	0.00	0.05	0.05	0.05	0.05	0.01	0.01	1	3	0.40	0.33
Cognitive Strategies	2.50	2.06	0.54	0.54	0.63	0.64	0.01	0.00	0.19	0.22	0.19	0.22	0.02	0.02	0	2	0.84	0.86
Satisfaction with Interactions	1.07	0.80	0.43	0.32	0.48	0.36	0.01	0.00	0.02	0.00	0.02	0.00	0.01	0.01	2	4	0.33	1.00
School Satisfaction	2.15	1.98	0.37	0.46	0.44	0.54	0.00	0.08	0.01	0.04	0.01	0.04	0.01	0.02	2	4	0.70	0.60
Attentional Control	4.76	3.99	0.59	0.59	0.71	0.71	0.00	0.01	0.27	0.28	0.27	0.28	0.03	0.03	0	2	0.87	0.76
Civic Engagement	4.61	3.84	0.59	0.59	0.71	0.72	0.00	0.01	0.27	0.29	0.27	0.29	0.03	0.03	0	2	0.87	0.63
Empathy	4.44	3.52	0.59	0.58	0.71	0.69	0.00	0.01	0.27	0.26	0.27	0.26	0.03	0.03	0	2	0.87	0.77
Assertive Communication	4.23	2.18	0.59	0.50	0.71	0.58	0.00	0.00	0.27	0.16	0.27	0.16	0.03	0.02	0	2	0.87	0.81
Leadership	4.17	4.56	0.58	0.60	0.69	0.74	0.00	0.04	0.25	0.30	0.25	0.30	0.03	0.03	0	2	0.90	0.54
Gender	0.68	0.45	0.48	0.38	0.52	0.44	0.00	0.00	0.05	0.02	0.05	0.02	0.01	0.01	1	1	0.33	0.00
Growth Mindset	3.99	2.79	0.52	0.55	0.66	0.66	0.00	0.00	0.25	0.24	0.25	0.24	0.03	0.02	0	2	0.89	0.80
Emotional Awareness	3.39	2.26	0.50	0.47	0.63	0.58	0.00	0.00	0.22	0.18	0.22	0.18	0.02	0.02	0	2	0.92	0.79
Household Composition	0.22		0.34		0.38		0.00		0.01		0.01		0.01		2		0.00	
Teacher Support	1.66	1.64	0.33	0.33	0.38	0.37	0.00	0.00	0.00	0.01	0.00	0.01	0.01	0.01	1	3	1.00	1.00
Environmental Safety	0.24	0.40	0.33	0.32	0.36	0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	1	3	0.00	0.00
Attendance and Punctuality	0.22		0.32		0.34		0.00		0.00		0.00		0.01		1		0.00	
Expectations	0.71		0.40		0.46		0.00		0.06		0.06		0.01		0		1.00	
Emotional Well-Being	4.74	4.65	0.70	0.72	0.78	0.82	0.32	0.33	0.24	0.31	0.24	0.31	0.03	0.04	1	3	0.40	0.41
Max	5.30	6.24	0.72	0.70	0.82	0.82	0.19	0.24	0.32	0.29	12.00	7.00	1.00	1.00	1.00	1.00	3.90	6.89
Min	1.76	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.00	2.00	0.00	0.00	0.00	0.00	1.25	1.25
$\widehat{\boldsymbol{\mu}}(\cdot)$	3.15	3.71	0.41	0.46	0.48	0.54	0.02	0.02	0.13	0.13	6.50	3.44	0.57	0.63	0.40	0.46	2.29	3.79
$\widehat{\boldsymbol{\sigma}}(\cdot)$	1.11	1.36	0.22	0.18	0.26	0.22	0.04	0.05	0.12	0.12	2.38	1.72	0.32	0.29	0.38	0.40	0.95	2.19

Note. All edges are undirected. Coefficients Closeness centrality, Harmonic closeness centrality and Betweenness centrality are normalize centralities $[0,1]$. All factors are supposed to be latent variables. Coefficients cannot be converted into an odds ratio.

Now, by modelling variables and agents as nodes in a network, we can unveil evidence of how educational exclusion is not just a lack of resources, but a specific structural trap created by the absence of bridging ties and the presence of reinforcing loops of disadvantage, a key milestone in education that can tackle the challenge of integrating topological, pedagogical, and sociological perspectives. In this regard, the analysis of official LSA's dataset for modelling complex networks can be accomplished by estimating the multiple rank-based correlation matrix among associated factors for fulfilling the adjacency matrix—a method developed in chapters 7 and 8 of this thesis—. This method is especially useful when it is a high-dimensional classification problem and using combinatorial produce noise accumulation, which also leads to a rise in the rate of spurious correlations among unrelated covariates [Masuda, N., et al, 2018; Schober, P., Boer C, Schwarte LA., 2018; Borsboom, D., et al, 2021; Masuda, N., et al, 2024].

Another case in where this method outperforms is whit Community detection algorithms, such as modularity maximization, when used to partition networks into groups of nodes—in communities or modules— that are more interconnected among themselves than with nodes outside those groups.

The robustness of network structures provided by correlational matrix allows finding their functional associations weighting the edges, enabling a new way for estimating the structure of interactions among latent variables with social, emotional, and cognitive skills, with the advantage of mapping and translating multivariate student resource profiles into interpretable topological architectures [Borsboom et al., 2021; Brinkhof et al, 2025], as well as to escalate modelling the entire multivariate resource ecosystem. As can be seen, Figure 6 and Table 1 correspond to the same two high-performance student networks—coming from two different socioeconomic backgrounds (D01 and D10), where color of the nodes represents the *Modularity class*, their size represents *Katz centrality* coefficient, and their label size *Betweenness centrality.* In this study, modularity is employed to identify densely connected communities of factors, providing evidence on whether the resilient group exhibits a highly integrated resource ecology—low modularity, typically indicating that resources operate synergistically—or a fragmented configuration [Masuda et al., 2024].

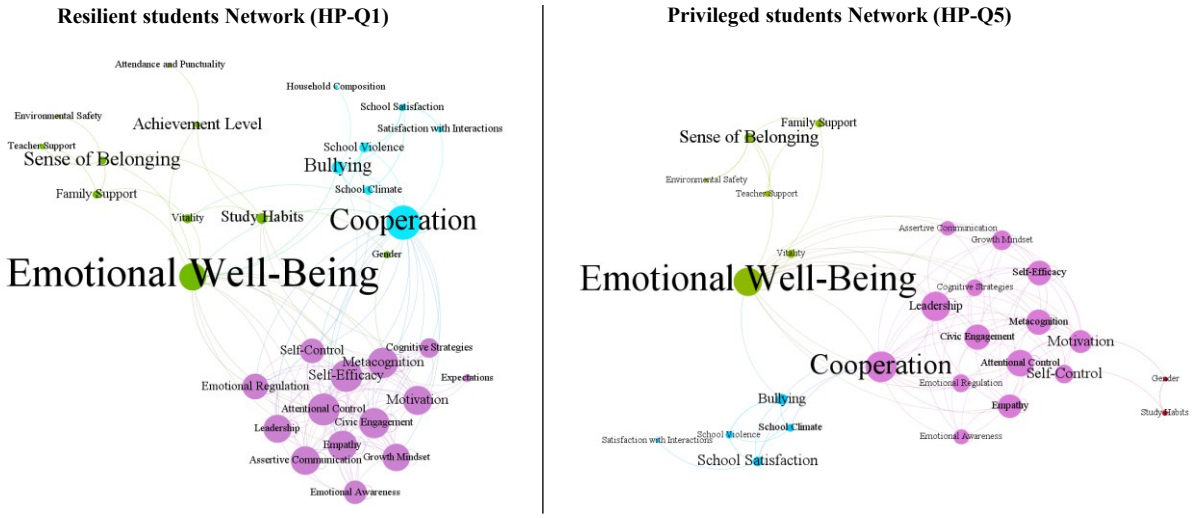


Fig. 6. Factor's Correlation-based Academic Networks. The figure illustrates the structural and topological differences in resource organization between resilient (HP-Q1, left) and privileged (HP-Q5, right) high-performing students, quantified by node centralities and modular partitioning.

As can be seen, this approach puts together the metrics for advancing from mere correlational observations to structural and causal inferences regarding the effectiveness of different policies. According to Tilly (1996), if inequality is persistent because it is embedded within educational structures, then we need new tools for quantifying multiple configurations for a small number of emergent properties. From this point of view, if it seems feasible to monitor educational inequality, why to remain attached to an "attribute-based" approach for just merely correlating poverty with low achievement, when what is needed is a framework for mapping the structural architecture of segregation? This calls for a shift from measuring the sheer volume of resources available to a student toward analyzing the efficiency of the network in which they are situated, raising a fundamental question: how do the topological properties of the educational system—its modularity, connectivity, and centralization—function as the invisible scaffolding that sustains or breaks these inequalities?

With the aim of mitigating the dissonance between the global mandate of SDG 4 and the conceptual and methodological limitations currently employed to monitor it, the proposal presented in this thesis suggests moving beyond metrics that describe the symptoms of inequality to diagnose its structural pathology. By recognizing education as a system whose complexity stems from its multiple interactions at different scales, it becomes clear that educational gaps are emergent properties and that their persistence is due to structural mechanisms that cannot be addressed simply by allocating resources or by focusing individually on those gaps that manifest in the most visible urban areas, instead of looking for solutions that require identifying and addressing the structural determinants that drive those dynamics. In this sense, most of the literature suggests that to truly address academic inequality, we must operationalize the educational landscape as an interdependent network of relationships. In this chapter, we argue that network architecture—defined by its connectivity, segregation, and modularity—constitutes the invisible scaffolding of opportunities to learn in significant ways. Therefore, subsequent chapters, based on published material, have attempted to move beyond theory and delve into empirical data. Along this path, we have dedicated considerable effort to developing methods for analyzing contrasts and relationships across topologies and other properties using Network Analysis.

To conclude this chapter, it is valuable to recall the words of R. Dum & J. Johnson (2016): "Science contributes by helping to generate and evaluate policies and by providing the best possible evidence-based answers to the question of how policymakers can ensure that their policies will have the desired outcomes and not undesirable ones." This underscores the importance of always placing evidence at the heart of the quality of our conclusions, and not

just human desires. Based on the results of the last years, we are convinced that unveil and understand the silent determinants of systemic fragility and resilience, is not a merely technical refinement, but a necessary epistemological shift, remembering that, as long as educational policies focus on isolated attributes, the mechanisms that underpin deprivation will remain invisible to decision-makers, and effective evidence-based policy will remain an utopia.

CHAPTER 2

Topological Properties of Inequality and Deprivation in an Educational System: Unveiling the Key-Drivers Through Complex Network Analysis

First Autor: Harvey Spencer Sánchez Restrepo
Harvey.Restrepo@iscte-iul.pt
Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL
1649-026, Lisbon, Portugal

Coauthor: Jorge Louçã
jorge.l@iscte-iul.pt
Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL
1649-026, Lisbon, Portugal

Running head: **Topological Properties of Inequality in an Educational System**

Article published in: *Advances in Intelligent Systems and Computing*, vol 1026.

Springer, Nature.

Topological Properties of Inequality and Deprivation in an Educational System: Unveiling the Key-Drivers Through Complex Network Analysis

Harvey Sánchez-Restrepo(✉) and Jorge Louçã(✉)

Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL, 1649-026 Lisbon, Portugal
{Harvey_Restrepo, jorge.l}@iscte-iul.pt

Abstract. This research conceives an educational system as a complex network to incorporate a rich framework for analyzing topological and statistical properties of inequality and learning deprivation at different levels, as well as to simulate the structure, stability and fragility of the educational system. The model provides a natural way to represent educational phenomena, allowing to test public policies by computation before being implemented, bringing the opportunity of calibrating control parameters for assessing order parameters over time in multiple territorial scales. This approach provides a set of unique advantages over classical analysis tools because it allows the use of large-scale assessments and other evidences for combining the richness of qualitative analysis with quantitative inferences for measuring inequality gaps. An additional advantage, as shown in our results using real data from a Latin American country, is to provide a solution to concerns about the limitations of case studies or isolated statistical approaches.

Keywords: Complex network · Educational deprivation · Inequality · Large-scale assessments · Policy informatics

1 Introduction

Since the Sustainable Development Goals were launched [1], policymakers have been increasing the use of analytical and statistical models for improving their knowledge about complex dynamics of the educational systems, especially for those with high ethnic-linguistic diversity [2], facing the challenge of developing coherent multilevel theories for making decisions and designing public policies [3]. The wide number of nonlinear relationships exhibited by multiple agents in different hierarchical levels interacting in education, demands the development of new instrumental and thought tools for modelling the variety of fluxes of energy - financial, human, physical and social resources—as well as a better understanding and measurement of the parameters related with the emergence of self-organized patterns and groups in different scales [4].

Network theory has enormously developed in this decade and is a leading scientific field for descripting and analyzing complex social phenomena [5]. Educational systems

© Springer Nature Switzerland AG 2020
T. Ahram et al. (Eds.): IHSED 2019, AISC 1026, pp. 1–7, 2020.
https://doi.org/10.1007/978-3-030-27928-8_72

exhibit different properties in many scales coming from same dynamics, students stablish relationships with other students and teachers, which might provide some insights about social characteristics about preferences, choices and interest on learning [6]. Actually, many interventions show that interactions impact instructional quality and learning outcomes [7] and there are some evidences on how social network structure becomes an important intermediate variable in education and that cultural, social and economic variables are related with educational deprivation and learning outcomes [8].

2 Modelling Framework

In this research we incorporate a rich framework for analyzing topological and statistical properties of educational deprivation in a Latin American country, as well as its relationship with social determinants as Socioeconomic Status (SES), Rurality of area where the school is located (RA), Type of school (TS), and self-identify student's Ethnicity (ET), for unveiling the key factors driving inequality gaps in learning outcomes. The model is developed through a network with different levels for analyzing properties and nodes exhibiting centrality and non-equilibrium parameters than might help to better understand the structure of the system and phenomena behind them.

Data Sources. A multivariate dataset related with learning outcomes of every student who has completed the *k-12* education process, estimated by scoring based on a census-based large-scale assessment carried out in Ecuador for 39 219 students in 2017, through a standardized computer-based test with psychometric parameters estimated by Item Response Theory with 2P-Logistic model. Raw scores were re-scaled to a *Learning index* (LI), a monotonous transformation of ability's parameter θ^i, where higher levels of learning are more likely to have higher scores.

The model was developed in four phases, the first one was psychometrical analysis for estimating scores and identify deprived students (L_0)—those with a LI below of the cut point L_0—and the intensity of deprivation $\lambda(LI_i)$, given by the distance from LI_i to L_0. The next three phases are directly based on the level-network (LN). In this way, the 1-LN is for disaggregating L_0-group by SES, each student is represented by a node and edges are directed to one of the SES-decile nodes $\left\{\theta^i \rightarrow L_j^i \rightarrow \left(SES_k^i\right)\right\}\forall i$, weighted by λ. In 2-LN we extend 1-LN for including RA, TS and ET to analyze their effects through the sequence $\left\{\theta^i \rightarrow L_j^i \rightarrow \left(SES_k^i\right) \rightarrow \left(RA_k^i, TS_k^i, ET_k^i\right)\right\}\forall i$. The 3-LN amplifies and strengthens the network through more than one hundred educational and non-educational factors associated with learning achievements [9], trough the sequence $\left\{\theta^i \rightarrow L_j^i \rightarrow \left(SES_k^i\right) \rightarrow \left(RA_k^i, TS_k^i, ET_k^i\right) \rightarrow AF_{m,n}^i\right\}\forall i$, a multi-dimensional system exhibiting educational deprivation at different levels. Network analysis was carried out by Gephi 0.9.2 and statistical estimations with Orange 3.3.8.

3 Empirical Results

Estimates indicate that, from 39 219 students, 8 438 were deprived, an absolute prevalence of 21.5%. The L_0-group has a $LI = 6.32$ and intensity of deprivation $\lambda = 0.22$, i.e., in average, each L_0-student lacks 0.68 standard deviations (SD) to the minimum level of learning for not being deprived.

3.1 Socioeconomic Status and Learning Deprivation

In all cases, inequality means asymmetries, in conditions of total equity—where socioeconomic factors would not produce differences—we might expect equal distribution of L_0-edges over the network, but 1-LN specification detects SES effects in nodes grouping L_0-students by deciles, driving the system out of equilibrium with a negative correlation between LI and SES ($R = –0.58$ ($p < 0.001$)). A 2-LN model shown in Fig. 1 integrates the different self-identified ethnic groups and disaggregate them by Rural-Urban areas and Public-Private schools, to identify the magnitude with which the lower deciles dominate the interactions through the edges.

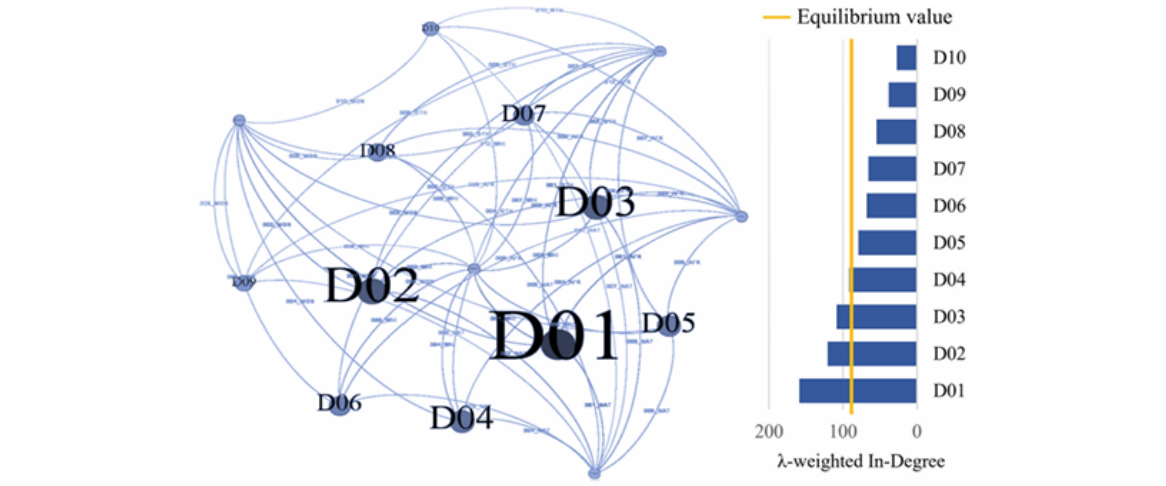


Fig. 1. Socioeconomic structure of learning deprivation (*left side*) and λ-weighted In-degree histogram (*right side*). Differential centrality of deciles shows a non-equilibrium L_0-system driven by ethnical groups dominated by poorest students.

As can be seen in Fig. 1, independently of RA, TS and ET, nodes corresponding to D01, D02 and D03—the poorest students—have stronger connections and dominate the network, being D01 the highest parameters in Prevalence rate ($DPR = 0.3860$), Hub ($H = 0.4887$), Weighted Degree ($WD = 158.9420$) and PageRank ($PR = 0.4289$). On the contrary, D10 is almost irrelevant for the network with parameters $DPR = 0.0800$, $H = 0.1133$, $WD = 27.222$ and $PR = 0.0289$.

3.2 Ethnicity and Type of School Financing

There is a negative correlation between SES and DPR ($R = –0.81$, $p < 0.001$) and, as can be seen in scatter-plot of Fig. 2, quintiles 1 and 2 have the highest lack of learning. The complementary Dendrogram was made by a hierarchical cluster analysis for $\left(SES_k^i, RA_k^i, TS_k^i, ET_k^i\right)$ and, in all cases, families are basically made of SES and TS where emerges an intra-class ranking ordered by ethnicity, meaning that closeness is a SES' function while distances are based on racial proximities, a structural inequality for the whole system. In this sense, scatter-plot shows that White-students (red circle) have the lowest deprivation rate among the poorest ($DPR = 0.402$), even lower than richest Afro-Ecuadorian students in blue circle ($DPR = 0.408$). For the whole network, Page Rank order of nodes and λ-intensity is based on students' ethnicity as follows: Afro-Ecuadorian (A), Montuvios (M), Indigenous (I), Other groups (O) and White (W). Furthermore, Private sector is dominated by richest students ($SES = 0.91$), while public schools serve to the poorest students ($SES = 0.33$) getting a $DPR = 0.257$, 2.12 times the rate of the private ones ($DPR = 0.121$), showing that SES is a key factor for educational deprivation due to the influence of cultural capital in learning outcomes.

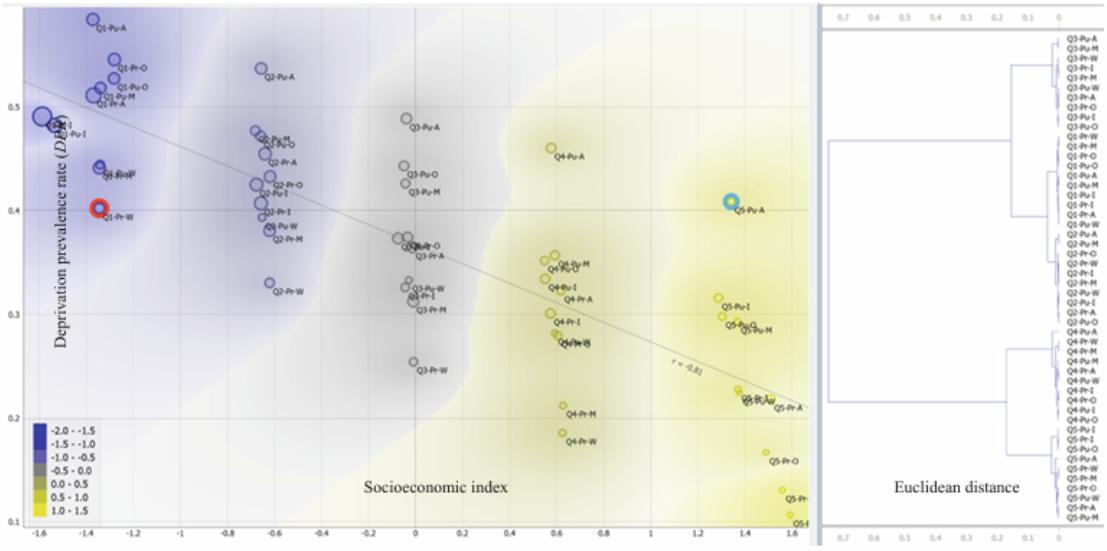


Fig. 2. Social determinants of learning deprivation. The scatter plot shows the strong negative correlation between SES and Deprivation prevalence rate (*solid line*). Highest deprivation rate in richest students (*blue circle)* is lower than the lowest rate of poorest students (*red circle*), pointing out a deep systemic racial discrimination.

3.3 Key-Factors for Public Policies

As 1-Ln and 2-LN are networks strongly connected, prevalence rate of L_0-students might be associated with Eigen-Centrality through measuring factor's influence for identifying how well connected a node is and how many links have its connections. In this way, building 3-NL for splitting L_0-students in communities becomes in a very valuable tool for developing group-oriented strategies to avoid implementing same actions for completely different populations and needs. One of the most useful

strategies for improving learning and closing gaps is micro-planning, i.e., implementing different policies at local level, however, to select the most relevant needs is a great deal, mainly because they use to interact and 'ceteris paribus hypothesis' seems to be too naive; through network simulation, this can be solved in a very easy way varying *AF*-parameters for recognizing and ranking the most relevant nodes (factors) to be attended by policymakers.

Figure 3 shows the 3-LN after a Modularity process (with parameter 0.073 at resolution of 0.254) for splitting richest from poorest students to find key factors for educational deprivation in both groups. As shown, richest students' community (D10) has just a few factors (20) and most of them have very low connectivity. On the contrary, poorest students (D01) have a lot of different sources provoking deprivation (57), reflected in the Degree of authority of the D01 (0.98743) versus D10 (0.158058).

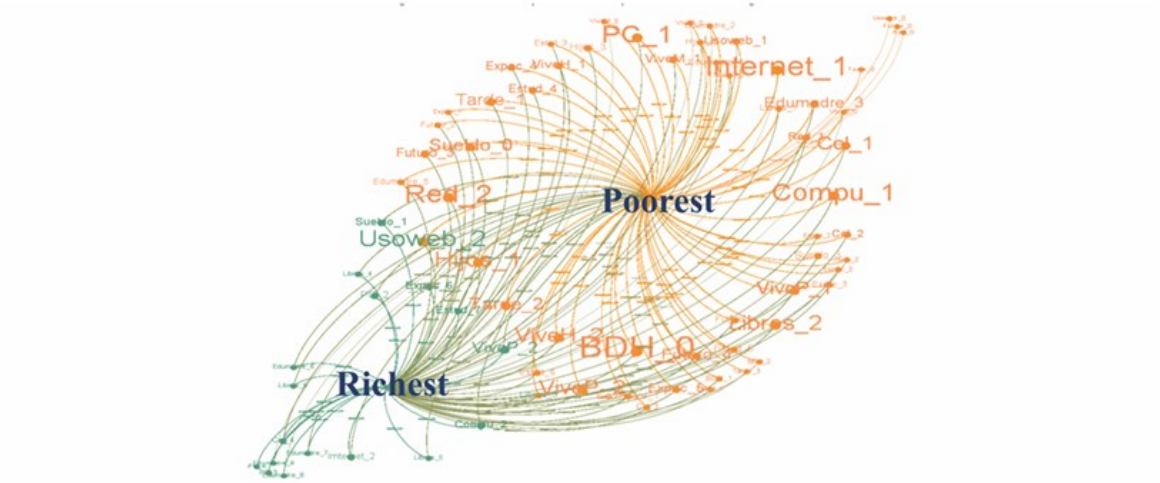


Fig. 3. Factor's network for richest and poorest students' communities. Centrality of nodes in each group points out the dissimilarities in factors provoking deprivation.

Thus, PageRank in 3-NL network helps to find and rank those factors dominating deprivation with the highest Degree of connectivity and centrality for each community ordering from the highest degree until the lowest for providing a very accurate and clear knowledge. In addition, it might be defined 'needs profiles' for focusing actions on those variables susceptible to be managed by policy, for example, a D01-profile might be: '*Members of a household who currently receives the Human Development Bond and needs to work for a wage. Their parents have a very low level of education, they do not have a desktop computer neither Internet connection, and also have no books or just a very few. In their school, teachers arrive late to class, are no committed with learning and have low expectations about student's future*'. As can be seen, this very detailed information is extremely helpful for developing based-evidence policies, to assign budget and have a successfully deployment.

4 Discussion

This approach provides a natural way to represent educational systems, bringing the opportunity of calibrating control-parameters for assessing order-parameters over time and in multiple territorial scales. Thus, the network analysis underpins current educational deprivation models [10] and provides a parametric way to estimate prevalence of inequality in learning outcomes for contexts with high levels of ethnic- diversity, a key aspect for understanding complexity in intercultural systems that also provides a solution to concerns about the limitations of case studies, the classical utility theory and isolated statistical analysis.

The findings offer evidence in the deep lack in equity that can help policymakers to identify those factors related with inequality in learning outcomes, as well as the magnitude of their relationship with deprivation. In this sense, running the modularity process for defining groups might help to identify those factors which are more relevant for one group and are not for others, avoiding statistical bias based on averages provoked by statistical multilevel modelling and bringing additional information about the order in which factors should be considered. Additionally, each LN can be stressed for calibrating boundaries and initial conditions, as well as to test policies at different levels, from-bottom-to-top and from-top-to-bottom and with real data, before being implemented.

These results confirm that network analysis is becoming fundamental for educational policy, specially linking microdata with other constructs and social macro-parameters. Finally, detecting how gaps in educational achievements are driven by students' context might highlight in a better way where policy can intervene properly, offering a series of unique advantages over classical analysis tools and allowing the intensive use of large-scale assessments and other datasets.

Acknowledgments. We would like to thank ISCTE's Programme in Complexity Sciences, the University of Lisbon and the Latin American Agency for Evaluation and Public Policy, for their support to the project as well as the authors of the articles mentioned in the references.

References

1. UN DESA. The Sustainable Development Goals Report 2018, UN, New York (2018)
2. Shee, N.K.: Karen Education Department's multilingual education for language maintenance, Kasetsart J. Soc. Sci. (2018)
3. Johnson, J.: Hypernetworks in the Science of Complex Systems. Imperial College Press, London (2015)
4. Sánchez, H.: Equity: the focal point of educational quality. Natl. Educ. Eval. Policy Gaz. **4**(10), 42–44 (2018)
5. Kivela, M., Arenas, A., et al.: Multilayer networks. J. Complex Netw. **2**(3), 203–271 (2014)
6. Mizuno, K., Tanaka, M., Fukuda, S., Imai-Matsumura, K., Watanabe, Y.: Relationship between cognitive functions and prevalence of fatigue in elementary and junior high school students. Brain Dev. **33**(6), 470–479 (2011)

7. Price, H.E., Moolenaar, N.M.: Principal-teacher relationships: foregrounding the international importance of principals' social relationships for school learning climates. J. Educ. Adm. **53**(1), 8–39 (2015)
8. Catalan Najera, H.E.: Multiple deprivation, severity and latent sub-groups: advantages of factor mixture modelling for analysing material deprivation. Soc. Ind. Res. **131**(2), 681–700 (2016)
9. OECD. Equity in Education: Breaking Down Barriers to Social Mobility, PISA, OECD Publishing, Paris (2018)
10. Gray, J., Kruse, S., Tarter, C.J.: Enabling school structures, collegial trust and academic emphasis: antecedents of professional learning communities. Educ. Manage. Adm. Leadersh. **44**(6), 875–891 (2016)

CHAPTER 3

Inequality in Learning Outcomes: Unveiling Educational Deprivation Through Complex Network Analysis

First Autor: Harvey Spencer Sánchez Restrepo
Harvey.Restrepo@iscte-iul.pt
Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL
1649-026, Lisbon, Portugal

Coauthor: Jorge Louçã
jorge.l@iscte-iul.pt
Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL
1649-026, Lisbon, Portugal

Running head: **Unveiling Educational Deprivation Through Complex Network Analysis**

Article published in: *Studies in Computational Intelligence*, vol 882. **Springer Nature**.

Inequality in Learning Outcomes: Unveiling Educational Deprivation Through Complex Network Analysis

Harvey Sánchez-Restrepo[1(✉)] and Jorge Louçã[2]

[1] University of Lisbon, Cidade Universitária, 1649-004 Lisbon, Portugal
Harvey_Restrepo@iscte-iul.pt
[2] Information Sciences, Technologies and Architecture Research Center, ISCTE-IUL, 1649-026 Lisbon, Portugal
jorge.l@iscte-iul.pt

Abstract. Understanding which factors are determinant to guarantee the human right to education entails the study of a large number of non-linear relationships among multiple agents and their impact on the properties of the entire system. Complex network analysis of large-scale assessment results provides a set of unique advantages over classical tools for facing the challenge of measuring inequality gaps in learning outcomes and recognizing those factors associated with educational deprivation, combining the richness of qualitative analysis with quantitative inferences.

This study establishes two milestones in educational research using a census high-quality data from a Latin American country. The first one is to provide a direct method to recognize the structure of inequality and the relationship between social determinants as ethnicity, socioeconomic status of students, rurality of the area and type of school funding and educational deprivation. The second one focus in unveil and hierarchize educational and non-educational factors associated with the conditional distribution of learning outcomes. This contribution provides new tools to current theoretical framework for discovering non-trivial relationships in educational phenomena, helping policymakers to address the challenge of ensuring inclusive and equitable education for those historically marginalized population groups.

Keywords: Educational network · Large-Scale Assessments · Policy informatics

1 Introduction

With the establishment of the Sustainable Development Goals (SDGs), the 193 countries attached to Unesco promulgated that 'education is a human right' [1] and that the two pillars of quality in education should be learning and equity [1, 2], recognizing that all human beings have the right to learn and that the State is obliged to guarantee to all citizens equally [3]. Education is also a source for social mobility: an additional year of quality education can increase a person's income up to 10% [4]. The lack of quality in education worldwide is of such magnitude that Unicef estimates that 250 million

© Springer Nature Switzerland AG 2020
H. Cherifi et al. (Eds.): COMPLEX NETWORKS 2019, SCI 882, pp. 325–336, 2020.
https://doi.org/10.1007/978-3-030-36683-4_27

children do not have the minimum learning and that more than half of them have been fooled systemically: despite having managed to attend school, many of them fail in developing the minimum learning such as reading, writing or performing basic operations. The causes for this deprivation of learning are multiple but they affect mainly to those belonging to historically marginalized population groups, almost always, the most impoverished [5–7]. Further, the modest improvements in learning achievements are usually by the hand of enormous inequalities among students, which has cast doubt on the government actions for improving the quality of education [8].

As a strategy to strengthening public policies, many countries have implemented national Large-Scale Assessments (LSA) to collect valid and reliable data on educational outcomes and Factors (variables) associated with Learning (FAL). The objective of LSA is to have evidence-based information and use quantitative methods to estimate educational changes when varying each factor [9] in many scales of time and across the territory [10], as well to modelling dynamic behavior and asymmetries at school and student level [11]. However, dominant models to study educational phenomena are based on tools that postulate that educational gaps can be explained just through the covariation between learning outcomes and each factor [12]. In this kind of studies, the linear analysis provokes a fragmented view of the system and dismiss very often the interactions between agents, FAL and educational phenomena [13, 14], dismissing the structure of the inequality in learning outcomes and its relationship with educational deprivation, which partially explains the lack in explanatory power of those models [15].

To represent the multiple interactions in a system and the emergence of collective properties at different scales from their constituents, several researchers have proposed the use of network theory to model social systems as a result of self-organized processes [17, 18]. Given that 'networks are at the heart of complex systems' [16], analyzing statistical and topological properties of the education system through network theory means studying the complexity of the system through its interactions.

Therefore, this research addresses empirical data for estimating educational deprivation in a Latin American country, as well as its relationship with the most relevant social determinants such as Socioeconomic status (SES) of the student and their families, Rurality in the area where the school is located (RA), the Type of school (TS) and Ethnic self-identification of the student (ET), for estimating topological features and order parameters of the network related to out-of-equilibrium states, providing a new kind of information about global and local properties of the structure of educational deprivation and helping to find those key factors driving inequality gaps in learning outcomes.

1.1 Dataset

In this model, a multivariate dataset integrates learning outcomes of every student who has completed the k-12 education process, estimated by the ability's parameter θ^j through a LSA carried out in Ecuador in 2017 using a standardized computer-based test[1] and integrated with a robust dataset with more than 140 variables coming from

[1] Full dataset is available in http://www.evaluacion.gob.ec/evaluaciones/descarga-de-datos/.

surveys to student's families and teachers. For building the scores, psychometric parameters were estimated by Item Response Theory through a 2P-Logistic model [19] following Eq. 1:

$$P(\theta_j) = \frac{e^{[\alpha_i(\theta_j - \beta_i)]}}{1 + e^{[\alpha_i(\theta_j - \beta_i)]}} \tag{1}$$

$$\text{Con} \quad \theta_j \in (-\infty, \infty), \alpha_i \in (-\infty, \infty) \; y \; \beta_j \in (-\infty, \infty).$$

Raw scores of θ^j were re-scaled to a Learning index ($LI_j \in [4.0, 10.0]$), a monotonous transformation, where higher levels of learning are more likely to have higher scores [19]. For measuring relative deprivation, this model uses the sociological proposal that 'needs, thresholds and satisfactions are determined by each society', while absolute deprivation proposes that 'there is an irreducible nucleus of needs that are common to every human being' [20]. All students are classified in levels of achievement (L_k) based on a standard Bookmark process for establishing psychometrical cut points s_i [19]. Students suffering learning deprivation are those that did not meet the minimum learning standards at the end of the compulsory education, denoted by L_0.

1.2 Deprivation Learning Index

To estimate this index, we use the family of scores $\{LI_j\} j = 1, \ldots, N$, of those students with low level of achievement L_0-class, s_1 is the first cut point —the minimum score to be located at level L_1—. For the L_0-class, absolute deprivation is given by $H = (n(L_0)/\sum n(LI_j)$, where $n(LI_j)$ represents the number of students below the first level of achievement, the intensity $\lambda(LI_j)$ is given by the distance to reach the first level L_1 and the Deprivation learning index (DLI) is given by $\delta_j = \mathrm{H} \cdot \lambda(LI_j)$, which represents a measure of the collective learning deficit, which considers the magnitude —the number of students with low performance— and intensity —how much below the minimum performance level are located [20].

1.3 Model Specification

For analyzing topological properties and pointing out those nodes holding the system out of statistical equilibrium, a three sequential steps model was developed. The first step is focused on disaggregating L_0-class by *SES*, each student is represented by a node and edges, weighted by $\lambda(LI_j)$, are directed to one of the SES-decile nodes $\{LI_j(\theta^j \rightarrow L_k^j \rightarrow (SES_d^j)\} \; \forall \, j$, a process which allows to analyze aggregated inequality at school level, as well as In-degree distribution for SES nodes. The second one is an extension for including RA, TS and ET to analyze their effects through the sequence $\{\theta^j \rightarrow L_k^j \rightarrow (SES_d^j) \rightarrow (RA_{C1}^j, TS_{C2}^j, ET_{C3}^j)\} \; \forall \, j$, where C denotes an index for each subcategory of the factors RA, TS and ET. Finally, the third step amplifies and strengthens the analysis through more than one hundred educational and non-educational factors associated with learning achievements trough the sequence $\{\theta^j \rightarrow L_k^j \rightarrow (SES_d^j) \rightarrow (RA_{C1}^j, TS_{C2}^j, ET_{C3}^j) \rightarrow FAL_{Cm}^j\} \; \forall \, j$, for integrating m different

educational and non-educational factors. Network analysis was carried out by Gephi 0.9.2 and statistical estimations and plots with R 3.5.0 and Orange 3.3.8.

2 Socioeconomic Status and Learning Deprivation

The first specification estimates the Weighted In-degree distribution of directed edges from social determinants nodes to those representing socioeconomic deciles, given by $\{SES_d\}d \in \overline{1, 10}$, where each edge represents one student in L_0-class. As inequality implies asymmetries, in conditions of total equity —where socioeconomic factors would not produce differences— we might expect equal distribution of L_0-edges over the network. Therefore, the study of equity can be deepened by analyzing the levels of absolute and relative deprivation experienced by different population groups and their relationship with the SES of the students.

According with LSA estimates, 8 438 of 39 219 students are in L_0-class, a prevalence rate of 0.215, a $LI = 6.32$ and intensity of deprivation $\lambda = 0.225$, i.e., in average, L_0-student lacks 0.68 standard deviations (SD) of the minimum learning. As shown in Fig. 1, distribution of In-degree $P(SES_d)$ is non-uniform and it decreases monotonically as the SES of the group increases.

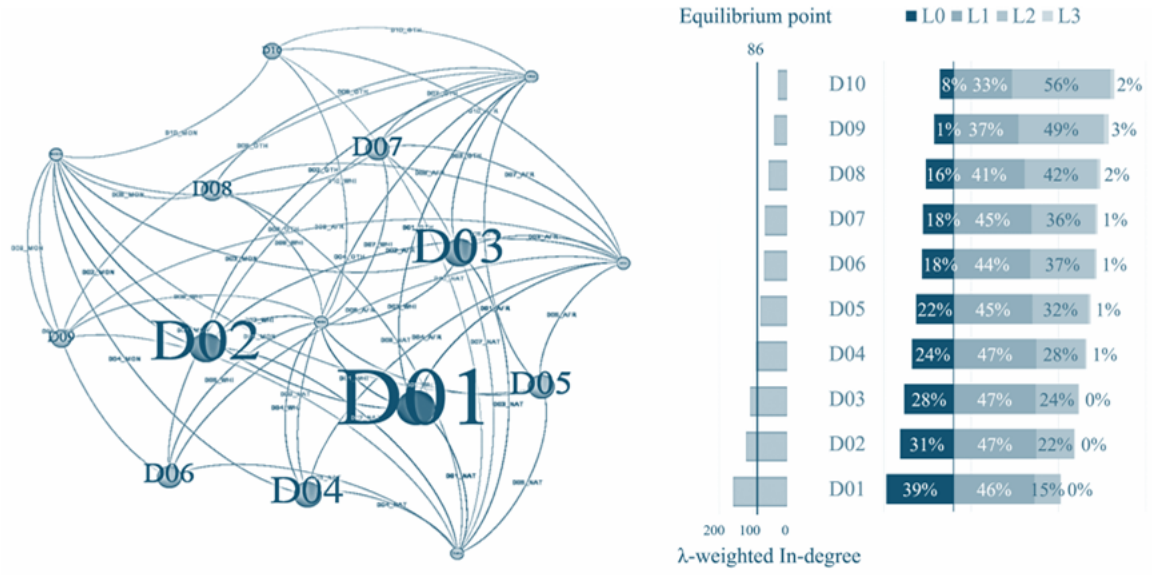


Fig. 1. Network of socioeconomic distribution with λ-weighted learning deprivation and distribution of students among levels of achievement.

For example, $P(SES_1) = 0.39$ and $P(SES_{10}) = 0.08$, this difference of 0.31 means that for each richest-family student who does not learn the minimum, there are 5 poorest-family students in the same situation. As we will see later, this unfortunate situation deepens in rural areas, where the ratio increases to one richest-student for every 7 poorest-students.

Differential centrality of deciles in Fig. 1 also shows a non-equilibrium system driven by SES where poorest students dominate the graph: nodes SES_1, SES_2 and SES_3 have stronger connections and the highest prevalence rate ($DPR = 0.3860$), Hub parameter ($H = 0.4887$), Weighted In-degree ($WID = 158.9420$) and PageRank ($PR = 0.4289$). On the contrary, the richest students grouped by SES_{10} are relative irrelevant for the network with parameters $DPR = 0.0800$, $H = 0.1133$, $WID = 27.222$ and $PR = 0.0289$. As can be seen, in-degree parameter provides an estimation for deprivation rates of each SES_d, showing the size of the gaps among them through detecting *SES* effects in nodes grouping L_0-students by deciles, pointed out by the negative correlation between *LI* and *SES* ($R = -0.58$, $p < 0.001$).

To estimate more accurately the cumulative effect of *SES* on learning outcomes, Fig. 2 shows *LI* (left plot) and intensity of deprivation (right plot), as functions of *SES* in two dynamical ways: (1) starting with the whole population of students and re-estimating *LI* and λ while excluding poor-students (black circles), and (2) starting with just poor-students and including richer students (white circles). As can be seen, the biggest circles indicate equal Global Average (GA) for *LI* (7.61) and λ (0.225), however, for case (1) the effect of removing poor students is that GA goes up immediately after removing SES_1-students, reaching a $LI = 8.21$ (0.65 SD away from GA), as well as λ diminishes 14%. On the contrary, when the estimation process starts with just the poorest students, $LI = 6.52$ (−1.09 SD below from GA) and starts going down after including SES_3 students. In summary, these opposite behaviors show a learning gap of 1.82 SD, equivalent to almost two years of formal schooling between richest and poorest students, and pointing out that, even among those deprived students, the most impoverished get the worst part suffering 14% deeper intensity of deprivation.

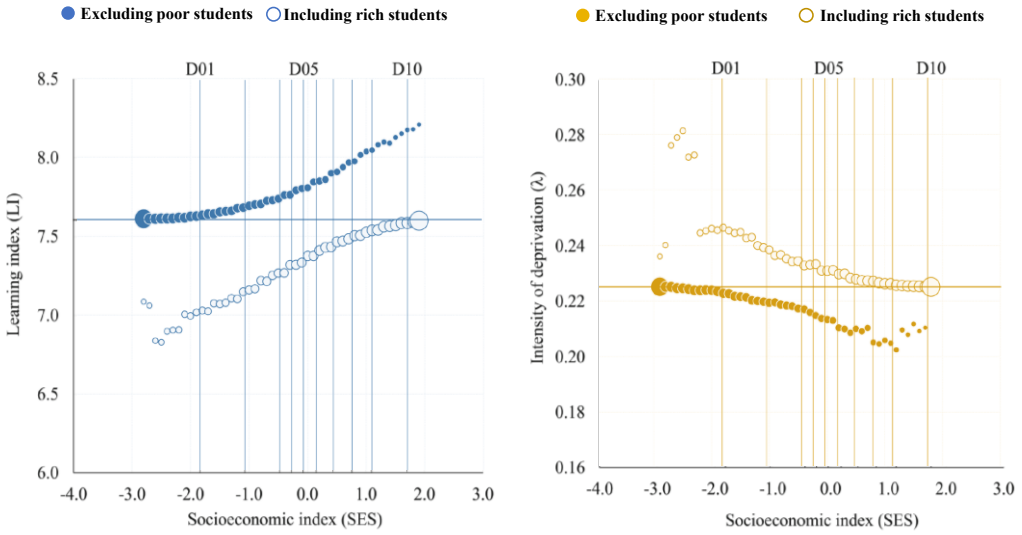


Fig. 2. Cumulative effect of SES in learning outcomes (*left side*) and intensity of deprivation (*right side*).

The previous results provide very clear information about learning inequalities at student level, however, to analyze an aggregated phenomenon, the school level shows local but systemic gaps to better understand the sources of structural deprivation. When studying schools as integrated units, the impact of *SES* becomes even more evident in learning outcomes, in left side of Fig. 3, schools are splatted between type of school — sources of funding—, here each school is represented by a circle whose size is proportional to its number of students. The average *SES* of students is located on the horizontal axis and *LI* is on the vertical one. As can be seen, private schools have higher *SES* (0.42) and *LI* (7.96) than the public ones (*SES* = −0.21) and (*LI* = 7.44), and there is also a strong positive relationship between *LI* and *SES*.

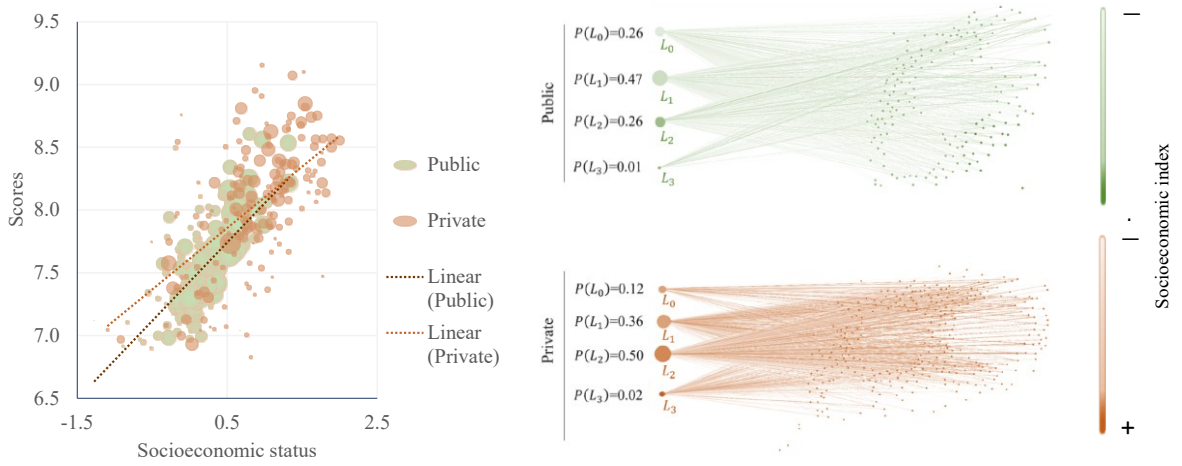


Fig. 3. Relationship between learning and socioeconomic indexes at school level (*left side*), and In-degree distribution of private and public networks (*right side*).

Moreover, right side of Fig. 3 shows a kind of bipartite data sources [21] for representing the networks of private and public schools, where darker edges refer to higher SES and the size of nodes are masses of probability for in-degree distribution $P(L_i)$ where Private-network has 309 nodes (schools) with 12 823 edges (41 per school) while Public-network has just 167 with 24 671 edges (148 per school), i.e., the private one has 85% more nodes than the public one, but its density is just 52%.

In addition, 70.6% of private schools are linked to L_0 through just 1 553 of their edges (H = 0.12 and λ = 0.227), while 97.0% of public are linked to the same node through 6 339 edges (H = 0.26 and λ = 0.223), a rate twice higher than in private sector, showing that SES is a key factor for educational deprivation due mainly to the influence of cultural capital, showing the lack in the capacity of the government to guarantee educational rights.

3 Ethnicity and Type of School Financing

The rurality of the area where the schools are located is also a factor that impacts on the learning gaps, its effect, combined with the type of school funding, indicates a huge variance among ethnic groups within the socioeconomic deciles. Figure 4 shows a Radviz plot, a non-linear multi-dimensional algorithm [22] and a dendrogram made with a cluster analysis based on $\left(SES_k^i, RA_k^i, TS_k^i, ET_k^i\right)$, this representation shows the hierarchical structure of inequality in learning deprivation for population groups, the points on the circle refer to three indicators attracting the students groups proportionally and the symbol size is proportional to the total deprivation given by δ.

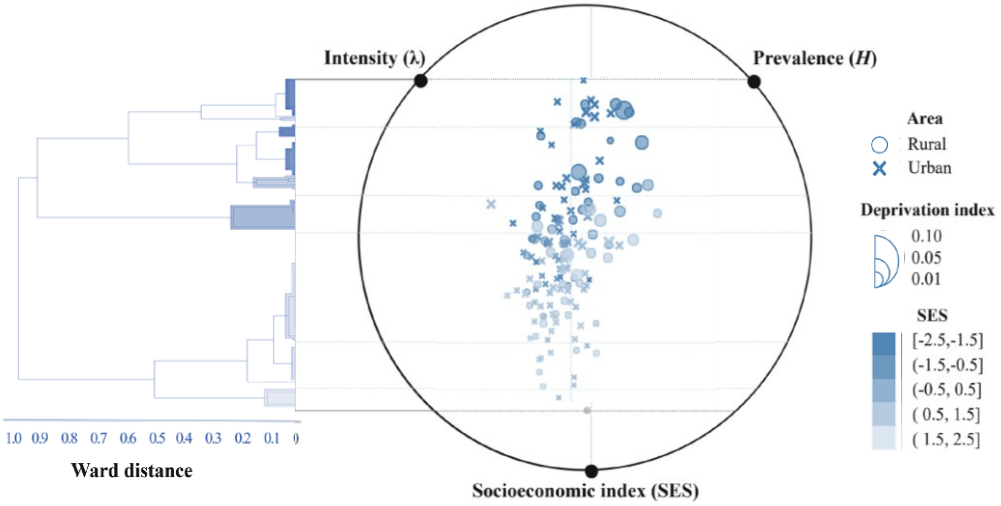


Fig. 4. Deprivation of learning by ethnicity, rurality and socioeconomic status.

As can be seen, the strong relationship between the SES and the prevalence distributes the groups in a vertical way showing borders between rural (circles) and urban areas (crosses).

Figure 4 shows that gaps between rural and urban areas are quite accentuated, especially among the poorest students in rural areas suffering the highest levels of prevalence and δ. In this sense, among the most deprived students in rural areas, Montubios exhibit the highest prevalence rates, followed by and Indigenous (H = 0.64) and Afro-Ecuadorians (H = 0.55). The clusters shown by the dendrogram points out that the highest SES students dominates private schools with the lowest levels of deprivation, with the exception of Montubios in rural areas, where no student attending private schools reaches the minimum level and where just 1 of 3 students in public schools is not deprived (H = 0.63), showing that social and ethnic classes are splatted into groups of students who have had different learning opportunities inside and outside of the schools, helping to understand how inequality evolves in a structural way [23].

4 Key-Factors for Public Policies

One of the most useful strategies for improving learning and closing gaps is micro-planning, however, to select which needs should be attended first for specific groups is always a great deal. For this reason, once the gaps in population groups are measured, we extend the network for introducing $\{FAL^{j}_{Cm}\}$ to the model and identify which variables are linked with specific population groups and recognizing those that should be intervened first, as well to simulating the variational aspects for establishing the order and control parameters for building more efficient and effective portfolios of policies at local level.

Prevalence rate of L_0-students are related with networks strongly connected and this might be associated with Eigen-Centrality through measuring the influence of each factor for identifying how well connected the j-th factor (FAL_j) is and how many links have its connections. In this way, splitting L_0-students in communities becomes in a very valuable tool for developing group-oriented strategies and avoid implementing the same actions for completely different needs. At this point, network simulation offers the advantage of building multiple scenarios across varying *FAL* parameters for recognizing and ranking the most relevant nodes to be attended by policymakers in a hypothetic but very specific situation.

Figure 5 shows a network build through the *ForceAtlas2* algorithm [24] to obtain the layout after a *Modularity process* (with parameter 0.073 at resolution of 0.254) for splitting richest from poorest students to find key factors for educational deprivation in both groups [25]. The Average Weighted Degree of the network is AWD = 426.9 and its density D = 5.617. Those large nodes appearing in the network as attractors correspond to two communities: the richest (SES_{10}), with degree of authority $A_{SES_{10}}$ = 0.158 and just 20 factors —most of them showing very low centrality parameters— and poorest students (SES_1), with $A_{SES_1} = 0.987$ and 57 different factors associated with their deprivation.

For assessing the quality of its connections, PageRank is a well-known algorithm for providing accurate and clear information when looking for the factors dominating deprivation in each community [25, 26], which also might order nodes parametrically for selecting those who susceptible to be managed by policy and define 'needs profiles' for focusing actions on those variables susceptible to be managed by policy for specific zones or groups.

For example, a SES_1 profile based on Fig. 5 using the 10 most relevant factors might be:

> 'Members of a household who currently receives the Human Development Bond and needs to work for a wage. Their parents have a very low level of education, they do not have a desktop computer neither Internet connection, and also have no books or just a very few. In their school, teachers arrive late to class, are not committed with learning and have low expectations about student's future'.

As can be seen, this very detailed information is extremely helpful for developing based-evidence policies, to estimate and assign budget and increase the chances of having a successful deployment, the most wanted tool from policymakers.

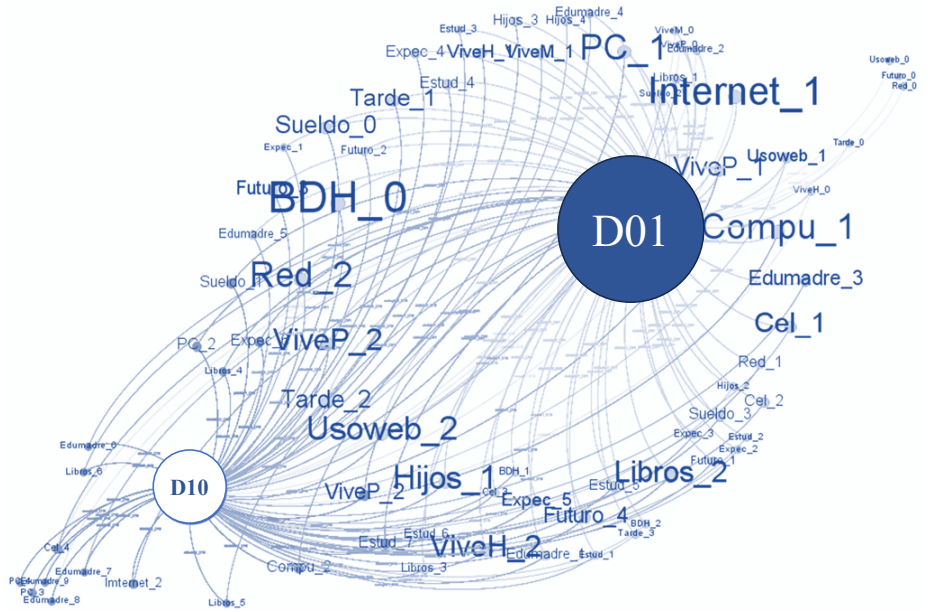


Fig. 5. Factor's network for richest and poorest students' communities. Centrality of nodes in each group points out the dissimilarities in factors provoking deprivation.

5 Discussion

This research finds a lot of evidence about the deep lack in equity in a Latin American country and its relationship with social determinants for synthetizing the complex structure of inequality on learning outcomes [27]. Measuring how far is each student of reaching the minimum level to ensure its right to education is crucial because the social deception of accessing education without guarantees of learning, converges at levels of precariousness similar to those found in people outside the scholar system [28]. Likewise, to identify a specific set of factors impacting a group of students that might not be relevant for others is a key issue for developing group-oriented policy [29].

This sequential network model allows to investigate the factors conditioning the exercise of the right to education and opens the possibility of reviewing how the type of school funding promotes inequality in different population groups at country level. Evidence found over deprivation among the socioeconomic groups shows the lack of attention to ethnicity, especially in the urban area. In this sense, when comparing the topologies of the networks of the richest versus the poorest was possible to estimate the gaps in educational outcomes that are associated to other factors like the influence of cultural capital, the context and the families of the students, beyond the schools.

Through this model is also possible to use LSA results for analyzing the interplay of socioeconomic status, rurality, type of school and ethnicity, bringing very useful information about the educational system at meso and macro levels from micro level interactions, offering valuable information for answering five main educational questions:

1. What is the level of deprivation of learning in students at the end of compulsory education?
2. Is learning deprivation associated with schools?
3. How deep is the inequality in the distribution of learning outcomes?
4. What factors are associated with educational deprivation in specific areas of the territory and population groups?
5. Considering limited resources and time, what factors and in what order should they be established in a group-oriented public policy?

Finding answers to these questions using a dataset from a LSA represents a tipping point for integrating thinking tools into current theoretical frameworks to find non-trivial relationships between social conditions and educational deprivation in historically marginalized population groups. In this sense, estimated gaps in deprivation due to rural area and ethnicity, point out the lack of effective inclusion policies, especially of those groups of students who have had different learning opportunities and highlights how inequality is structurally generated in the country.

However, though it is possible to infer the central role played by the schools in determining the more relevant factors for educational deprivation in different SES classes, more research is needed to understand the interactions between the school context and ethnic diversity. Nevertheless, the implications for public policies are: (1) programs aimed at indigenous self-identified students from the poorest quintile should be reviewed in all its pedagogical aspects to guarantee the achievement of meaningful learning; (2) the interculturality programs in the urban area should be broadened so that the ethnic groups can reach the minimum standards at same rates than the other ethnical groups and, (3) it is necessary to better regulate the private schools in the rural area to help close the gap in the absolute deprivation rate.

There is no doubt that the network analysis shows a great vein of scientific development that has not yet been explored to improve knowledge about educational systems, addressing the greatest challenges of educational research, and helping policymakers to develop strategies for an inclusive and equitable education for all.

References

1. OECD: Equity in education: breaking down barriers to social mobility. PISA, OECD Publishing, Paris (2018)
2. UN DESA: The sustainable development goals report 2018, UN, New York (2018)
3. Sánchez-Restrepo, H.: Equity: the focal point of educational quality. National Educational Evaluation Policy Gazette in Mexico, Year 4. (10), pp. 42–44 (2018)
4. United Nations Children's Fund (Unicef): The investment case for education and equity, Washington, DC (2015)
5. Keeley, B., Little, C.: The State of the World's Children 2017: Children in a Digital World. UNICEF, New York (2017)
6. Samman, E.: SDG progress: fragility, crisis and leaving no one behind: report (2018)
7. Gray, J., Kruse, S., Tarter, C.J.: Enabling school structures, collegial trust and academic emphasis: antecedents of professional learning communities. Educ. Manag. Adm. Leadersh. **44**(6), 875–891 (2016)

8. Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación – LLECE. Agenda (2018). http://www.unesco.org/new/en/santiago/press-room/newsletters/newsletter-laboratory-for-assesSSent-of-the-quality-of-education-llece/
9. Motl, T.C., Multon, K.D., Zhao, F.: Persistence at a tribal university: factors associated with second year enrollment. J. Divers. High. Educ. **11**(1), 51 (2018)
10. Tracey, B., Florian, K. (eds.): Educational Research and Innovation Governing Education in a Complex World. OECD Publishing, Paris (2016)
11. Davis, B., Sumara, D.: Complexity and Education: Inquiries into Learning, Teaching, and Research. Routledge, New York (2014)
12. Fox, J.: Applied Regression Analysis and Generalized Linear Models. Sage Public, Thousand Oaks (2015)
13. Moll, K., Göbel, S.M., Gooch, D., Landerl, K., Snowling, M.J.: Cognitive risk factors for specific learning disorder: processing speed, temporal processing, and working memory. J. Learn. Disabil. **49**(3), 272–281 (2016)
14. van der Maas, H., Kan, K., MarSSan, M., Stevenson, C.: Network models for cognitive development and intelligence. J. Intell. **5**(2), 1–17 (2017). https://doi.org/10.3390/jintelligence5020016
15. Gelman, A., Imbens, G.: Why high-order polynomials should not be used in regression discontinuity designs. J. Bus. Econ. Stat. 1–10 (2018)
16. Barabási, Albert-László, Pásfai, Márton: Network Science. Cambridge University Press, Cambridge (2016)
17. Johnson, J., Fortune, J., Bromley, J.: Systems, networks, and policy. In: Non-Equilibrium Social Science and Policy, pp. 111–134. Springer, Cham (2017)
18. Grandjean, M.: Complex structures and international organizations [Analisi e visualizzazioni delle reti in storia. L'esempio della cooperazione intellettuale della Società delle Nazioni]. Memoria e Ricerca **25**(2), 371–393 (2017)
19. Borsboom, D., Molenaar, D.: Psychometrics. In: Wright, J. (ed.) International Encyclopedia of the Social & Behavioral Sciences, vol. 19, pp. 418–422. Elsevier, Amsterdam (2015)
20. Bracho, T.: Índice de déficit en competencias? Avanzamos hacia la garantía del derecho a la educación? en Reformas y Políticas Educativas, No. 4, September–December 2017. FCE, México (2017)
21. van Borkulo, C.D., Borsboom, D., Epskamp, S., Blanken, T.F., Boschloo, L., Schoevers, R. A., Waldorp, L.J.: A new method for constructing networks from binary data. Sci. Rep. **4** (5918), 1–10 (2014)
22. Leban, G., Zupan, B., et al.: Vizrank: data visualization guided by machine learning. Data Min. Knowl. Discov. **13**(2), 119–136 (2006)
23. Koller, D., Friedman, N.: Probabilistic Graphical Models: Principles and Techniques. MIT Press, Cambridge (2009)
24. Jacomy, M., Venturini, T., Heymann, S., Bastian, M.: ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. PLoS ONE **9**(6), 1–18 (2013)
25. Forbush, K., Siew, C., Vitevitch, M.: Application of network analysis to identify interactive systems of eating disorder psychopathology. Psychol. Med. **46**(12), 2667–2677 (2016)
26. Isvoranu, A.M., van Borkulo, C.D., Boyette, L., Wigman, J.T.W., Vinkers, C.H., Borsboom, D., Group Investigators: A network approach to psychosis: pathways between childhood trauma and psychotic symptoms. Schizophr. Bull. **43**(1), 187–196 (2017)
27. Sánchez-Restrepo, H., Louçã, J.: Topological properties of inequality and deprivation in an educational system: unveiling the key-drivers through complex network analysis. In: International Conference on Human Systems Engineering and Design: Future Trends and Applications, pp. 469–475. Springer, Cham, September 2019

28. Tsatsaroni, A., Evans, J.: Adult numeracy and the totally pedagogised society: PIAAC and other international surveys in the context of global educational policy on lifelong learning. Educ. Stud. Math. **87**(2), 167–186 (2014)
29. Menefee, T., Bray, T.M.: Education in the Commonwealth: Quality Education for Equitable Development. Commonwealth Secretariat (2015)

CHAPTER 4

A Complex Network Approach to Structural Inequality of Educational Deprivation in a Latin American country

First Autor: Harvey Spencer Sánchez Restrepo
Harvey.Restrepo@iscte-iul.pt
Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL
1649-026, Lisbon, Portugal

Coauthor: Jorge Louçã
jorge.l@iscte-iul.pt
Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL
1649-026, Lisbon, Portugal

Running head: **Educational Deprivation in Latin America**

Article published in: *Springer Proceedings in Complexity*. **Springer Nature**.

A Complex Network Approach to Structural Inequality of Educational Deprivation in a Latin American country

Harvey Sánchez-Restrepo [1][0000-0001-7311-8622] and Jorge Louçã [2][0000-0003-4766-5627]

[1] Faculty of Sciences, University of Lisbon, Cidade Universitária, 1649-004, Lisboa, Portugal
harvey@comunidad.unam.mx
[2] Information Sciences, Technologies and Architecture Research Center. ISCTE-IUL,1649-026, Lisbon, Portugal
jorge.l@iscte-iul.pt

Abstract. To guarantee the human right to education stablished by the fourth UNESCO's Sustainable Development Goal, a deep understanding of a big set of non-linear relationships at different scales is need it, as well as to know how they impact on learning outcomes. In doing so, current methods do not provide enough evidence about interactions and, for this reason, some researchers have proposed to model education as a complex system for considering all interactions at individual level, as well as using computer simulation and network analysis to provide a comprehensive look at the educational processes, as well as to predict the outcomes of different public policies.

The highlight of this paper is modeling the structure of the inequality of a national educational system as a complex network from learning outcomes and socio-economic, ethnicity, rurality and type of school funding, for providing a better understanding and measuring of the educational gaps. This new approach might help to integrate insights improving the theoretical framework, as well as to provide valuable information about non-trivial relationships between educational and non-educational variables in order to help policymakers to implement effective solutions for the educational challenge of ensuring inclusive and equitable education.

Keywords: Structural network, Large-scale assessments, Policy informatics.

1 Introduction

The 193 countries attached to Unesco promulgated the Sustainable Development Goals (SDG), the fourth goal (SDG-4) establishes that ''education is a human right' [1,2] and that the essential axes of quality in education must be learning and equity, since that all human beings have the right to learn, the State is obliged to guarantee the exercise of this right to all citizens equally [3,4]. However, Unesco points out that inequality in education has progressively increased and that the most impoverished bear the worst consequences [5,6]. In Latin America, the results of the last Large-scale Assessment of learning (LSA) show that most of the countries have high percentages of children with low-level learning after several years of schooling [7].

At global level, the lack of quality in education is of such magnitude that Unicef estimates that 250 million children, after having attended school, fail to develop the minimum learning in Literacy and Numeracy, both necessary to continue learning at the following educational levels [8]. For facing this challenge, many governments have promoted multiple reforms to improve the quality of their education systems, however, the modest improvements in learning have been accompanied by huge inequalities between different population groups, raising many questions to the policymakers [9,10].

In this context, to design better public policies, some countries have created national evaluation systems for measuring outcomes results and a pool of associated factors to learning (FAL) [4]. Most of those systems used standard tests based on psychometrical analysis and statistical models oriented to recognize the factors with a significant covariation, the main idea is estimating the average educational gains that a system might experience by varying one factor at a time through a specific policy [11,12]. Of course, the connection between constructs and observations has been a constant hypothesis of multiple experimental researches, however, one of the main questions to these actions is that, in daily practice, a fragmented vision of the system predominates, in which the interactions between the factors and the educational phenomena in their different scales are not considered [13,14].

Like most social systems, the educational ones exhibit non-linear relationships among their multiple agents with different levels of organization and time horizons that allow the emergence of self-organized phenomena, producing dynamic equilibria at different scales. However, dominant models for studying educational phenomena are based on reductionist tools that postulate that educational systems can be understood as the sum of their constituents where interactions among them are irrelevant. Therefore, the observed phenomenon is understood and modeled using one-to-one covariations for explaining lack or improvements in learning outcomes, dismissing the information coming from interactions between variables and their relevance in the structure of the entire system. Despite the historical success in some social research, in education this approach is, most of the times, just descriptive and lack of explanatory power [15].

In the last years, models based on Network science have been proposed as an alternative representation of systems, overall, because many systems can be described by complex interconnected networks as a result of self-organized processes [16]. One of the main advantages of modeling educational systems as a network, or multiple networks, is establishing connections between population characteristics of individuals as random phenomena with probabilities of occurrence given by data. This approach allows the identification of the factors that influence learning based on both topological and statistical parameters for unveiling some hierarchical structures related with inequality and educational deprivation.

For providing a robust model for better understanding inequality gaps, in this research we use concepts from network science as a tool for studying complexity and global and local properties of the structure of inequality in learning outcomes observed in a Latin American country. The analysis is based on statistical properties of the networks related with low-level-of-proficiency students and the Socio-economic Status

(SES) of the student's family, Rurality of the area where the school is located (RA), Type of school (TS), and Ethnicity (ET) for analyzing out-of-equilibrium states [19].

1.1 Dataset

For developing the model, a multivariate dataset integrates learning outcomes of every student who has completed the *k-12* education process, estimated by the ability's parameter θ^j through a LSA carried out in Ecuador in 2017, using a standardized computer-based test[1] and integrated with a robust dataset with more than 240 variables coming from surveys to student's families and teachers. For building the scores, ability and psychometric parameters were estimated by Item Response Theory as usual, through a 2P-Logistic model [19,20], following equation 1:

$$P(\theta_j) = \frac{e^{[\alpha_i(\theta_j-\beta_i)]}}{1+e^{[\alpha_i(\theta_j-\beta_i)]}} \text{ with } \theta_j, \alpha_i, \beta_j \in (-\infty, \infty) \quad (1)$$

After estimation process, raw scores were re-scaled to a standardized Learning index ($LI_j \in [4.0,10.0]$), a monotonous transformation of θ^j, where higher levels of learning are more likely to have higher scores [19,20]. The scores are on a continuous scale corresponding to four levels of achievement, according with the LSA and national standards, according with a technical description[2], all students are classified in Levels of Achievement (LA) stablished by a Bookmark process carried out by an expert pedagogical group on each subject [19]. The three psychometrical cut points s_1, s_2, s_3 correspond to four LA, where L_0 corresponds to those who have not reached a minimum level of learning, L_1 to the minimum acceptable, L_2 at the level of achievement raised by the system and L_3 corresponds to a performance higher than the standard.

1.2 Deprivation learning index

For estimating the Deprivation Learning Index (DLI), we use the family of scores $\{LI_j\}\, j = 1, \dots, N,$ of those students with low level of achievement L_0-class where s_1 is the first cut point —the minimum score to be located at level L_1. They are students suffering learning deprivation according with the sociological proposal that states 'there is an irreducible nucleus of needs that are common to every human being', while relative deprivation, estimated by LI_j for each student, becomes from 'needs, thresholds and satisfactions are determined by each society' [21], both established by the Bookmark process.

For the L_0-class, absolute deprivation is given by H=$(n(L_0)/\sum n(LI_j)$, where $n(LI_j)$ represents the number of students below the first LA, the intensity $\lambda(LI_j)$ is given by the distance to reach the first level L_1, then DLI is given by δ_j=H·$\lambda(LI_j)$, which

[1] Full dataset is available in http://www.evaluacion.gob.ec/evaluaciones/descarga-de-datos/, selecting the option *Ser Bachiller 2017-2018* and *microdato* for downloading the full data.

[2] Technical and pedagogical details about design can be found in http://www.evaluacion.gob.ec/evaluaciones/ser-bachiller/

represents a measure of the collective learning deficit, which considers the magnitude —the number of students with low performance— and intensity —how much below the minimum performance level are located [21].

As can be seen, histogram in Figure 1 shows that 22% does not meet the learning minimums at the end of the compulsory cycle.

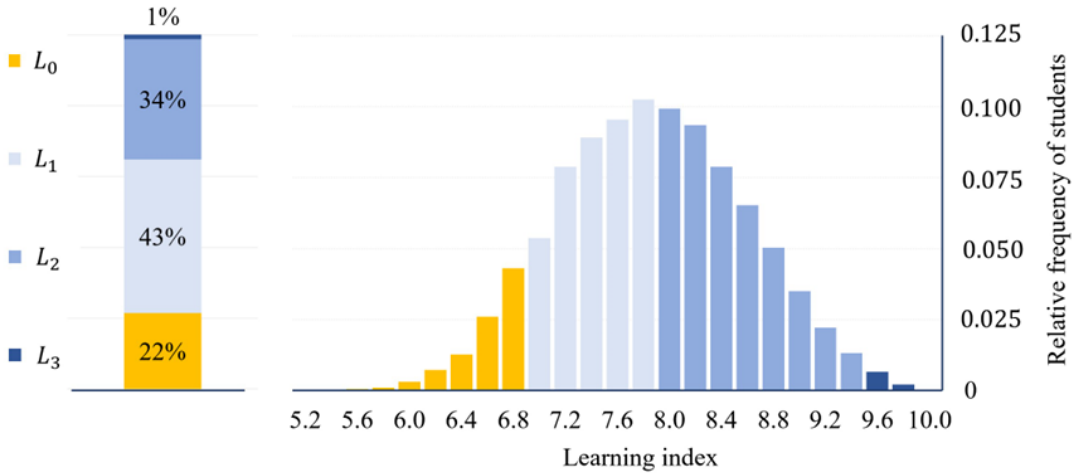


Fig. 1. Distribution of the students among Learning index (scores).

Scores distribution in Figure 1 allow the study of equity and it can be deepened by analyzing the levels of deprivation —absolute and relative— experienced by different population groups and their relationship with the socioeconomic status and ethnicity of students.

1.3 Model specification

In the last years, models based on graph theory have been proposed as a parallel representation of psychometric constructs such as intelligence, leadership or depression [22]. These models have in common that the covariance among the observable variables could be explained from the identification of patterns found among a set of interactions between these variables, measured through a set of informative items. The model for building the network is based on interactions between two nodes, representing the level of learning of each student directed to the set of each factor categories, where the edge is weighted by $\lambda(LI_i)$ [23].

For carrying on the analysis, the model runs in three phases: 1) analyzing the scores for assigning a LA to each student for identifying those located in level L_0, 2) estimating the SES for aggregated levels and subpopulation groups using the cut points for splitting in deciles, and 3) analyzing the associated factors to learning for creating the family of sequences $\{\theta^j \rightarrow L_k^j \rightarrow (SES_d^j)\}\ \forall\ j$ [24,25] for each student, where L_k^j is the LA, and SES_d^j corresponds to the SES decile of the j-th student.

To carry out a refinement of the variables that increase the deficit of basic skills in the population, the SES and the estimated deciles for the general population are preserved during all the stages, all other estimates are made again over the L_0 −group.

As each student is represented by a node, a set of edges, weighted by $\lambda(LI_j)$, are first directed to one of the SES-decile nodes $\{\theta^j \rightarrow L_k^j \rightarrow (SES_d^j)\}\, \forall\, j$, a process which allows to analyze aggregated inequality at school level, as well as In-degree distribution for SES nodes.

For extending the model and knowledge about social determinants, TS, RA and ET are included in the analysis one by one for analyzing their effects through the sequence $\{\theta^j \rightarrow L_k^j \rightarrow (SES_d^j) \rightarrow (TS_{C2}^j, RA_{C1}^j, ET_{C3}^j)\}\, \forall\, j$, where C denotes an index for each subcategory of the factors RA, TS and ET. Network analysis was carried out by Gephi 0.9.2 and statistical estimations and plots with R 3.5.0 and Orange 3.3.8.

2 Socioeconomic status and student's learning outcomes

To estimate the size of the gap at the macro level, the first network shown in Figure 2 integrates the Weighted In-degree distribution of directed edges from nodes indicating subpopulation groups to those representing SES deciles, given by $\{SES_d\}\, d \in \overline{1{,}10}$, where each edge represents one student in L_0-class. As inequality implies asymmetries, in conditions of total equity —where socioeconomic factors would not produce differences — we might expect equal distribution of L_0-edges over the network for all deciles, but the distribution is not like that. Therefore, the study of equity can be deepened by analyzing the levels of absolute and relative deprivation experienced by different population groups and their relationship with the SES of the students.

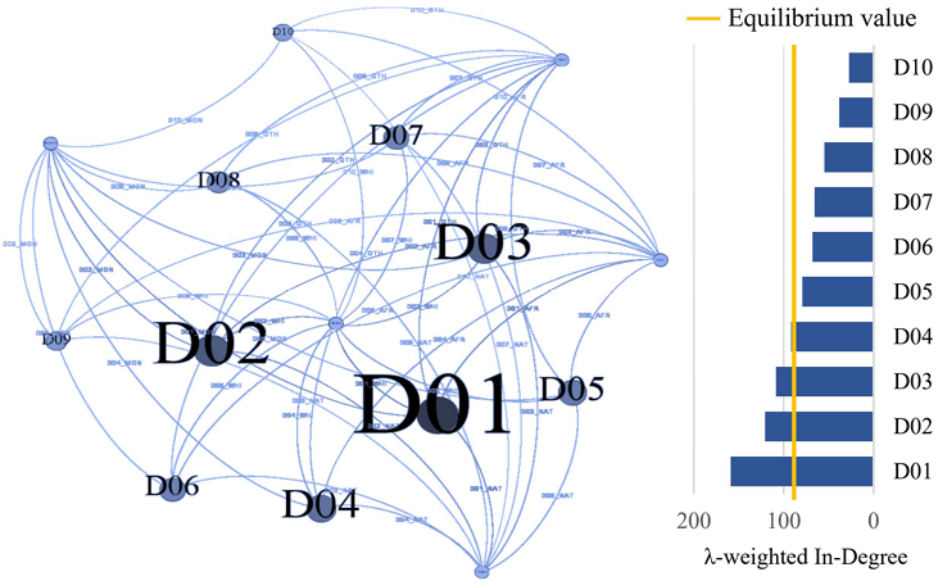


Fig. 2. Network for Weighted Out-degree and its edges' histogram shows socioeconomic status distribution of deprived students (L_0-group).

According with estimates, 21.5% of students are in L_0-class, a prevalence rate of 0.215 corresponding to a $LI = 6.32$ and shows an intensity of deprivation λ=0.225, i.e., in average, L_0-student lacks 0.68 standard deviations (SD) of the minimum learning.

To estimate the size of the gap at the macro level, Figure 2 also shows the percentages of students in each level of achievement for each SES decile. As can be observed, the proportion of students in each decile decreases monotonically as the SES of the group increases, being for the first decile (D01), 39% of students, and only 8% in D10. This difference of 31 percentage points is equivalent to the fact that for each rich family student who does not learn the minimum, there are 5 poor in the same situation. As will be shown later, this situation deepens in rural areas, where the ratio increases to one rich student for every 7 poor students.

3 The relationship between Type of school and SES

When studying schools as integrated units, the impact of SES becomes even more evident, in Figure 3 each school is represented by a circle whose size is proportional to the number of its students enrolled, the source of funding is distinguished by the color: green for private, blue for public. The average SES of students is located on the horizontal axis and the average score of LSA on the vertical axis. In the right side, the two whisker-box plots show dispersion for both indexes.

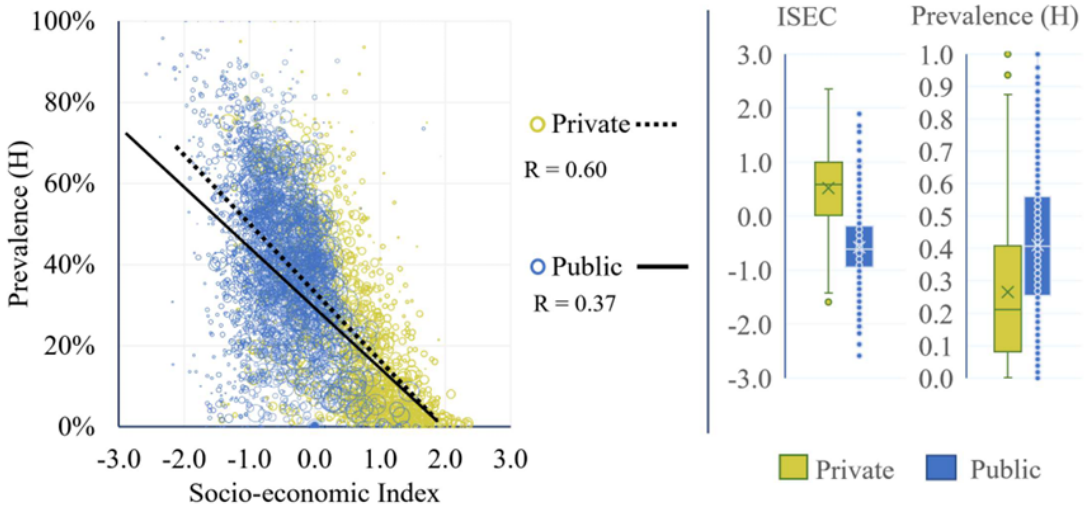


Fig. 3. Relationship between learning prevalence of deprivation and socioeconomic status at school level and box plots disaggregated by type of funding.

The negative correlation between SES and prevalence rate (R= —0.55, p<0.001) shows the separation of SES classes in groups of students who have different learning opportunities inside and outside schools, which helps to understand how inequality is gestated in a structural way in the country: schools with high SES predominate in the private sector and it is also there that the lowest levels of deprivation are presented, in this sector the correlation coefficient between the SES and H index is (R= —0.60, p<0.001). On the contrary, public schools that serve the poorest students have higher prevalence rates and a lower correlation (R = —0.39, p <0.001), which could indicate

that, a weight of which the deprivation of learning is higher for the whole group, there is less inequality motivated by the socioeconomic origin of the student.

The socioeconomic gap between private (0.64) and public (—0.18) schools accumulates 0.82 standard deviations (SD), in addition, the prevalence rate in the public sector (H=0.219) is 1.7 times that of the private sector (H= 0.374). So, this confirms that public schools not only serve the poorest students in the country, but as they do so in the most depressed and most difficult areas, attendance is an extra challenge reflected in prevalence rates, while private schools concentrate on students in the top quintiles and the prevalence in most cases does not exceed 30%.

4 Inequality gaps and marginalized population groups

For having a more detailed and in-depth analysis, a selection of the two opposite SES population groups were selected as attractor nodes in the network —deciles D01 (Green) and D10 (Pink)—. In Figure 4, the network integrates the different ethnic groups of the country, disaggregated by rural and urban areas. Both, the nodes and the labels, are proportional to the *Weighted Out-Degree*, and $\lambda(LI_j)$ is weighting the edges.

The representation is based on the Eigen Centrality to measure the influence of a factor in terms of the number of edges with the population groups and that appear as other nodes within the network [27,28]. This measure is very valuable because, by counting how well connected a node is, and how many links have its connections through the network [29], the preponderance of the factors in the population groups becomes very clear in identifying the effect of the three educational deprivators: 1) rural areas, 2) types of funding of schools, and 3) ethnicity.

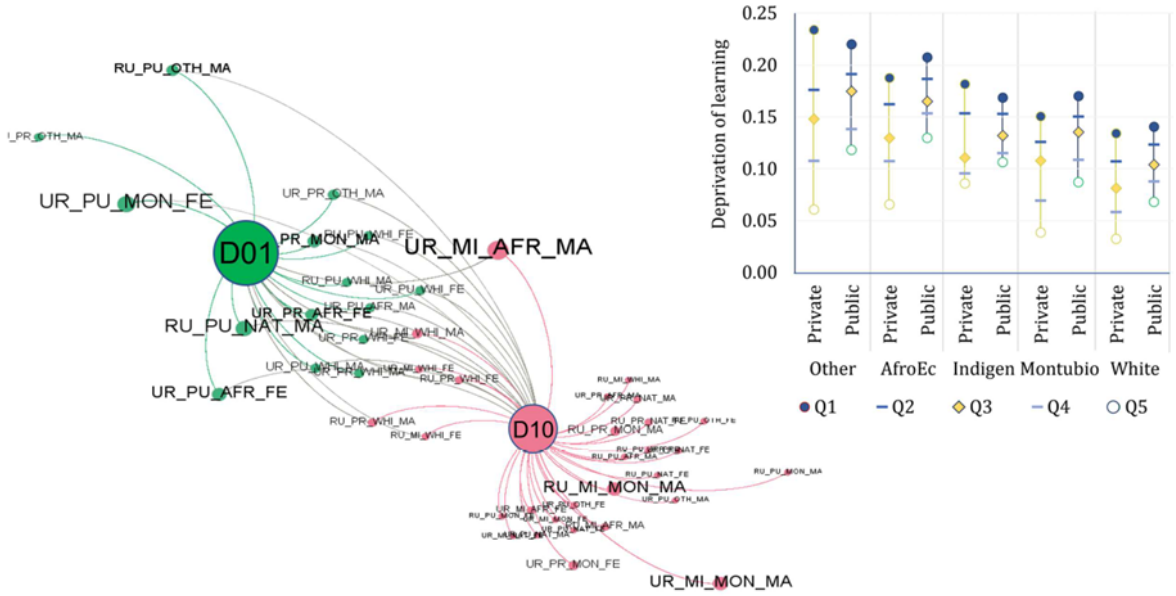


Fig. 4. Network of subpopulation groups splatted in rich and poor students with λ-weighted learning scores and its gaps synthesis.

For complementing information, the plot in the right side of Figure 4 shows that gaps between public and private sectors are quite pronounced, especially among the poorest students in both systems. In addition, although in the public sector there is less variability among quintiles, in all cases, Q1 and Q2 suffer the highest levels of deprivation. As can be seen, the lack of learning is found in the Afro-Ecuadorian population, Montubio's people, indigenous and other groups, before than people identified as White, the gaps are greater than 40% and show that a concentration in these historically vulnerable groups, mainly in the Montubios and Afro-Ecuadorians from rural areas, as well as indigenous people who attend school in urban areas.

Of special interest is the case when the modularity is represented, in this network, the parameter estimated was 0.073 and resolution -0.254, which produces two communities: the richest and the poorest. This result is particularly important for public policies because it might allow policymakers to work directly with families in a group-oriented strategy to avoid presenting same actions for completely different problems.

It is also remarkable that private schools serving indigenous and other minority groups coming from Q1, show higher deprivation rate than graduates of the public system in the same Q1 level. This result points out a tremendous social deception and suggests an urgent migration of those students to the public system in order to review the operation of these schools, given their low performance in a group with so many disadvantages. A racial dramatic case is also found: Afro-Ecuadorians with the lowest level of deprivation are those located in Q5, however, the poorest white students attending public schools show a level of deprivation equivalent.

This approach allows to measure the magnitude with which the lower deciles dominate in the interactions with the population groups through the edges [26]. Furthermore, the network has directed edges and its average weighted Out-Degree might be seen as a covariation-measure of the deciles with population of non-learning students, so, it is possible to compare the values thrown by the network with the value that could be expected on this parameter in conditions of equity where the factors would not produce differences.

5 Discussion

With this new kind of analysis, we have developed a model for finding answers to classical questions in educational research using free available data for a Latin American country, providing a direct method to recognize the structure of inequality, as well as the relationship between social determinants for educational deprivation and the conditional distribution of learning outcomes.

Given that equity is a major focus of government policies around the world and that it is promoted by international agencies with the aim of transforming educational systems, attending the wide diversity of students in each country and the whole region is a big challenge and in this paper we have presented an analysis that offers a lot of valuable evidence showing the deep lack in this dimension, highlighting that is a structural problem that goes beyond educational policy.

Using the network concept of modularity in defining groups with the same kind of challenges, might help to policymakers in selecting those factors that might be so much relevant for one group than to others, overall in some areas where the intra-class variability is very low using complementarian topological and statistical analysis.

Given the relationship between the DLI and SES, and that deprivation is almost eight times higher for poorest than for richest students, it is confirmed that the exercise of educational rights is a function of SES and that the gap is wider when considering the types of financing and that this phenomenon gets even worst for minority ethnic groups.

Extending the model for including a large set of factors should be the next step for improving the analysis and offer useful information about the educational system, as well as developing based-evidence public policies. This can be achieved introducing a dataset at micro level for building the network for meso and macro levels, a very challenging task when considering the number of students in a national education system and all the connections they may have with educational and non-educational variables. However, this point is crucial for policy because the social deception of accessing to a school without the guarantees of learning translates in levels of precariousness similar to those who never attend the school [30], creating circles of poverty that systematically impact to the impoverished people, increasing the gaps and structural inequality.

References

1. OECD (2018). Equity in Education: Breaking Down Barriers to Social Mobility, PISA, OECD Publishing, Paris
2. UN DESA (2018). The Sustainable Development Goals Report 2018, UN, New York
3. MCGRATH, Simon; NOLAN, Aoife. SDG 4 and the Child's Right to Education. NORRAG NEWS 54, 2017, p. 122.
4. Sánchez-Restrepo, H. (2018). Equity: the focal point of educational quality. National Educational Evaluation Policy Gazette in Mexico, Year 4. (10), pp. 42-44.
5. United Nations Children's Fund (Unicef). The investment case for education and equity. DC, Washington: 2015
6. Keeley, B., & Little, C. (2017). The State of the World's Children 2017: Children in a Digital World. UNICEF. 3 United Nations Plaza, New York, NY 10017.
7. Samman, E. (2018). SDG progress: fragility, crisis and leaving no one behind: report.
8. Gray, J., Kruse, S., & Tarter, C. J. (2016). Enabling school structures, collegial trust and academic emphasis: Antecedents of professional learning communities. Educational Management Administration & Leadership, 44(6), 875-891
9. Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación – LLECE (2018). Agenda 2018. http://www.unesco.org/new/en/santiago/press-room/newsletters/ newsletter-laboratory-for-assesSSent-of-the-quality-of-education-llece/.
10. Hopfenbeck, Therese N., et al. (2018). Lessons learned from PISA: A systematic review of peer-reviewed articles on the programme for international student assessment. Scandinavian Journal of Educational Research, vol. 62, no 3, p. 333-353.
11. Motl, T. C., Multon, K. D., & Zhao, F. (2018). Persistence at a tribal university: Factors associated with second year enrollment. Journal of Diversity in Higher Educ, 11(1), 51.
12. Moll, K., Göbel, S. M., Gooch, D., Landerl, K., & Snowling, M. J. (2016). Cognitive risk factors for specific learning disorder: processing speed, temporal processing, and working memory. Journal of learning disabilities, 49(3), 272-281.

13. Tracey, B., & Florian, K. (Eds.). (2016). Educational Research and Innovation Governing Education in a Complex World. OECD Publishing.
14. Fox, J. (2015). Applied regression analysis and generalized linear models. Sage Public.
15. Davis, B., & Sumara, D. (2014). Complexity and education: Inquiries into learning, teaching, and research. Routledge.
16. Barabási, Albert-László, and Márton Pásfai. Network Science. Cambridge: Cambridge University Press, 2016.
17. van der Maas, H., Kan, K., MarSSan, M., & Stevenson, C. (2017). Network models for cognitive development and intelligence. Journal of Intelligence, 5(2), 1–17. doi: 10.3390/jintelligence5020016
18. Johnson, J., Fortune, J., & Bromley, J. (2017). Systems, networks, and policy. In Non-Equilibrium Social Science and Policy (pp. 111-134). Springer, Cham.
19. Borsboom, D., & Molenaar, D. (2015). Psychometrics. In J. Wright (Ed.), International encyclopedia of the social & behavioral sciences. Vol.19 pp.418–422.Amsterdam: Elsevier
20. LI, Feiming, et al. (2009). Model selection methods for mixture dichotomous IRT models. Applied Psychological Measurement, vol. 33, no 5, p. 353-373.
21. Bracho, T (2017), Índice de déficit en competencias ¿Avanzamos hacia la garantía del derecho a la educación? en Reformas y Políticas Educativas; No. 4, septiembre-diciembre 2017, FCE, México.
22. Marsman, D. Borsboom, J. Kruis, S. Epskamp, R. van Bork, L. J. Waldorp, H. L. J. van der Maas & G. Maris (2018) An Introduction to Network Psychometrics: Relating Ising Network Models to Item Response Theory Models, Multivariate Behavioral Research, 53:1, 15-35.
23. Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mottus, R., Waldorp, L.J., & Cramer, A.O.J. (2015). State of the Art personality research: a tutorial on network analysis of personality data in R. Journal of Research in Personality, 54, 13–29.
24. Kossakowski, J.J., Epskamp, S., Kieffer, J.M., van Borkulo, C.D., Rhemtulla, M., & Borsboom, D. (2016). The application of a network approach to health-related quality of life (HRQoL): introducing a new method for assessing HRQoL in healthy adults and cancer patient. Quality of Life Research, 25, 781–92.
25. Forbush, K., Siew, C., & Vitevitch, M. (2016). Application of network analysis to identify interactive systems of eating disorder psychopathology. Psychological Medicine, 46(12), 2667–2677.
26. Gibson, L., & Koenig, A. (2012). Neighboring groups and habitat edges modulate range use in Phayre's leaf monkeys (Trachypithecus phayrei crepusculus). Behavioral Ecology and Sociobiology, 66(4), 633-643.
27. van Borkulo, C.D., Borsboom, D., Epskamp, S., Blanken, T.F., Boschloo, L., Schoevers, R.A., & Waldorp, L.J. (2014). A new method for constructing networks from binary data. Scientific Reports, 4(5918), 1–10.
28. Forbush, K., Siew, C., & Vitevitch, M. (2016). Application of network analysis to identify interactive systems of eating disorder psychopathology. Psych Med, 46(12), 2667–2677.
29. van der Maas, 2017)
30. Menefee, T., & Bray, T. M. (2015). Education in the Commonwealth: Quality education for equitable development. Commonwealth Secretariat.

CHAPTER 5

Educational Deprivation in Latin America: Structural Inequality Beyond Borders

First Autor: Harvey Spencer Sánchez Restrepo
Harvey.Restrepo@iscte-iul.pt
Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL
1649-026, Lisbon, Portugal

Coauthor: Jorge Louçã
jorge.l@iscte-iul.pt
Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL
1649-026, Lisbon, Portugal

Running head: **Educational Deprivation in Latin America**

Article published in: *Advances in Intelligent Systems and Computing, vol 1211.* **Springer Nature**.

Educational Deprivation in Latin America: Structural Inequality Beyond Borders

Harvey Sánchez-Restrepo(✉) and Jorge Louçã

Information Sciences, Technologies and Architecture Research Center (ISTAR), ISCTE-IUL, Lisbon 1649-026, Portugal
{Harvey_Restrepo, jorge.l}@iscte-iul.pt

Abstract. This is a secondary research that uses the results of a large-scale assessment to those children attending school in sixth grade, coordinated by UNESCO in 15 Latin American countries, for estimating the level of learning achievements in the scientific domain. The quality of the data allows estimating deprivation of learning in each country, offering high-level evidence on the deep lack of equity in educational systems and unveiling the structure of learning inequality through statistical and network analysis. This study offers unique evidences about the inequality gaps in learning outcomes and the incidence of educational deprivation for improving empirical knowledge about the regional interdependence between learning, cultural capital and the socioeconomic status of student's families. In addition, this contribution provides new tools to the current theoretical framework for assessing the challenges faced by the countries in the region and establishes three milestones in educational research. The first one is to provide a direct method to estimate educational deprivation. The second one is to recognize the structure of inequality through decomposition in regional socioeconomic groups with comparable measures. The third one is estimating learning gaps and their association with the socioeconomic status of students' families beyond borders.

Keywords: Network Analysis · Inequality in Latin America · Large-Scale assessments · Educational deprivation · Public policy

1 Introduction

The 193 countries attached to UNESCO promulgated that 'education is a human right' that must be guaranteed to all citizens equally [1, 2]. According to the Sustainable Development Goals (SDGs), the two pillars of quality in education are learning and equity [3–5]. However, UNICEF estimates that 250 million children who attend school do not develop the minimum level of learning and that more than half of them fail to read, write or perform basic operations [6].

Although the challenge of improving the level of childhood learning is global, in Latin America and the Caribbean (LAC) it is even greater, since it is the most unequal region in the world [5–7]. In LAC are 72 million children living in poverty, only 6 out of 10 children aged 3 to 4 receive early childhood education [7]. In addition, more than 14 million children and adolescents are outside the education system, 3.6 million of

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2020
S. Nazir et al. (Eds.): AHFE 2020, AISC 1211, pp. 92–98, 2020.
https://doi.org/10.1007/978-3-030-50896-8_14

them in the primary stage, and 7.7% are outside of secondary education, while in developed countries it is 2.1% [6].

Although huge investment efforts have been made in recent years, modest improvements in learning achievement have been followed by deep inequalities [8]. For example, children belonging to the richest families are 2.5 times more likely of attending the school that those of the poorest families [6], and the majority of students with low levels of achievement come from poor families living in precarious conditions at home, deep food deficiencies and little or no access to health and employment systems [7], which has provoked discontent, and even protests against different governments that have increased the fragility and instability of educational systems [8, 9].

In order to strengthen their public policies, most LAC countries are members of the UNESCO's Latin American Laboratory for Educational Quality Assessment (LLECE) and had participated in the Regional Comparative and Explanatory Study (ERCE), a large-scale evaluation (LSA) to collect valid and reliable data on educational outcomes and a large number of Factors associated with learning (FAL) [10]. The main objective of this program is to create evidence and detect those variables driving or conditioning educational results for developing policies, programs and strategies focused on closing gaps between poor and rich students in all countries and across the region.

2 Materials and Methods

Data Sources. The TERCE database was produced with high quality standards in all its stages of sample design, data collection, processing and statistical estimates, so they are useful and sufficiently robust to estimate and compare educational deprivation in those 15 LAC countries, as well as its relationship with the student's Socioeconomic Status (SES) and other relevant social determinants. The dataset is multivariate and has a relational structure that links the learning outcomes of the students —represented by a score and grouped into four levels of achievement— with a unique ID for the dimensions of the explicit and implicit strata, FALs and the SES segmented by the cut-off points for the regional deciles $\{Dk\}k = 1, \ldots, 10$, which might be join to the estimated ability θ^j and the plausible values $\{PV_k^j(\theta^j)\}k = 1, \ldots, 5$ under the matrix of the LSA.

For estimating the deprivation of learning, model developed by Bracho [10] was used for creating a new representation of the set $\{\theta^j\}$. From the percentage of students in each stratum, the vector $\overrightarrow{\omega_{Di,k}}$ was estimated, where i refers to the i-th decile and k to the k-th country. In this way, the selected the subset L_0 of those students below the first level of achievement as students deprived of learning —who did not reach the minimum expected learning.

From the above, the *rate of learning deprivation* is given by $H(C_k^j) = n(C_k^j, L_0)/\sum_j n(C_k^j, L_j)$, where $n(C_k^j, L_j)$ represents the cardinality of the group C_k^j at the achievement level L_j. The *Intensity of learning deprivation* $\lambda_j(C_k^j, \theta^j)$ is estimated by distance to reach the first level of achievement L_{II}. Thus, the *Deprivation of Learning Index* (DLI) is given by ${}_j(C_k^j, \theta^j) = H_k^j \cdot \lambda_j(C_k^j, \theta^j)$ [11]. Statistical analysis was made with Julia 1.3.0 and Orange 3.3.8, while Network Analysis was implemented in Gephi 0.9.2.

3 Structural Inequality in Learning Deprivation

According to Royce [12], structural inequality is 'a situation in which a group of people recognizes a state of inequality in relation to other groups', which is perpetuated and reinforced by a confluence of unequal relationships in roles, functions, decisions, rights and opportunities [13]. In Latin America, barriers to schooling and the unequal offer of educational services to groups of population according to the postcode constitute the first sources of structural inequality determined by SES, which is expressed with different magnitudes in each country [14].

In poor areas, the lack of education enrolment is often accompanied by low quality services that prevent students from developing minimum learning. In this way, children who, despite their marginalization conditions, gain access to school in the early years, begin a ladder in which the school system helps very little to reverse the disadvantages in opportunities for prosperity [15], limiting the richness of the school experience to a certification process with insignificant learning, a very unlikely scenario for improving welfare and having access a better life, as shown in Fig. 1.

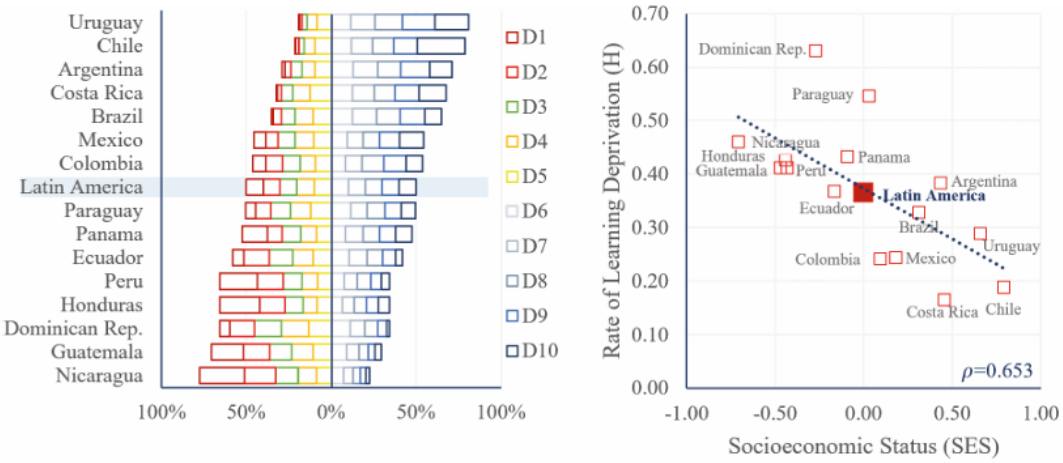


Fig. 1. The bar chart shows the proportion of students in each socioeconomic decile with regional cut-off point. The scatter plot shows the relationship between *SES* and *H* at country level.

As can be seen, 36 out of every 100 children in LAC do not develop minimum scientific learning at the end of basic education. Despite this high rate, the variability and the gap between countries is huge, for example, in Dominican Republic $H_{DOM} = 0.631$ and in Paraguay $H_{PAR} = 0.546$ while in Chile and Costa Rica it is $H_{CHI} = 0.189$ and $H_{CRI} = 0.166$, respectively. In addition, the correlation between *SES* and *H* es $\rho_{SES,H} = 0.653$ ($p < 0.001$), which indicates a degree of relevant association between both dimensions of precariousness. So, countries with lower SES are also those that exhibit higher rates of deprivation, with the exception of Panama, which despite having a socioeconomic level close to the Latin American average and higher than Ecuador or Peru, exhibits similar deprivation rates to Honduras and Guatemala.

To understand the socioeconomic structure of the distribution of L_0, a directed weighted network was designed $\{C_K^j\} \rightarrow \{Dk\}$ $\forall j$-students as edges, k-countries as source-nodes and deciles as sink-nodes. The weight of each edge ε_j^i is the normalized value of $\lambda(\cdot)$, so this arrangement allows you to find class properties and find four groups of Learning Deprivation: *Generalized*, *recurring, Incipient* and *Controlled*. After iteration process, the *In-degree* parameters of the nodes give the estimates for the level of connectiveness of each node $\{H_{Dk}\}k = 1, \ldots, 10$, that might be understood as how far from the equilibrium is each SES group, as shown in Fig. 2.

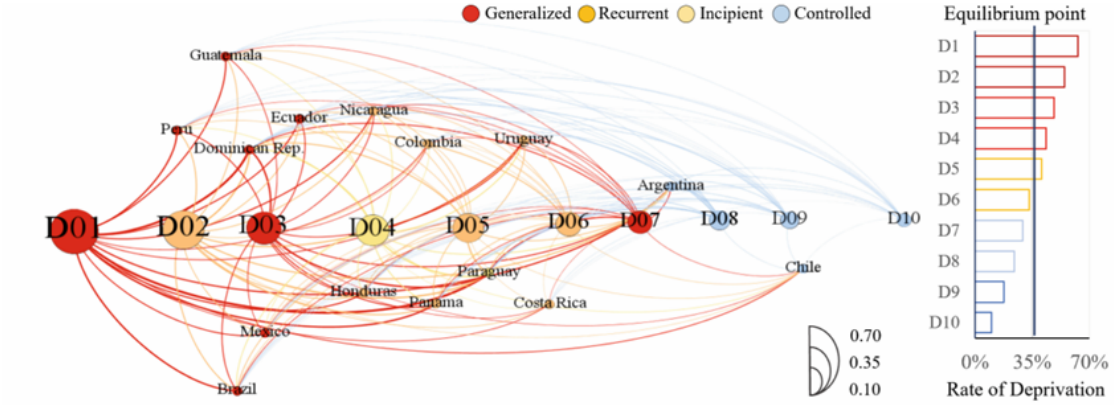


Fig. 2. The network shows the level of deprivation of learning in each country based on the distribution and degree of connectivity of its student's population in L_0, splatted by deciles.

As can be seen, the H function decreases monotonously as the SES increases. While 63 of each 100 poor students do not learn the minimum, in rich families the rate is 10 per 100, which indicates the gap between D01 and D10 ($H_{\mathrm{D01}} = 0.631$ and $H_{\mathrm{D10}} = 0.099$). Despite the dramatic difference between the deciles, the gap deepens much more between the countries in the region. For example, among the richest students in Mexico (D_{MEX}^{10}), 5 out of 100 students are in L_0, while in the group of the poorest in the Dominican Republic (D_{DOM}^{01}), 86 of 100 are located in the same group. Furthermore, the level of deprivation in both groups is very different: $\lambda\left(C_{MEX}^{D10}\right) = 0.08$ and $\lambda\left(C_{DOM}^{D01}\right) = 0.22$, which means that the effort of the first group to reach L_{II} would be only a third of the second for guarantying the right to education to the whole group of students, given that the Mexican group is much closer to the minimum threshold than the Dominican one, as can be seen in Fig. 3.

As mentioned before, not only family SES determines the levels of learning, but also the social and cultural context of the countries has a significant impact on learning opportunities. Figure 3 shows that deprivation has two sources of variability: 1) intragroup—within the deciles—with a gap of 46.6% in decile 1 that decreases progressively to 22.4% and 2) intergroup, which presents a gap of 46.5% between countries.

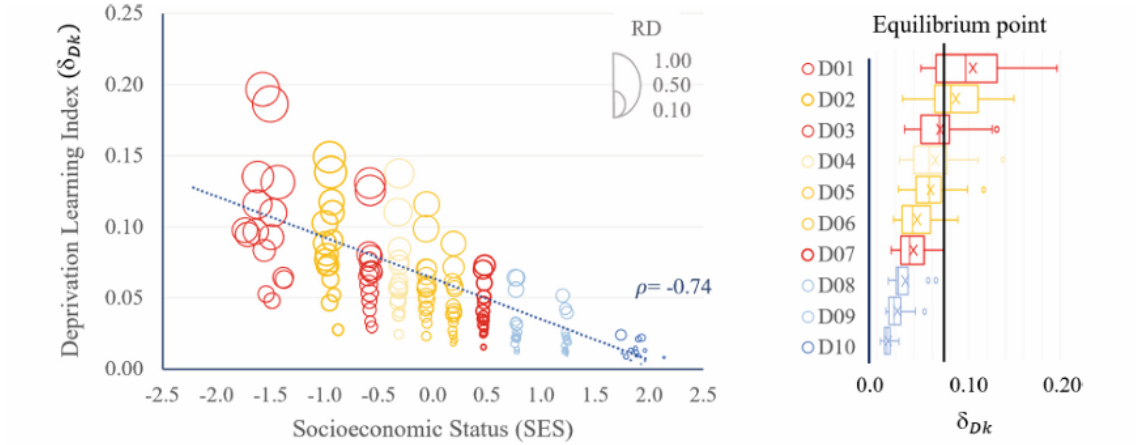


Fig. 3. The scatter plot shows DLI estimates for the deciles of all countries. The Box-and-whisker Plot shows intra-group and inter-group variance showing the distances to the equilibrium point.

As can be seen, the SES gap goes beyond geographical barriers: students from middle-class families in the poorest countries develop equivalent learning to students with the lowest SES in more privileged countries, which can be interpreted as a positive externality of living in a higher-level economy. For example, the incidence of deprivation in students in deciles D06, D07 and D08 of the Dominican Republic, Paraguay and Argentina corresponds to that of students in decile D01 of Chile, Uruguay and Costa Rica.

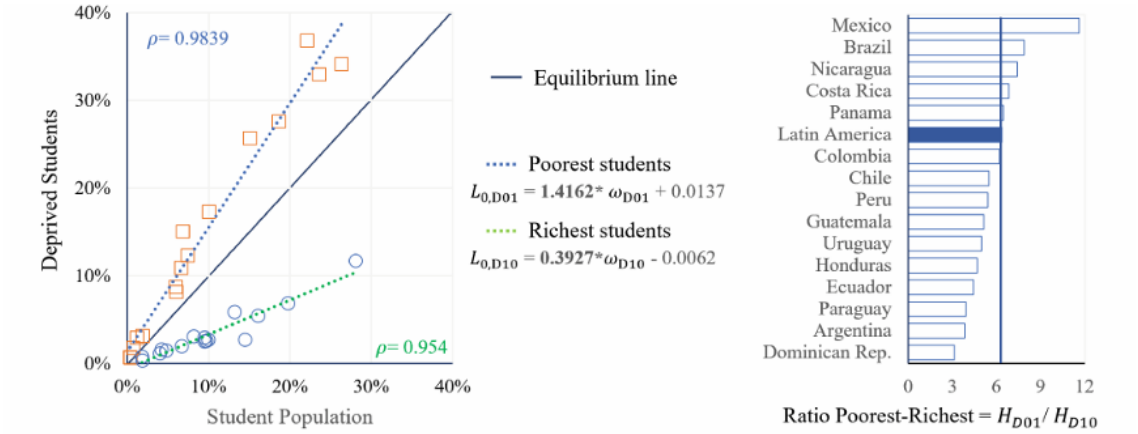


Fig. 4. The scatter plot shows the relationship between ω_{Dk} and H function for D01 (squares) and D10 (circles) for each country. The bars on the right indicate the ratio DR_{D01}/DR_{D10}.

As can be seen in Fig. 4, the regression line of students in D01 shows a homogeneous increase on rates in all countries $\rho_{SES,H} = 0.9839$ ($p < 0.001$). The slope of 1.4162 implies an average increment the region of 41.62% for poor students in the group $L_{0,D01}$ with respect to their initial weight ω_{D01}. On the other hand, the D10 regression line is also quite homogeneous $\rho_{SES,H} = 0.954$ ($p < 0.001$), but its slope of

0.3927 implies that the percentage of rich students with deprivation is reduced above 60% of the initial weight. Obviously, this lack of capacity for providing to all students with the same learning varies widely among countries, but in all countries, variations are far away from the equity situation represented by the equilibrium line, as well as the intensity of inequality between rich and poor students. In this sense, Mexico shows the highest gap: the total percentage of students in deciles 1 and 10 is $\omega_{\mathrm{D01}} = 6.9\%$ and $\omega_{\mathrm{D10}} = 14.5\%$, but there is an incidence of $H_{D01} = 0.535$ and $H_{D10} = 0.046$, which gives 11.6 poor students for each rich in L_0 and a deprivation ratio of $_{r(D01/D10)} = {}_{D01}/_{D10} = 23.5$. On the contrary, Ecuador has $\omega_{\mathrm{D01}} = 6.7\%$, $\omega_{\mathrm{D10}} = 4.3\%$, $H_{D01} = 0.109$ and $H_{D10} = 0.016$, with a ratio $H_{D01}/H_{D10} = 4.4$ and $_{r(D01/D10)} = 10.6$, a still dramatic learning deprivation gap, but only half of the deep challenge that Mexico must face.

4 Discussion

Latin American countries share many aspects of their economic, social and cultural reality, which is also reflected in the deep inequality of opportunities offered by their educational systems and the low learning outcomes obtained by their students, mainly those from poorest families. Although this problem is common to all the countries in the region, there are observed significant differences related with the socio-demographic composition of their school population, a phenomenon that is particularly reflected in the distribution of the students in their educational systems [16].

Historically, most measures of learning achievement have been used to contrast scores as the final goal of the system, rather than reaching the minimum level to guarantee the right to education. In this research, clear and conclusive evidences are offered on the deep lack of equity between different socioeconomic classes, a phenomenon present in all countries [13]. This point is crucial because the social deception of accessing education without learning guarantees translates later into a convergence in precarious levels, similar to those children outside of the school system, which creates circles of poverty increasing structural inequality [16].

This research also provides evidence on the deep lack of equity in LAC countries and their relationship with social determinants for synthesizing the complex structure of inequality in learning outcomes as a consequence of socio-economic and cultural inequality, estimating the capacity of educational systems to bridge the gaps [10]. So, with this analysis, three milestones in educational research are established using the best data available for 15 Latin American countries. The first is to provide a direct method to estimate educational deprivation. The second is to recognize the structure of inequality through the decomposition into regional socio-economic groups with comparable measures. The latter focuses on estimating learning gaps and their association with the socioeconomic status of students' families beyond borders.

This study contributes significantly to integrating thinking tools into current theoretical frameworks and finding non-trivial relationships and to estimate the magnitude of the correlation between social conditions and educational deprivation in historically

marginalized population groups. Given the multiple evidences on the adverse effects derived from the lack of quality of school processes in LAC, equity should be at the core of educational policies. In this sense, the information shown might be helpful to policymakers for creating better strategies while developing inclusive and equitable education for all.

References

1. Desa, U.N.: The Sustainable Development Goals Report 2018. UN, New York (2018)
2. UN (2015), Incheon Declaration. Education 2030: Towards inclusive and equitable quality education and lifelong learning for all. In: World Education Forum, pp. 19–22
3. OECD: Equity in Education: Breaking Down Barriers to Social Mobility. PISA, OECD Publishing, Paris (2018)
4. UNICEF, et al.: The investment case for education and equity. Unicef (2015)
5. Sánchez-Restrepo, H.: Equity: the focal point of educational quality. National Educational Evaluation Policy Gazette in Mexico, Year 4, no. 10, pp. 42–44 (2018)
6. Keeley, B., Little, C.: The State of the World's Children 2017: Children in a Digital World. UNICEF. 3 United Nations Plaza, New York, NY 10017 (2017)
7. Salazar, M.C., Glasinovich, W.A. (ed.) Child work and education: Five case studies from Latin America. Routledge (2018)
8. Samman, E.: SDG progress: fragility, crisis and leaving no one behind report (2018)
9. Tracey, B., Florian, K. (eds.): Educational Research and Innovation Governing Education in a Complex World. OECD Publishing, Paris (2016)
10. Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación – LLECE (2018). Agenda 2018. http://www.unesco.org/new/en/santiago/press-room/newsletters/newsletter-laboratory-for-assesSSent-of-the-quality-of-education-llece/
11. Bracho, T,: Índice de déficit en competencias "Avanzamos hacia la garantía del derecho a la educación" en Reformas y Políticas Educativas; No. 4, septiembre-diciembre 2017, FCE, México (2017)
12. Sánchez-Restrepo, H., Louçã, J.: Topological properties of inequality and deprivation in an educational system: unveiling the key-drivers through complex network analysis. In: International Conference on Human Systems Engineering and Design: Future Trends and Applications, pp. 469–475. Springer, Cham, September 2019
13. Royce, E.C.: Poverty and Power: The Problem of Structural Inequality (2009)
14. Dani, A.A., Arjan De, H., (eds.): Inclusive States: Social Policy and Structural Inequalities. World Bank, Washington (2008). 3
15. Willms, J.D.: Educational Prosperity: An assessment strategy for supporting student learning in low-income countries. Learning at the Bottom of the Pyramid: Science, Measurement and Policy in Low-Income Countries, UNESCO-IIEP, Paris (2018)
16. Lanham: Rowman & Littlefield: Combating Poverty and Inequality: Structural Change, Social Policy. Report no. 978–92-9085-076-2. Geneva: United Nations Research Institute for Social Development. Web. 3 Nov. 2011 (2010)

CHAPTER 6

Educational Gaps in Latin American childhood: Regional Inequality in Learning Deprivation with Local Modulations

First Autor: Harvey Spencer Sánchez Restrepo
Harvey.Restrepo@iscte-iul.pt
Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL
1649-026, Lisbon, Portugal

Coauthor: Jorge Louçã
jorge.l@iscte-iul.pt
Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL
1649-026, Lisbon, Portugal

Running head: **Regional Inequality in Learning Deprivation with Local Modulations**

Article published in: *Journal of Physics: Conference Series 1828.* **Journal of Physics**.

Educational Gaps in Latin American childhood: Regional Inequality in Learning Deprivation with Local Modulations

Harvey Sánchez-Restrepo* **and Jorge Louçã**

Information Sciences, Technologies and Architecture Research Center (ISTAR), ISCTE-IUL, Lisbon, 1649-026, Portugal

*Email: Harvey_Restrepo@iscte-iul.pt

Abstract. This research is a secondary analysis of a multivariate dataset coming from a large-scale assessment, coordinated by UNESCO in 15 Latin American countries (LAC), which was carried out for estimating the level of learning achievements of students at the end of primary education. The objective is to estimate the lack of quality in education across the region, based on the sociological concepts of absolute and relative deprivation, through a new learning deprivation index, as well as to offer robust evidence on the deep inequalities in learning outcomes among groups of population. Main results show strong evidences on how big gaps in learning outcomes are associated with socioeconomic status of students, pointing out that the profound lack of equity in LAC has become structural inequality beyond borders. This study establishes three milestones in educational research using the best available data. First, provides a direct method for estimating learning deprivation. Second, intra-and inter-country inequality is estimated from comparable socioeconomic deciles at regional level. Third, it was found that the intensity of learning deprivation, as a function of the SES, has an interplay with the wealth of the country and it is not possible to apply same strategies in all countries for developing inclusive and equitable educational for all.

1. Introduction

The unequal exercise of the right to education is a source of social and economic inequality. In the scenario prior to the current COVID pandemic, it was estimated that, worldwide, the primary total net enrollment rate was 92% and the primary completion rate was only 84%, with Sub-Saharan Africa being the most depressed region, with rates of 81% and 63%, respectively [1]. Furthermore, in the same period, 258 million children and youth were out of school [2], of which 59 million were of primary age, 62 million in lower secondary age, and 138 million of upper secondary age. To deal with this critical situation, the United Nations (UN) launched in 2015 the Sustainable Development Goal 4 (SDG-4) as part of the 2030 Agenda, from which 193 countries committed to 'Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all', which in practical and conceptual terms translates into focusing efforts on two pillars: learning and equity [3-5].

In Latin America and the Caribbean (LAC), the challenges to comply with the SDG-4 are enormous, since the primary total net enrollment rate is 95% and the primary completion rate is only 91% [6]. Additionally, in LAC are 72 million children living in poverty, only 6 of each 10 children between 3 and 4 years old receive early education, and more than 14 million children and adolescents are outside the educational system, 3.6 million of them in the primary stage and 7.7% outside of secondary education, while in developed countries this percentage is 2.1% [7]. This situation has multiple consequences in inequality, for example, children belonging to the richest families are 2.5

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

times more likely to attend school than those of the poorest families [6]. With these antecedents, LAC countries face new threats of deepening inequality in post-COVID era due mainly to the differential effect of the pandemic in absolute and relative learning deprivation, as well as the magnitude of the gaps between students coming from rich and poor families before and after the pandemic.

2. Material and Methods

With the aim of strengthening their educational policies, most of the countries in LAC are members of UNESCO's Latin American Laboratory for the Evaluation of Educational Quality (LLECE), 15 of them participated in the Third Regional Comparative and Explanatory Study —TERCE, for its acronym in Spanish—, a large-scale assessment (LSA) designed through a sample design that assesses students to collect valid and reliable data on educational outcomes, along with a large number of factors associated with learning (FAL) [8]. The main objective of this project is to generate evidence and detect those variables that drive or limit learning outcomes of students in Numeracy, Literacy and Science for developing evidence-based policies that might help to close the gaps between students of all socioeconomic strata and among the countries of the region [8].

This research carries out a secondary analysis of this multivariate data set developed by UNESCO from TERCE's results, a database produced with high quality standards in all its stages of sample design, data collection, processing and statistical estimates [9]. Thus, data is useful and sufficiently robust to estimate and compare learning outcomes and educational deprivation in those 15 LAC participating countries, as well as its relationship with the student's Socioeconomic Status (SES) and other relevant social determinants [10].

For estimating the raw scores, psychometric parameters were estimated by Item Response Theory (IRT) through a 1P-Logistic model, specifically a Rasch model [9], following equation (1):

$$P(Y_{i,j}|\theta_i,\beta_j) = \frac{e^{(\theta^i-\beta_j)}}{1+e^{(\theta^i-\beta_j)}} \quad (1)$$

where $Y_{i,j}$ corresponds to the response of student i to item j; $\theta^i \in (-\infty,\infty)$ refers to the ability of the student i and $\beta_j \in (-\infty,\infty)$ represents the difficulty of item j [11]. As raw scores of θ^i are in logits, mean and standard deviation are centered in $\mu = 0$ and $\sigma = 1$, however, final scores are linear transformed scores with $\mu = 500$ and $\sigma = 100$ for stablishing a *learning index* $\ell^i = T(\theta^i)\ \forall i$, where student with higher levels of learning are more likely to have higher scores [11]. As the dataset has a relational structure, it is easy to join learning outcomes of the students with a unique ID for the dimensions of the explicit and implicit strata, FAL and the SES. Thus, SES have been properly segmented by regional cut-off points for producing regional deciles $\{\mathcal{D}_j\}\ j = 1,\ldots,10$.

3. Structural Inequality in Learning Deprivation

According to Royce [12], structural inequality is 'a situation in which a group of people recognizes a state of inequality in relation to other groups', a state that might be perpetuated and reinforced by a confluence of unequal relationships in roles, functions, decisions, rights and opportunities. This idea helps to explain the big barriers to schooling and the unequal offer of educational services to groups of population according to the postcode, one of the main sources of structural inequality in LAC [12].

For measuring the level of learning deprivation, is necessary to measure it in two ways: *absolute* and *relative*. The first concept proposes that 'there is an irreducible nucleus of needs that are common to every human being' [13]. For attending this point, LSA's results are classified for all students in m levels of achievement L_m, $m = 0,1,2,3$, a standard procedure based on a *Bookmark process* [9] developed from psychometrical and qualitative analysis for establishing the m cut off points ψ_m in the same scale of ℓ^i, where $\psi_0 = \alpha = \min(\ell_i)\ \forall i$, so the following equality holds

$$n(\varepsilon_i,\mathcal{D}_j) = \sum_{k=1}^{p} n_k(\varepsilon_i,\mathcal{D}_j) = \sum_{k=1}^{p} n(C_k,\varepsilon_i,\mathcal{D}_j) = \sum_{k=1}^{p}\sum_{r=1}^{m} n_k(\varepsilon_i,\mathcal{D}_j;L_r) \quad (2)$$

In equation (2), the number $n(\varepsilon_i, \mathcal{D}_j)$ refers to valid students assessed at the region in stratum ε_i coming from the decile j, $n_k(\varepsilon_i, \mathcal{D}_j) = n(C_k, \varepsilon_i, \mathcal{D}_j)$, a number that identifies a group bounded by $(C_k, \varepsilon_i, \mathcal{D}_j)$ inside the country k; p is the total number of participating countries and L_m refers to one of the m levels of achievement reached by each group. Students suffering learning deprivation are those that did not meet the minimum standards according with the regional curriculum, grouped by L_0. Thus, the *rate of learning deprivation* in the country k for the j-th decile is given by equation (3).

$$H_k(\varepsilon_i, \mathcal{D}_j) = n_k(\varepsilon_i, \mathcal{D}_j; L_0)\left[\sum_{r=1}^{m} n_k(\varepsilon_i, \mathcal{D}_j; L_r)\right]^{-1} \quad (3)$$

Figure 1 shows the distribution of the student population among the four achievement levels, disaggregated by socioeconomic deciles estimated from the SES. As can be seen, in the region, 37% of the students do not develop the minimum learning in the scientific field, 39% are located at level L_1 — elementary learning—, 17% at L_2 —the expected of an appropriate and timely educational process— and 7.4% reach level L_3, which exceeds standards provided by the regional curriculum [9]. However, the same Figure 1 shows how the relative participation of the private learning group varies according to the SES, being 10% for the $\mathcal{D}_{10}$ —those from families with greater socioeconomic advantages— and 63% for $\mathcal{D}_{01}$ —the group with the worst situation—.

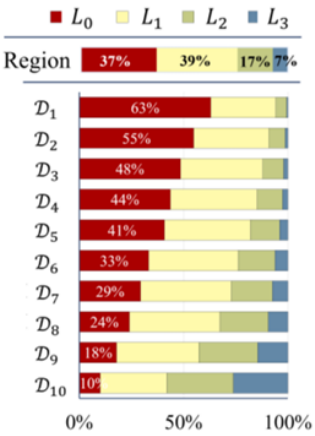


Figure 1. Distribution of levels of achievement among SES deciles.

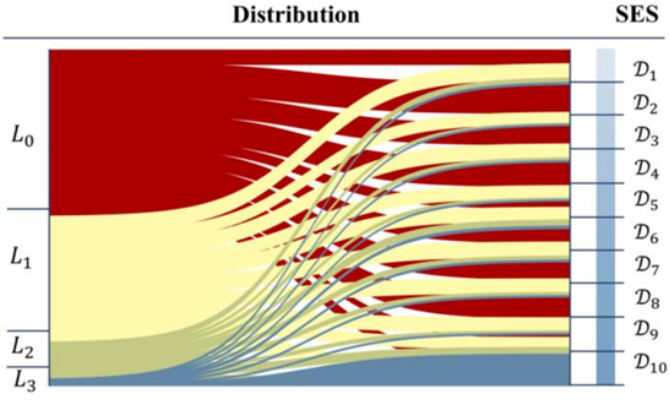


Figure 2. Relative flows of the distribution of levels of achievement among SES deciles.

Figure 2 shows the synthesis of absolute deprivation for each decile of the *SES*, which highlights the profound inequality in the distribution of learning levels among socioeconomic groups, which explains why the *H* function decreases monotonously as the *SES* increases. For example, 63 of each 100 poor students do not learn the minimum (H_{D01}=0.631), while in rich families the rate H_{D10}=0.099, this means that, of every 100 students who do not develop elementary learning, 32 correspond to the first two deciles $(\mathcal{D}_1 + \mathcal{D}_2)$, while only 8 belong to the two highest deciles $(\mathcal{D}_9 + \mathcal{D}_{10})$. On the contrary, of every 100 students in the outstanding group L_3, only 3 come from the deciles $\mathcal{D}_1$ and $\mathcal{D}_2$, while 55 come from the groups $\mathcal{D}_9$ and $\mathcal{D}_{10}$, 20 and 35, respectively, indicating the big gap between students coming from poor and rich families.

Now, regarding the estimation of relative deprivation, the sociological proposal is that 'needs, thresholds and satisfactions are determined by each society', a sentence that might be interpreted in this context as 'how deep is deprivation', which can be estimated through the function $\lambda^i(C_k, \varepsilon_i, \mathcal{D}_j, \mathrm{T}(\theta^i)) \in [0,1]$, a real value measuring the intensity of learning deprivation in the i-th student, measured as the distance from the score reached by her or him to the first cut-off point ψ_1, considering the trajectory given by the function $S_{L_0}(\alpha, \vec{\beta})$: $\mathbb{R} \to [0,1]$, as equation (4) shows

$$\lambda^i\left(C_k, \varepsilon_i, \mathcal{D}_j, \mathrm{T}(\theta^i)\right) = \int_{\rho^i}^{\psi_1} dS_{L_0}\left(\alpha, \vec{\beta}; s\right) \tag{4}$$

From these two measurements, *Index of Learning Deprivation* (ILD) results as a composition of absolute and relative deprivation [13], which is equivalent to adding the intensity functions of all students in L_0 and weighting the result for all students grouped by $(\varepsilon_i, \mathcal{D}_j)$, as equation (5) indicates

$$n_k\left(\varepsilon_i, \mathcal{D}_j; L_0\right)\left[\sum_{r=1}^{m} n_k\left(\varepsilon_i, \mathcal{D}_j; L_r\right)\right]^{-1} \cdot \frac{1}{L_0}\left[\sum_{i=1}^{L_0} \lambda^i\left(C_k, \varepsilon_i, \mathcal{D}_j, \mathrm{T}(\theta^i)\right)\right] \tag{5}$$

From the above, the ILD is given by equation (6)

$$\delta_k\left(C_k, \varepsilon_i, \mathcal{D}_j, \mathrm{T}(\theta^i)\right) = H_k\left(\varepsilon_i, \mathcal{D}_j\right) \cdot \Lambda\left(C_k, \varepsilon_i, \mathcal{D}_j, \mathrm{T}(\theta^i)\right) \tag{6}$$

4. Deprivation in Socioeconomic Classes Beyond Borders

As the ILD is a standardized and bounded metric, its estimates are useful to compare different groups and countries in a stable way [14], which allows, among other things, to assess the size of the gaps over time and territory. The horizontal axis of Figure 3 shows the rate of learning deprivation (H_k), the vertical axis represents the intensity of learning deprivation (Λ_k), both at country level. As can be seen, there is a fairly significant positive correlation between the metrics ($\rho_{H,\Lambda}$=0.90, p<0.001), i.e., those countries with the highest percentage of students with learning deprivation also tend to exhibit more intense deprivation: in those countries with greater challenges, such as the Dominican Republic, Paraguay, Panama, Peru, Honduras and Nicaragua, not only have more students lacking the most basic education, but the deficiency is broader and deeper, so reducing the gaps with the mixed effect may require greater efforts than those made by other countries in better financial conditions.

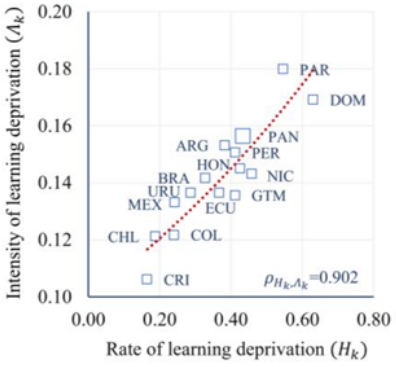


Figure 3. Relationship between absolute and relative deprivation.

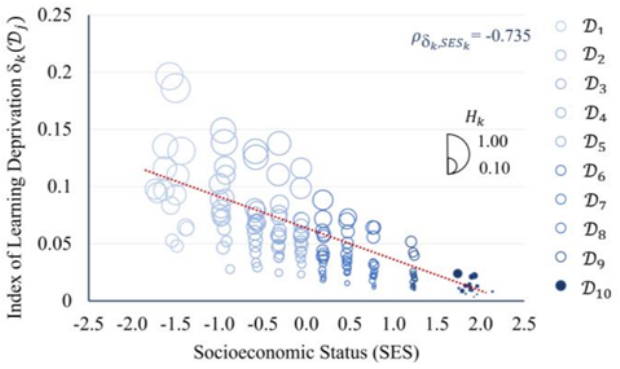


Figure 4. ILD gaps among countries grouped by SES deciles.

Figure 4 presents the SES on the horizontal axis and the ILD δ_k on the vertical axis to show the wide disparity in learning deprivation within the region. First, it points out an enormous heterogeneity among the lowest deciles, highlighting that the poorest students in Chile, Colombia, Costa Rica and Uruguay got a $\delta_k(\mathcal{D}_{01}) < 0.065$, while Dominican Republic and Paraguay tripled that value ($\delta_k(\mathcal{D}_{01}) > 0.185$). Second, it indicates that, regardless of the country, as the SES increases, not only does the IDL decrease, but the distance between countries tends to zero, which implies that countries with less resources serve the poorest students so much worst. Third, the low variability in the upper deciles suggests that, regardless the country, students belonging to $\mathcal{D}_{10}$ get a minimum level of learning deprivation and a very low probability of suffering from it.

Despite the dramatic difference between the deciles, the gap deepens much more between the countries in the region. For example, among the richest students in Mexico, 5 out of 100 students are in L_0, while in the Dominican Republic, 27 of 100 students form the same socioeconomic group are located in the same level of achievement. Despite the dispersion of deprivation levels among the

lowest deciles, a close negative relationship between *SES* and *ILD* (ρ_{δ_k, SES_k}= -0.735, p<0.001) allows to estimate the sizes of the intra-and inter-country gaps, as well as to analyze its intensity, unveiling a structural deprivation beyond borders.

For recognizing the deprivation structure of each country, it might be useful to build a directed weighted network integrating only the students whose scores were located in L_0 and relate the nodes that represent the countries with the SES deciles through multiple edges. In this network, each student is an edge such that $\{C_K^j\} \rightarrow \{\mathcal{D}_j\}$, where the k-countries are source-nodes and deciles are sink-nodes [15]. By its continuous nature, the normalized value of $\lambda(\cdot)$ is the weight of each edge E_j^i. Figure 5 shows the network of this model where it is possible to find class properties and identify degrees of learning deprivation: from *Generalized* to *Controlled*. After iteration process, the In-degree parameters of nodes estimate the level of connectiveness of each node $\{H_k(\mathcal{D}_j)\}$ $j = 1, \ldots, 10$, represented by the size of the node and pointing out how far from the regional average is each SES-decile.

As can be seen, Honduras, Dominican Republic, Nicaragua, Guatemala and Peru concentrate in low deciles through strong connections, showing generalized deprivation in poorest students, while Panama, Colombia, Mexico and Brazil deprived students are distributed along poor and middle SES; Uruguay, Costa Rica and Chile diversify deprivation in a better way, being Argentina and Chile the more balanced countries, having a better control over SES impact. Furthermore, countries with lower SES also exhibit higher rates of deprivation, with the exception of Panama, which despite having a higher SES level than Ecuador or Peru, exhibits similar deprivation rates as Honduras and Guatemala.

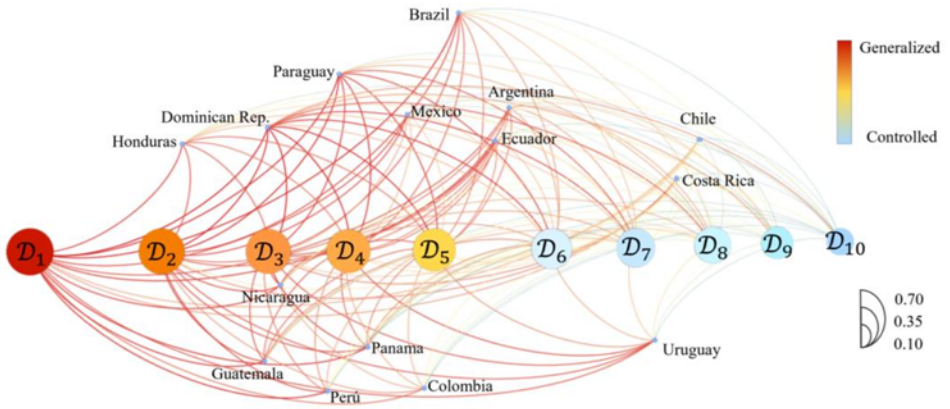


Figure 5. Network representation of the level of learning deprivation in each country based on the distribution and degree of connectivity of students splatted by deciles and $\delta_k(C_k, \mathcal{D}_j)$.

5. How Big Are the Learning Gaps?

One way to assess the learning gaps between socioeconomic groups is from the speed with which this deprivation changes as the SES varies. In this way, the regression coefficient between deprivation levels and SES is a measure of speed and could indicate the intensity of the gap with respect to socioeconomic status. As has been seen, the countries that present higher absolute levels of deprivation also exhibit higher relative levels of deprivation, however, the intensity of the gaps in both dimensions is a measure that offers novel information on the learning distance that separates students by SES, allowing to estimate the expected effects when applying variational methods to policies.

In this sense, Figure 6 shows a strong exponential relationship (ρ=0.902, p<0.001) between the intensity of the absolute and relative deprivation gaps when estimating the regression equation over SES. The results indicate that Chile, Costa Rica and Colombia, exhibits shorter learning distances among their students, when compared to the Dominican Republic and Paraguay, countries that are much less economically capable than Chile, but in which students from the richest families learn the same. In this sense, it is interesting how the intensity of the gap in Panama is one of the largest in the

region—similar to that of Paraguay—, but it is located in the LAC average in the intensity of deprivation.

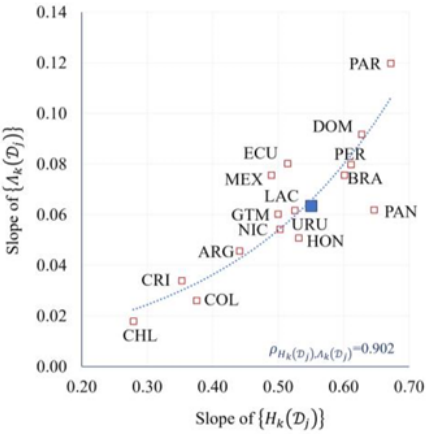


Figure 6. Relationship between intensities of absolute and relative deprivation gaps.

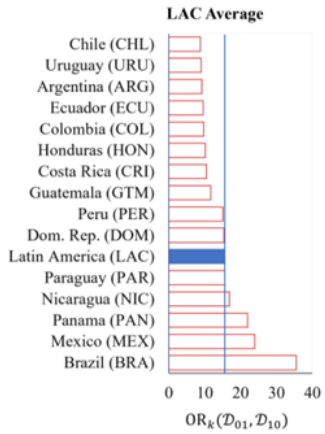


Figure 7. Magnitude of the gaps in each country through odds-ratio $\mathcal{D}_{01}$-$\mathcal{D}_{10}$

To clarify the magnitude of the gaps in each country, the odds-ratio between two deciles $\mathcal{D}_i$ and $\mathcal{D}_j$ in the stratum ε_i can be estimated for country k, i.e., the quotient between odds of $\mathcal{D}_{01}$ and $\mathcal{D}_{10}$, denoted by $\mathrm{OR}_k(\varepsilon_i; \mathcal{D}_i, \mathcal{D}_j)$, which indexed the size of the gap as a function of how many students from the lowest decile are expected to have in L_0, for each student from the highest decile. Figure 7 shows the calculation for all countries, the OR in the region is 15 poorest students for each one from the richest families, Chile, Uruguay and Argentina show the smallest gaps, followed by Ecuador and Colombia, being Panama, Mexico, and Brazil those countries with the largest OR highlighting the worst situation in LAC.

6. Discussion

Analyzing the average scores of the countries for estimating the gaps between them, leaves aside the rich qualitative information derived from achievement levels. For attending both key aspects, this research offers a new metric based on the concepts of absolute and relative deprivation to generate robust evidence on the profound lack of equity between different socioeconomic groups, a phenomenon rooted in all LAC countries [16]. Shown results are absolutely relevant because are more and better associated with the SDG-4 and the right to education. This study establishes three milestones in educational research using the best available data. The first one is to provide a direct method for estimating educational deprivation. The second one is to estimate structural inequality by decomposing into regional socioeconomic deciles with comparable measures. The third one is to provide an estimation of the intensity of learning deprivation as a function of SES, producing useful information for helping policymakers to develop an inclusive and equitable education system for all.

7. Conclusion

Most of students with low levels of performance come from poor families living in precarious conditions at home, profound nutritional deficiencies, and little or no access to health systems and employment [17,18]. So, tracking the lack of equality in the students' right to learn is crucial because it will be reflected later in multiple inequalities, deepening the systematization of precariousness, generating circles of poverty, reducing social mobility and promoting a stronger structural inequality [19], extinguishing the few possibilities of millions of children of developing skills for getting higher opportunities to a better life [20].

References

[1] UNESCO Institute for Statistics 2019 Combining Data on Out-of-school Children, Completion and Learning to Offer a More Comprehensive View on SDG 4. Information Paper No. 61 October 2019. Montreal: Canada.
[2] UN DESA 2018 *The Sustainable Development Goals Report 2018* (UN, New York)
[3] UN 2015 Incheon Declaration. Education 2030: Towards inclusive and equitable quality education and lifelong learning for all *In World Education Forum* p 19-22
[4] Sánchez-Restrepo H 2018 Equity: the focal point of educational quality *National Educational Evaluation Policy Gazette in Mexico* **10** 42-44
[5] UIS 2019 Education Data Release: Sep. 2019. Information Paper No. 59. UIS/2019/ED/IP/59
[6] UNICEF et al 2015 The investment case for education and equity. *Unicef*, 2015
[7] Keeley B and Little C 2017 The State of the World's Children 2017: Children in a Digital World. UNICEF. 3 (United Nations Plaza, New York)
[8] UNESCO-OREALC 2018 Agenda 2018. http://www.unesco.org/new/en/santiago/press-room/newsletters/ newsletter-laboratory-for-assesSSent-of-the-quality-of-education-llece/.
[9] UNESCO-OREALC 2016 Reporte Técnico. Tercer Estudio Regional Comparativo y Explicativo (TERCE. Santiago, Chile)
[10] Hattori H 2019 Demographic and socioeconomic determinants of school attendance: An analysis of household survey data (UIS/UNICEF, 2015). (Montreal: UNESCO-UIS)
[11] Kolen M J and Brennan R L 2014 Test equating, scaling, and linking. (Springer, New York)
[12] Royce E C 2009 Poverty and Power: The Problem of Structural Inequality
[13] Bracho T 2017 Índice de déficit en competencias ¿Avanzamos hacia la garantía del derecho a la educación? en *Reformas y Políticas Educativas* No. 4, Sep-Dec 2017 (FCE, México)
[14] Sánchez-Restrepo H, Louçã J 2020 Topological Properties of Inequality and Deprivation in an Educational System: Unveiling the Key-Drivers Through Complex Network Analysis. In: Ahram T., Karwowski W., Pickl S., Taiar R. (eds) *Human Systems Engineering and Design II. IHSED 2019. Advances in Intelligent Systems and Computing*, Vol 1026 (Springer, Cham)
[15] van Borkulo C D, Borsboom D, Epskamp S, Blanken T F, Boschloo L, Schoevers R A, and Waldorp L J 2014 A new method for constructing networks from binary data. *Scientific Reports* **4**(5918) 1–10
[16] Dani A A, and Arjan D H, 2008 *Inclusive States: Social Policy and Structural Inequalities.* (Washington, D.C: World Bank)
[17] OECD 2018 Equity in Education: Breaking Down Barriers to Social Mobility, (PISA, OECD Publishing, Paris)
[18] Salazar M C and Glasinovich W A 2018 Child work and education: Five case studies from Latin America. (Routledge)
[19] Sánchez-Restrepo H, Louçã J 2020 Inequality in Learning Outcomes: Unveiling Educational Deprivation Through Complex Network Analysis. In: Cherifi H., Gaito S., Mendes J., Moro E., Rocha L. (eds) *Complex Networks and Their Applications VIII. Studies in Computational Intelligence*, vol 882. (Springer, Cham)
[20] Samman E 2018 SDG progress: fragility, crisis and leaving no one behind report.

CHAPTER 7

Topological Substitution: Mapping the Structural Mechanisms of Academic Resilience via Complex Network Analysis

First Autor: Harvey Spencer Sánchez Restrepo[1,2]
Harvey.Restrepo@iscte-iul.pt

[1] Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL 1649-026, Lisbon, Portugal

[2] Agencia Latinoamericana de Evaluación y Política Pública (Aleph)
14140, Mexico City, Mexico

Peer-reviewed chapter of the book already accepted and in press, in: *Socioeconomic Studies of Today's Mexico: A Mathematical Overview*. **Springer Nature**.

Topological Substitution: Mapping the Structural Mechanisms of Academic Resilience via Complex Network Analysis

Harvey Spencer Sánchez-Restrepo [1,2] [0000-0001-7311-8622]

[1] ISTAR Research Center. ISCTE-IUL,1649-026, Lisbon, Portugal
[2] Agencia Latinoamericana de Evaluación y Política Pública, 14140, Mexico City, Mexico
Harvey.Restrepo@iscte-iul.pt

Abstract. The persistence of the global learning crisis [UNESCO Institute for Statistics, 2024; The World Bank, UNESCO, & UNICEF, 2021; 2022] demands that educational policy move beyond simply describing achievement gaps toward understanding the structural mechanisms of resilience, which is the quantifiable capacity of students to sustain positive outcomes despite exposure to chronic adversity [Rutter, 2012]. Hypotheses regarding resilience often invoke dynamic models, emphasizing the interactive balance of protective and risk factors [Luthar, S. S., 2020]. However, the precise structural roles and interplays of socio-cognitive and socio-emotional assets and contextual factors for processes leading to positive outcomes are often difficult to discern. These structural dynamics often elude traditional linear models, which are unable to account for the multiple interdependencies and non-linear feedback that shape real-world processes. Here, we apply complex network analysis to model factors as nodes and their functional associations as edges derived from rank-based statistical correlations, thereby translating multivariate student resource profiles into interpretable topological architectures (Borsboom et al., 2021; Brinkhof et al, 2025). Our aim is to identify the specific topological mechanism that distinguishes academically resilient, high-performing students from their more privileged peers. We find that resilience is not a fixed trait but an emergent strategy of topological substitution: resilient networks compensate for socioeconomic disadvantage by concentrating influence on a reduced set of internally controllable assets (for example, self-efficacy and attentional control), which occupy intermediary positions with elevated centrality and foster local cohesion via increased clustering. This pattern implies that resilient systems achieve higher structural efficiency, a phenomenon consistent with empirical findings that centrality and local clustering jointly index nodes' capacity to mediate influence within differing network topologies.

Keywords: Academic Resilience, Complex Network, Socioemotional skills, Large-scale assessments, Adjacency-correlation matrix, Educational Policy.

1. Introduction

The 2030 Agenda's Sustainable Development Goal 4 (SDG 4) commits 193 countries to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" (UNESCO, 2015). Education is recognized not only as a fundamental human right but also as a transformative engine of development—reducing

poverty, advancing health, promoting gender equality, and strengthening peace and stability [Tonegawa, 2023; UNESCO, 2024]. Despite these lofty ambitions, most education systems persistently fail to provide to all students with the foundational knowledge and competencies required to thrive academically [Hanushek, E. A., & Woessmann, L., 2020; Mishchuk et al., 2024; The World Bank, UNESCO, & UNICEF, 2022].

As a key strategy for advancing evidence-based public policy, many governments have institutionalized *Large-Scale Assessments* (LSAs) to gather valid and reliable data on educational outcomes [Senior, C., & Sahlberg, P., 2025]. These assessments also capture a range of academic and non-academic variables, commonly referred to as Factors Associated with Learning (FAL), which provide deeper insights into the conditions influencing student performance [Motl, T. C. et al, 2018]. The primary objective of LSAs is to furnish high-quality, quantitative evidence to measure educational progress, estimate the causal or correlational impact of varying each FAL across different time scales and territorial domains [Tracey, B, & Florian, K., 2016; Davis, B., & Sumara, D., 2014]. Historically, the dominant analytical paradigm used to study educational phenomena postulating that educational gaps are predominantly explained through simple linear covariations between learning outcomes and individual factors [Moll, K. et al, 2016]. However, the methodological efficacy of LSAs in informing policy has faced increasing scrutiny [Davis, B., & Sumara, D., 2014; Brinkhof, L. P. et al, 2025; Sykes, C., et al 2025]. This traditional approach is increasingly recognized as both, useful and insufficient [Johnson, J. H., 2016; Brinkhof, L. P. et al, 2025], probably because its almost exclusive reliance on linear regression models has been leading to a fragmented view of the highly interrelated educational ecosystem, and the wide range of complex educational phenomena [van der Maas, H., 2017]. Crucially, these simplified models often fails to capture the deep-seated, structural nature of inequality in learning outcomes and their inherent relationship with socioeconomic and educational deprivation [15]. Probably, this methodological limitation—the inability to model complexity, non-linearity, and interactions—is a primary reason for the observed lack of explanatory power in many educational policy models [Johnson, J. H, 2017].

Worldwide, students from socioeconomically disadvantaged backgrounds persistently face structural and pedagogical barriers that restrict educational access and diminish learning outcomes [Brinkhof et al., 2025; Luthar, 2020]. However, this deterministic perspective is challenged by a small yet significant subset of learners from disadvantaged socioeconomic backgrounds—often described as *academically resilient learners* (henceforth: ARL)—who attain high levels of academic success despite facing severe social and economic adversity [Rudd, G., Meissel, K., & Meyer, F., 2021; Yan, Y., & Gai, X., 2022]. Nevertheless, much of the existing literature continues to rely on linear modeling approaches that obscure the systemic interdependencies underlying these processes. In contrast, complex network theory offers a robust analytical framework capable of capturing the non-linear, reciprocal relationships among diverse variables, strata, and contexts that might be helpful to shape students' capacity to learn and succeed under adversity [Siew, 2020; Saqr et al., 2024]. This perspective provides fertile ground for advancing equity-oriented educational research and informing more contextually responsive policy and practice [OECD, 2021; 2025].

2. Literature review

2.1. Academic resilience

Academic performance is widely seen as the primary indicator of educational success and has been widely used in the academic resilience literature [Martin, A. J., & Marsh, H. W., 2006; Rudd, K. L., & Yates, T. M., 2020; Wang, F., 2025]. In educational research, family Socioeconomic Status (SES) has been identified as one of the most significant factors affecting students' educational experiences. According to Marta & Michelle (2023), approximately 15% of the variance in academic achievement can be explained by family SES. Low SES is often associated with challenging learning conditions and reduced motivation, contributing to lower academic achievement, and students achieving better than expected educational outcomes, use to be defined as high-level academic performers [Wang, F. 2025; Sirin, S. R. 2005].

The notion of resilience finds its conceptual roots in psychology, where it emerged as an effort to explain why some individuals achieve positive developmental or psychological outcomes despite exposure to significant stress or adversity (Luthar et al., 2020; Rutter, 2012; Xenofontos & Mouroutsou, 2023). Early formulations tended to describe resilience as a stable personal trait—a characteristic possessed by certain individuals who managed to "do well" under pressure. Academic resilience refers to "the heightened likelihood of success in school and other life accomplishments, despite environmental adversities brought about by early traits, conditions, and experiences" [Wang, F., 2025]. Defined broadly as the capacity to rebound from adversity and to maintain engagement in learning despite challenging circumstances, resilience encompasses qualities such as persistence, adaptability, and self-regulation [Rudd, G., Meissel, K., & Meyer, F., 2021]. Research conducted in highly complex educational contexts revealed that students' resilience strategies often emerge through leveraging social support, recognizing institutional affordances, and reframing identity narratives around persistence [Rudd, G., Meissel, K., & Meyer, F., 2021]. For instance, a longitudinal diary study with undergraduate students found that greater weekly variability in resilience was associated with improved short-term academic performance, mediated by self-regulation behaviours, and further moderated by grit and social support [Delprato, M., 2025; Veenhoven, R., 2024].

One of the most common approaches in measuring resilience involves asking participants to answer self-report questionnaires about traits that are usually related to resilience (also known as the latent construct approach). For example, Martin and Marsh (2006) used an academic resilience survey to capture students' perceptions of their ability to overcome academic challenges. Another approach that generally is much more objective uses two predetermined criteria in measuring academic resilience (also known as the *definition-driven approach* (DDA), [Rudd, R. E., Wingo, J., & Bynum, L., 2021]), and its main criterion is the presence of chronic adversity that hinders students' academic success. The DDA integrates students' relative positions in family SES and academic achievement compared to their peers, which has been widely used in previous

the SES distribution yet attain academic outcomes within the highest achievement percentiles of their respective country or region. From a global assessment perspective, this methodological framework substantially enhances cross-national comparability by standardizing resilience metrics across heterogeneous educational contexts. Such an approach is regarded as particularly rigorous [Rudd & Yates, 2020], as it relies on objective performance indicators rather than self-reported measures, thereby reducing susceptibility to socially desirable responding and cultural bias, contributing to a more robust and comprehensive understanding of structural resilience within educational systems.

2.2. Learning poverty

The construct *Learning Poverty* (LP) refers to "The ability to read with comprehension [as] a foundational skill that every education system around the world strives to impart by late in primary school […] as it's essential for all future learning and development" [World Bank, UNESCO & UNICEF, 2022]. The indicator is a composite measure combining the proportion of primary-aged children who are out-of-school (Schooling Deprived, SD) with the proportion of students in school who perform below minimum proficiency levels in reading (Learning Deprived, LD). In the context of the SDG 4, according with the *Global Alliance to Monitor Learning* (GAML, 2021), learning deprivation is defined as share of children at the end of primary who read at below the minimum proficiency level [Azevedo, J. P., 2020]. This combined measure effectively captures the dual imperative of educational policy: ensuring "more schooling" (access) and guaranteeing "better learning" (quality), thus confirming that time spent in the classroom translates into essential skills and capabilities.

Current global estimates for the 2019–2025 period indicate that nearly 60% of all 10-year-olds worldwide—equivalent to more than 390 million children—experience learning poverty [World Bank, UNESCO & UNICEF, 2021]. The situation is considerably more severe in low-and middle-income countries, where over 70% of children lack basic reading comprehension skills, a condition that has been further exacerbated by the prolonged school closures and learning disruptions caused by the COVID-19 pandemic [UNESCO Institute for Statistics, 2024; World Bank, UNESCO & UNICEF, 2022]. In Latin America and the Caribbean, learning poverty has escalated to approximately 79%, representing one of the most dramatic declines in foundational learning outcomes globally and effectively erasing two decades of educational progress. In Mexico, recent joint assessments by the World Bank and UNESCO estimate that, at the end of primary school, around 77% of children are unable to read and comprehend a simple text [World Bank, UNESCO & UNICEF, 2022].

2.3. Complex networks

Network science has become an invaluable tool across scientific fields for solving complex problems, offering a robust set of topological, statistical, and combinatorial measurements capable of differentiating unique characteristics within the data

[Barabási, A. L., 2003; Liu, Y. Y., & Barabási, A. L. 2016]. Key topological metrics frequently used to diagnose network behavior involve analyzing paths and connectivity—the sequence and existence of edges linking nodes [Oldham et al., 2019]. These include essential measures such as node centrality (e.g., Degree, Betweenness), graph distances (which rely on path length), and metrics for clustering and community structure (modularity) [Malliaros & Vazirgiannis, 2016].

A graph remains the fundamental mathematical structure consisting of nodes (objects) and edges (relationships), weighted to quantify connection strength [Newman, M. E. J., 2018; Brass, D. J., & Borgatti, S. P., 2020]. Due to the inherent complexity of the systems they represent, these mathematical structures are often termed complex networks [Serrano, M. A., et al 2007]. Complex networks arise as a tool to geometrically represent the mutual inter-connections among the elements of a real-world system. From a mathematical viewpoint, a network is defined as a graph $G(N_v, N_e) = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1,2, \dots, N_v\}$ is a set of N_v labeled nodes (or vertices) and the a of N_e links (or edges) are represented by $\mathcal{E} = \{1,2, \dots, N_e\}$ [Newman, M. E. J., 2018]. The interconnections between nodes are usually represented by means of a binary and symmetric adjacency matrix, $A_{i,j}$, defined as

$$A_{i,j} = \begin{cases} 0, \text{if } i = j \text{ or } (i,j) \notin \Omega(\Phi, C_k) \\ \omega, \qquad \text{otherwise.} \end{cases} \tag{1}$$

The matrix defined in (1) is the simplest representation of a network preserving only the presence or absence of a connection between two nodes. Once the adjacency matrix is established, each variable becomes a node, and edges connect nodes with a "1" in the matrix. This representation enables network-based analyses, including community detection. Algorithms such as modularity maximization partition the network into communities—groups of nodes that are more densely connected internally than externally. Modularity serves as a measure of partition quality, with higher values indicating stronger community structure. To ensure robustness, researchers often compute network metrics across multiple thresholds to avoid results being driven by arbitrary choices [Newman, M. E. J., 2016; 2018]. Although optional, this step can be replaced by constructing weighted networks, which are computationally more demanding. Threshold selection is a critical and often debated issue, as it strongly influences network topology; factors such as sample size, type I error rate, network density, and data characteristics typically guide this choice (Masuda et al., 2024).

An adjacency-correlation matrix **R** is a particular type of adjacency matrix that uses to be very useful for uncovering latent structures among a large set of variables while aiding dimensionality reduction [Borsboom, D., 2021]. This symmetric matrix has diagonal elements $r_{i,i} = 1$, and off-diagonal elements $r_{i,j} = r_{j,i} \epsilon [-1,1]$ representing the strength and direction of the association between two ranked variables V_i and V_j [Masuda, 2024]. In contemporary network analysis, the matrix **R** captures the statistical associations among observed variables and enable their representation as interconnected nodes within a network [Newman, M. E. J., 2018]. Once established, these networks can be examined to identify community structures—clusters of nodes that exhibit stronger internal connections relative to their links with the rest of the system.

Community detection methods, such as modularity maximization, provide a quantitative means to assess and refine these structures, where higher modularity values indicate more coherent internal organization [Newman, M. E. J., 2018].

For this study, each network is developed on correlations among latent variables [Bolt T.S., 2016] based on a wide dataset of variables representing students and schools. The rationale for modeling multivariate educational data using a correlation matrix, rather than traditional linear methods, lies in the paradigm shift toward viewing psychological and social phenomena as inherently networked systems. While methods like factor analysis examine latent constructs and regression assesses prediction, they fail to capture the structure of interdependency among variables [Borsboom et al., 2021]. The increased use of this framework in education allows researchers to move beyond traditional linear models by transforming the pairwise statistical relationships into functional edges. This also allows for the direct application of network topology metrics—such as centrality and modularity—to reveal how factors cluster and propagate influence [Oldham et al., 2019], as well as capturing the systemic interdependencies and structural vulnerabilities driving outcomes like academic resilience [Martin, A. J., & Marsh, H. W., 2006; Delprato, M., 2025].

2.4. Dataset

This study draws on official datasets from the state of Nuevo León, collected during the year 2025 (https://www.aleph.la/Aprende/NLA/data/2025database). The official dataset NLA25-D is derived from a census-based process that encompasses both public and private schools, ensuring comprehensive coverage of all municipalities in the state and including approximately 97% of schools [NLA, 2025]. This multivariate dataset incorporates an index that captures the learning outcomes of ninth-grade students at the end of the lower secondary education. According to official reports, the assessment results were analysed using Item Response Theory (IRT) to validate the scales, establish construct validity, and examine measurement invariance over time [NLA, 2025]. The methodology adheres to international standards and demonstrates high reliability, as evidenced by Cronbach's alpha coefficients exceeding 0.80, indicating that the instruments consistently and accurately measure their intended constructs over the period 2021-2025 [NLA, 2025].

The primary NLA25D dataset comprises 83 800 student records from 1 066 schools, all assessed at the end of ninth grade. It includes high-quality linked microdata, such as responses from the associated factors survey [NLA, 2025]. Students' abilities for three main areas: Literacy, Numeracy and Science, were estimated using IRT with a two-parameter logistic model (2PL), where higher learning levels correspond to higher ability estimates $\theta_i \in (-\infty, \infty)$ [Borsboom, D., & Molenaar, D., 2015]. To identify low-level learners (LLL) and high-level learners (HLL), learning outcomes were classified into m achievement levels (L_m).

To contextualize student achievement within broader societal differences, and to address critical questions regarding educational opportunity and inequalities in learning outcomes, this LSA incorporates a comprehensive socioeconomic index, a composite index, identified as the *Index of Economic, Social, and Cultural Status* (ISEC),

constructed from three key components: 1) parental educational attainment and occupational status; 2) home possessions (as and book count); and 3) home services (including dwelling materials, access to sanitation, waste collection, internet and telephony). The estimates of the ISEC score and its corresponding SES deciles $(D_j), j = 1, \ldots, 10$ are directly employed in this study. Complete methodological documentation and raw datasets are openly available through the official platform, ensuring transparency and enabling full replicability of the analyses [NLA, 2025]. Data analyses were conducted using R (version 4.4.3), with the 'corrplot' package employed to estimate and visualize the correlation matrix. Network analysis was performed using Gephi (version 0.10). Statistical inference followed algorithms proposed by Zhang & Peixoto (2020), while betweenness centrality based on correlation adjacency matrices was computed using the approach described by Wang (2025). School weights were applied to correct sampling errors and enhance the representatives of the findings. This database offers a unique opportunity to examine associations among a wide range of variables at the end of the low secondary school.

3. Measuring learning

3.1. Levels of achievement

Following the definition of learning poverty [Azevedo, J. P. 2020], students' literacy abilities are seen as the main students' learning outcomes. For this reason, they are analyzed as dependent variables of a wide variety of academic and non-academic variables collected for the NLA25D dataset [NLA, 2025]. Within NLA framework, literacy is conceptualized as the comprehension and command of the language system and its linguistic conventions —skills that constitute the foundation of effective communication and, ultimately, lifelong learning [Aleph, 2022]. In this sense, literacy functions as a transversal competency, shaping both students' academic development and their broader social participation. This domain encompasses spelling, morphology, and syntax, while also emphasizing students' capacity to communicate with clarity and precision in both oral and written forms.

To interpret students' mastery in literacy, NLA has developed a classification for all students [Aleph, 2022] in three levels of achievement $\{L_1, L_2, L_3\}$, stablished by a *Bookmark process* [Borsboom, D., & Molenaar, D., 2015], a standard procedure based on psychometrical and qualitative analysis for establishing the threshold for splitting students who do not meet the minimum proficiency level specified by the cut off point ψ_1 [Borsboom, D., & Molenaar, D., 2015]. Students performing at Level 1 (L_1) exhibit significant learning deficits that hinder continued educational growth, situating them in conditions of learning deprivation and, potentially, social vulnerability. Conversely, students performing at Level 3 (L_3) belong to the highest level, corresponding to consolidated knowledge that enable the acquisition of more complex competencies. The distribution of students across these levels—expressed as rates—serves as a key diagnostic indicator of educational equity and helps to highlight gaps within the system (Sánchez-Restrepo, H. & Carmona-Soto, M.B., 2023).

constructed from three key components: 1) parental educational attainment and occupational status; 2) home possessions (as and book count); and 3) home services (including dwelling materials, access to sanitation, waste collection, internet and telephony). The estimates of the ISEC score and its corresponding SES deciles $(D_j), j = 1, \dots, 10$ are directly employed in this study. Complete methodological documentation and raw datasets are openly available through the official platform, ensuring transparency and enabling full replicability of the analyses [NLA, 2025]. Data analyses were conducted using R (version 4.4.3), with the 'corrplot' package employed to estimate and visualize the correlation matrix. Network analysis was performed using Gephi (version 0.10). Statistical inference followed algorithms proposed by Zhang & Peixoto (2020), while betweenness centrality based on correlation adjacency matrices was computed using the approach described by Wang (2025). School weights were applied to correct sampling errors and enhance the representatives of the findings. This database offers a unique opportunity to examine associations among a wide range of variables at the end of the low secondary school.

3. Measuring learning

3.1. Levels of achievement

Following the definition of learning poverty [Azevedo, J. P. 2020], students' literacy abilities are seen as the main students' learning outcomes. For this reason, they are analyzed as dependent variables of a wide variety of academic and non-academic variables collected for the NLA25D dataset [NLA, 2025]. Within NLA framework, literacy is conceptualized as the comprehension and command of the language system and its linguistic conventions —skills that constitute the foundation of effective communication and, ultimately, lifelong learning [Aleph, 2022]. In this sense, literacy functions as a transversal competency, shaping both students' academic development and their broader social participation. This domain encompasses spelling, morphology, and syntax, while also emphasizing students' capacity to communicate with clarity and precision in both oral and written forms.

To interpret students' mastery in literacy, NLA has developed a classification for all students [Aleph, 2022] in three levels of achievement $\{L_1, L_2, L_3\}$, stablished by a *Bookmark process* [Borsboom, D., & Molenaar, D., 2015], a standard procedure based on psychometrical and qualitative analysis for establishing the threshold for splitting students who do not meet the minimum proficiency level specified by the cut off point ψ_1 [Borsboom, D., & Molenaar, D., 2015]. Students performing at Level 1 (L_1) exhibit significant learning deficits that hinder continued educational growth, situating them in conditions of learning deprivation and, potentially, social vulnerability. Conversely, students performing at Level 3 (L_3) belong to the highest level, corresponding to consolidated knowledge that enable the acquisition of more complex competencies. The distribution of students across these levels—expressed as rates—serves as a key diagnostic indicator of educational equity and helps to highlight gaps within the system (Sánchez-Restrepo, H. & Carmona-Soto, M.B., 2023).

$$E[\mathcal{N}_t(\Phi, L_1, C_k)] = \sum_{n=0}^{\infty} E\left[\sum_{j=1}^{\mathcal{N}_t(C_k)} \mathbb{I}_{i,t}(\theta_i; \Phi,\ C_k)|\mathcal{N}_t(C_k) = n\right] \cdot P[\mathcal{N}_t(C_k) = n] \quad (3)$$

$$= \sum_{n=0}^{\infty} E\left[\sum_{j=1}^{E[\mathcal{N}_t(C_k)]} \mathbb{I}_{i,t}(\theta_i; \Phi, L_1, C_k)|\mathcal{N}_t(C_k) = n\right] \cdot P[\mathcal{N}_t(C_k) = n]$$

$$= \sum_{n=0}^{\infty} n_t \cdot E\left[\mathbb{I}_{i,t}(\theta_i; \Phi, L_1, C_k)\right] \cdot P[\mathcal{N}_t(C_k) = n]$$

$$= E\left[\mathbb{I}_{i,t}(\theta_i; \Phi, L_1, C_k)\right] \cdot \sum_{n=0}^{\infty} n_t \cdot P[\mathcal{N}_t(C_k) = n]$$

$$\mathbb{E}[\mathbb{I}_t(\theta_i; \Phi, L_1, C_k)] \cdot \mathbb{E}[\mathcal{N}_t(\Phi, C_k)] = \boldsymbol{P}[(\theta_i < \psi_1; \Phi, L_1, C_k)] \cdot n_t(\Phi, C_k)$$

$$= p_{L_1}(\Phi, C_k) \cdot n_t(\Phi, C_k) \quad (4)$$

Given that $\Omega(\Phi, C_k)$ is the total population of students in the community C_k and $n_t(\Phi, C_k)$ is the expectation of this RV, then $p_{L_1}(\Phi, C_k)$ is the probability of any student in the subset C_k might be learning deprived, i.e., that it is located at level L_1, then, we have the following general equation

$$p_{L_1}(\Phi, C_k) = \frac{n_t(\Phi, L_1, C_k)}{n_t(\Phi, C_k)} = H(\Phi, C_k) \quad (5)$$

Equation (3) establishes the statistical way for estimating the RAD for any subset $\Omega_{C_k}(\Phi)$. In fact, the equation is held for any family of thresholds $\{L_k\}_{k>0}$. Considering different communities, stratification, SES categories, and all levels of achievement, the general expression for estimating $n(\Phi; L_1, \varepsilon_g, D_j, C)$ becomes

$$n(\Phi, L_1, \varepsilon_g, D_j, C) = \sum_{p=1}^{q} n_p(\Phi, L_1, \varepsilon_i, D_j C) = \sum_{w=1}^{q} \sum_{k=1}^{T} n_w(\Phi, L_1, \varepsilon_i, D_j, C_k) \quad (6)$$

In equation (4), the number $n(\Phi; L_1, \varepsilon_g, D_j, C)$ represents the weighted population group of students divided by a partition with q categories, bounded by the stratum ε_g, coming from the SES-decile D_j, collecting all cases through the space $k = 1,2,3 \ldots T$, where T is a finite number of communities among students suffering learning deprivation, grouped by L_1.

3.3. Relative Deprivation

After having a partition of all those students deprived of learning, a natural question is how deep the level of deprivation of each of them is, a concept linked with sociological proposal that states that 'needs, thresholds and satisfactions are determined by each society', objectively defined by the cut-off point ψ_1. For each deprived student, i.e.,

each student in the L_1-group has a measure of the deepeness of her/his deprivation given by the *Relative deprivation index* (Red) (Sánchez-Restrepo, H., 2018; 2021), which can be estimated through the RV

$$\Lambda_i(\theta_i; \varepsilon_g, D_j, C_k) \in [0,1] \tag{7}$$

This RV measures the intensity of learning deprivation as a real value showing the gap between the ability' student and the first cut-off point ψ_1, named the *relative deprivation index* (Red_i), a distance function of the ability θ_i considering the trajectory given by the function (SF) $S_{\Lambda_i}(\theta_i; \Phi, C_k) = 1 - F_{\Lambda_i}(\theta_i; \Phi, C_k)$, the complement of the probability distribution function (PDF), $F_{\Lambda_i}(s; \varepsilon_g, D_j, C_k) = \boldsymbol{P}[(\theta_i < s | L_1, \varepsilon_g, D_j, C_k)]$, where the inclusion of L_1 as a restriction of the measure space implies the boundary $\psi_0 < s < \psi_1$. Equation (6) shows the analitycal relationship among functions

$$S_{\Lambda_i}(\theta_i; \Phi, C_k) = 1 - \int_{\psi_0}^{\theta_i} dF_{\Lambda}(s; \Phi, \varepsilon_g, D_j, C_k) ds = \int_{\theta_i}^{\psi_1} dF_{\Lambda}(s; \Phi, \varepsilon_g, D_j, C_k) ds \tag{8}$$

Now, as define for equation (1), every student in the sub-space $\Omega(\Phi, C_k)$ follow the same distribution of abilities and given that the subset of RV are i.i.d, their individual expectation $\Lambda(\Phi, \varepsilon_g, D_j, C_k)$ is given by the equation (7)

$$\mathbb{E}[\Lambda(\Phi, \varepsilon_g, D_j, C_k)] = \int_{\psi_0}^{\psi_1} s \cdot dF_{\Lambda}(s; \Phi, \varepsilon_g, D_j, C_k) ds = \lambda(\Phi, \varepsilon_g, D_j, C_k) \tag{9}$$

3.4. Learning Deprivation Index

From the two measures explained in the previous sections, it might be useful to have a synthetic index for representing learning deprivation of groups of students for different strata, population groups, or territorial demarcations of interest. The Learning Deprivation Index (LDI) is a composition of RAD and Red indexes as a compound stochastic process (CSP) of both stochastic processes [Sanchez-Restrepo, H. & Jorge Louçã, 2019], which is equivalent to adding the intensity functions of all students in L_1, and weighting the result for all students grouped by $\mathcal{N}_t(\Phi; \varepsilon_g, D_j, C)$. As shown in equation (7), $LDI_t(\Phi, C_k)$, denoted by $\delta_t(\Phi, C_k)$, It is a synthetic measure of total learning deprivation (TLD), observed in the student universe $\Omega(\Phi, C_k)$, at point *t*, which can be estimated by aggregating the individual deprivation of each of the realizations, i.e., the expectation of the CSP, which can be written as follows

$$\delta_t(\Phi, C_k) = \mathbb{E}\left[\sum_{\forall i \in L_1}^{\mathcal{N}_t(C_k)} \mathbb{I}_t(\theta_i; \Phi, L_1, C_k) \cdot \Lambda(\theta_i; \Phi, C_k)\right] \tag{10}$$

Given the counting process $\mathcal{N}_t$ is independent generated from the random variable $\Lambda(\theta_i; \Phi, C_k)$, and $\mathbb{I}_t(\theta_i; \Phi, L_1, C_k)$, equation (9) holds for any sub-space (Φ, C_k)

$$\delta_t(\Phi, C_k) = \mathbb{E}[\mathcal{N}_t(\Phi, C_k)] \cdot \mathbb{E}[\mathbb{I}_t(\theta_i; \Phi, L_1, C_k)] \cdot \mathbb{E}[\Lambda(\theta_i; \Phi, C_k)]$$
$$= p_{L_1}(\Phi, C_k) \cdot \lambda_i(\Phi, C_k) = H(\Phi, C_k) \cdot \lambda_i(\Phi, C_k) \quad (11)$$

Because the LDI is a standardized and bounded metric, it provides a consistent basis for comparing different subsets of students. This stability makes it particularly valuable for assessing the magnitude of learning gaps across time and regions, both in terms of the absolute rate of learning deprivation and the intensity of that deprivation (Λ_k).For NLA25-9, $n(\Omega, L_1) = 43\,280$, $n(\Omega, L_2) = 17\,661$ and $n(\Omega, L_3) = 22\,859$. Thus, the rate of absolute deprivation $H(\Omega) = 0.5165$ points out that a little bit more than 1 in 2 students have not learned the minimum expected. Regarding gender, the rates $H(\Omega, Male) = 0.5428$ and $H(\Omega, Female) = 0.4901$ states that men experience a higher rate of deprivation, the gap of -0.0527 represents 10.8% of more male students in this group, confirming a global phenomenon: women tend to develop better learning in literacy [Marta, E. M., & Michelle, C., 2023].

4. Unveiling the key factors

4.1. Socioeconomic status

The analysis of educational deprivation across socioeconomic deciles reveals a pronounced and consistent gradient, underscoring how socioeconomic status fundamentally shapes vulnerability, further compounded by persistent gender disparities. Table 1 presents the statistics of the three main indicators of deprivation. At student-level, results indicated that ISEC was negatively correlated with absolute learning deprivation ($r_M = -0.233$ and $r_F = -0.262$, $p < 0.001$), and once being deprived also with the intensity of deprivation ($r_M = -0.125$ and $r_F = -0.135$, $p < 0.001$). This monotonous inverse relationship shows that educational risk peaks at the lowest end of the socioeconomic spectrum.

When grouping by deciles, correlations among SES, RAD, Red and LDI become higher than 0.95, giving a clear opportunity for measuring gaps between advantaged and disadvantaged students. For instance, the RAD metric—reflecting systemic failure to ensure foundational learning—registers a value nearly 2.3 times greater for the most socioeconomically disadvantaged male students (D01: 0.689), than for the most advantaged female students (D10: 0.235). If one compares the gender gaps found, it is interesting to note that the pattern persists in simple and composite indicators. For the poorest male students, the LDI exceeds three times the deprivation observed among the wealthiest male students ((D01: 0.138 vs D10: 0.045), highlighting the cumulative and disproportionate impact of deprivation on the poor. However, a closer look at relative deprivation states that disadvantaged groups falls within a closer range (D01 Male: 0.200 vs. D10 Male: 0.148). Both metrics indicate that poverty is the dominant structural determinant, while mechanisms influencing outcomes among the deprived appear less sensitive to extremes of SES distribution [Hanushek & Woessmann, 2020].

Table 1. Socioeconomic distribution of learning deprivation indicators by gender.

	RAD				Red				LDI			
	Male		Female		Male		Female		Male		Female	
	$\hat{\mu}(\cdot)$	$\widehat{\text{se}}(\cdot)$	$\hat{\mu}(\cdot)$	$\widehat{\text{se}}(\cdot)$	$\hat{\mu}(\cdot)$	$\widehat{\text{se}}(\cdot)$	$\hat{\mu}(\cdot)$	$\widehat{\text{se}}(\cdot)$	$\hat{\mu}(\cdot)$	$\widehat{\text{se}}(\cdot)$	$\hat{\mu}(\cdot)$	$\widehat{\text{se}}(\cdot)$
State	**0.543**	**0.002**	**0.490**	**0.002**	**0.176**	**0.001**	**0.156**	**0.001**	**0.095**	**0.001**	**0.077**	**0.001**
D01	0.689	0.008	0.673	0.007	0.200	0.003	0.178	0.003	0.138	0.002	0.120	0.002
D02	0.658	0.008	0.628	0.007	0.188	0.003	0.168	0.003	0.124	0.002	0.105	0.002
D03	0.639	0.007	0.588	0.008	0.186	0.003	0.165	0.003	0.119	0.002	0.097	0.002
D04	0.628	0.007	0.540	0.008	0.180	0.003	0.156	0.002	0.113	0.002	0.084	0.002
D05	0.579	0.008	0.529	0.008	0.177	0.003	0.157	0.002	0.102	0.002	0.083	0.002
D06	0.570	0.008	0.490	0.008	0.173	0.003	0.149	0.002	0.098	0.002	0.073	0.002
D07	0.541	0.008	0.456	0.008	0.169	0.003	0.144	0.002	0.091	0.002	0.065	0.002
D08	0.462	0.008	0.401	0.008	0.162	0.002	0.143	0.002	0.075	0.002	0.057	0.002
D09	0.387	0.007	0.323	0.007	0.147	0.002	0.134	0.002	0.057	0.002	0.043	0.002
D10	0.304	0.007	0.235	0.007	0.148	0.002	0.122	0.002	0.045	0.001	0.029	0.001

Note. SES-deciles are estimated among all participant students, including public and private schools. RAD and Red were estimated directly from cut-off point of Literacy ability's continuous latent variable. Standard errors are calculated at $p < 0.05$.

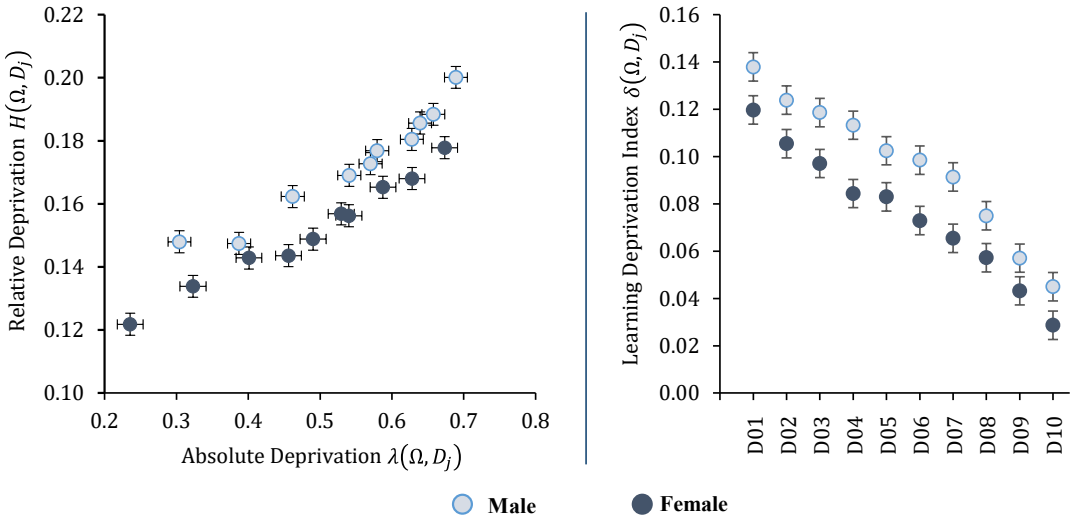


Fig. 1. In the left side, plotting RAD against Red shows socioeconomic gradients. The right plot displays the LDI, highlighting gender disparities across SES deciles.
Note: gender is coded as female = 1 and male = 0. Standard errors are reported at a significance level of $p < 0.05$.

Regarding gender disparity, Figure 1 illustrates a clear, statistically significant pattern across the entire socioeconomic distribution: males consistently exhibit higher deprivation metrics than females in nearly all deciles. This trend is evident in RAD and Red indexes, as well as in the compound index LDI. This robust finding reinforces the interpretation of gender as an independent axis of educational vulnerability that interacts with, but is not fully explained by, socioeconomic stratification. The data suggest

the need for interventions that specifically address the unique risk factors contributing to male underperformance in foundational skills.

4.2. Resilient and Privileged Networks

Following the data-driven definition approach of ARL, the HLL dataset (L_3) was constructed by selecting students from the two lowest and two highest SES deciles. Subsequently, students within L_3 were classified into two groups: privileged (L_3^{Q5}) and resilient (L_3^{Q1}). From the initial population of 83,600 students, 16,760 (20%) were selected from 859 schools (80.6%), with female students representing 53.7% of the group. Among these low-SES students, 2 673 (16.1%) were identified as ARL, coming from 603 (56.6%) schools, where 55.2% of them are females. For avoiding bias due to possible exceptional conditions associated with private schools, 113 students (0.7%) were removed from the analysis. For estimating the matrix **R,** all entries were estimated by the Spearman rank, a method that provides a valuable nonparametric approach for identifying structured dependencies in complex datasets [Zhu, J. et al 2023]. By ranking rather than directly comparing raw data values, it effectively captures the monotonic association between variables—reflecting consistent directional changes, whether increasing or decreasing, across attributes [Hauke & Kossowski, 2011; Schober P, Boer C, Schwarte LA, 2018], a property especially suitable for exploring nonlinear or irregular patterns that still preserve ordered relationships.

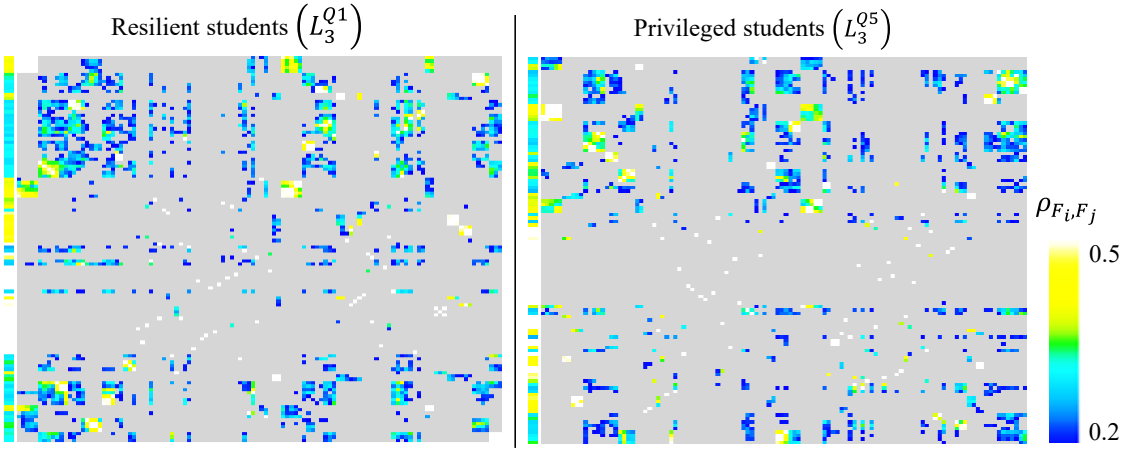


Fig. 1. Correlation matrices among factors for two student groups: Resilient students (HP-Q1, *left side*) and Privileged students (HP-Q5, *right side*). Only correlations between factors where $\rho_{i,j} > 0.2$ ($p < .001$) were considered for analysis.
Note. For SES, L1C=1, L3=5. For SES, Q1=1, Q5=5. SES and Level of Performance are partitions based on cut-off points of continuous latent variables.

The matrix for the whole dataset has 114 variables and its 6 555 entries includes autocorrelations. As can be expected, statistical significance of the covariances among factors is conditional to each subpopulation group, that help to explain why it is highly informative about the underlying structure of each group of students. For analyzing

equivalent to control the level of achievement for Q1 and Q5 and grouping them by SES to contrast. Figure 2 illustrates that the two matrices exhibit distinct patterns and varying levels of correlation, providing an effective way to identify and emphasize the factors common to both groups. This way of segmenting is highly informative about the underlying structure of the group, as the statistical significance of the covariances among factors is conditional to the dataset of students [Schober P, Boer C, Schwarte LA, 2018]. For removing noise and excess of misinformation, both matrices $\mathbf{R}\left(L_3^{Q1}\right)$ and $\mathbf{R}\left(L_3^{Q5}\right)$ were reduced for being just positive defined matrices—based on the sign of correlations— for all pairs of variables that were statistically significant ($p<0.001$) and they became the adjacency matrices, where the remaining number of positive entries for $\mathbf{R}^{+}\left(L_3^{Q1}\right)$ is 967 (14.75%), and for $\mathbf{R}^{+}\left(L_3^{Q5}\right)$ is 783 (11.95%).

4.3. Network Topology Divergence

The academic factor networks were constructed by defining every factor, represented by a related variable with a statistically significant correlation, as a node in the network. Edges were established between any pair of nodes with a calculated value in the adjacency matrix, where the weight of each edge corresponds to the correlation coefficient between those nodes (Bolt T.S., et al., 2016). This methodology laid the fundamental procedure for carrying out the subsequent comparative computation of both local and global metrics across the two segmented networks. The initial adjacency matrices for the resilient and privileged students revealed highly dense network structures. Specifically, the $\mathbf{N}^{+}\left(L_3^{Q1}\right)$ matrix initially comprised 853 edges among 31 nodes, while the $\mathbf{N}^{+}\left(L_3^{Q5}\right)$ matrix featured 669 edges among 28 nodes. To ensure topological clarity and efficiency in the subsequent analysis, the methodology of averaging parallel edges was applied, reducing the networks' dimensionality to 158 non-parallel edges (31 nodes) and 124 non-parallel edges (28 nodes), respectively. Following this crucial data simplification, we proceeded to calculate the core centrality measures and Betweenness centrality was prioritized among these metrics, as it is essential for identifying the specific nodes that mediate the flow of influence, thereby acting as indispensable structural bridges within each resource network [Oldham et al., 2019].

The comparative analysis of centrality metrics—Closeness, Harmonic closeness, and Betweenness (see Table 2)—confirms a fundamental divergence in structural strategy underpinning high academic performance between the resilient and privileged students. In both networks, *Emotional Well-Being* acts as the single most critical bottleneck ($Q1: 0.32$; $Q5: 0.33$), affirming its role as the non-negotiable structural prerequisite for the flow of all other resources, regardless of socioeconomic status [Newman, M. E. J., 2018]. The capacity to regulate emotion and maintain well-being appears like the single most important factor for controlling the flow of all other academic and social resources into the high-performance system. However, the resilient network employs topological substitution [Liu, C., Ding, J., Song, Y., & Li, Y., 2024] by forcing key internal assets into mediating roles: the factors associated with internal effort, such as *Metacognition* ($Q1: 0.007$ vs. $Q5: 0.00$), and *Study habits* ($Q1: 0.06$ vs. $Q5: 0.00$), are

structurally integrated as necessary bridges within the $Q1$ system, yet remain peripheral in the $Q5$ system.

While the privileged network exhibits slightly higher overall closeness centrality for key socioemotional factors—suggesting marginally faster access to core resources—the $Q1$ system demonstrates a pronounced functional necessity by structurally integrating active self-management processes (e.g., *Emotional Regulation*, $Q5: 0.02$ vs. $Q5: 0.00$) to overcome resource scarcity [Borsboom et al., 2021]. The comparison here is not simply in the value, but in the distribution: a high, concentrated Betweenness in the resilient network implies dependence on structural compensators (e.g., self-regulation skills), whereas a low, distributed Betweenness in the privileged students network signals robust redundancy across resources. Closeness centrality further quantifies the efficiency of access, determining how quickly a factor can activate the entire resource network, thereby measuring the overall speed of adaptive.

Table 1. Network parameters for resilient and privileged students.

Factor	Weighted Degree		Closness centrality		Harmonic closness centrality		Betweeness centrality		Authority		Modularity class		Clustering		Eigen centrality		Katz Centrality	
	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1
Emotional Well-Being	4.74	4.65	0.70	0.72	0.78	0.82	0.32	0.33	0.24	0.31	0.24	0.31	0.03	0.04	1	3	0.40	0.41
Cooperation	5.87	4.92	0.70	0.70	0.82	0.81	0.21	0.23	0.29	0.32	0.29	0.32	0.04	0.04	2	2	0.43	0.45
Sense of Belonging	2.29	2.20	0.48	0.46	0.54	0.53	0.10	0.10	0.03	0.03	0.03	0.03	0.02	0.02	1	3	0.40	0.50
Bullying	2.85	2.15	0.52	0.51	0.59	0.57	0.09	0.06	0.07	0.07	0.07	0.07	0.02	0.02	2	4	0.50	0.60
Achievement Level	1.30	0.58	0.45	0.00	0.50	0.00	0.07	0.00	0.02	0.00	0.02	0.00	0.02	0.01	1	0	0.33	0.00
Study Habits	2.55	1.18	0.49	0.38	0.58	0.43	0.06	0.00	0.07	0.02	0.07	0.02	0.02	0.01	1	1	0.29	0.67
Self-Efficacy	5.04	3.66	0.61	0.59	0.74	0.71	0.04	0.01	0.28	0.28	0.28	0.28	0.03	0.03	0	2	0.63	0.77
Motivation	4.54	3.41	0.61	0.58	0.74	0.69	0.04	0.07	0.28	0.25	0.28	0.25	0.03	0.03	0	2	0.63	0.70
Metacognition	4.82	3.78	0.60	0.59	0.72	0.71	0.04	0.01	0.26	0.28	0.26	0.28	0.03	0.03	0	2	0.63	0.69
Family Support	2.22	1.80	0.45	0.46	0.51	0.51	0.03	0.03	0.02	0.03	0.02	0.03	0.01	0.01	1	3	0.60	0.67
Self-Control	3.25	2.35	0.60	0.57	0.70	0.67	0.02	0.07	0.24	0.22	0.24	0.22	0.03	0.02	0	2	0.85	0.71
Emotional Regulation	4.00	2.72	0.59	0.52	0.69	0.60	0.02	0.00	0.24	0.18	0.24	0.18	0.03	0.02	0	2	0.80	0.82
School Violence	2.29	1.70	0.50	0.36	0.56	0.42	0.02	0.00	0.05	0.01	0.05	0.01	0.02	0.01	2	4	0.67	1.00
School Climate	1.28	1.02	0.45	0.46	0.53	0.53	0.01	0.01	0.05	0.04	0.05	0.04	0.01	0.01	2	4	0.60	0.83
Vitality	2.21	1.83	0.49	0.44	0.54	0.48	0.01	0.00	0.05	0.05	0.05	0.05	0.01	0.01	1	3	0.40	0.33
Cognitive Strategies	2.50	2.06	0.54	0.54	0.63	0.64	0.01	0.00	0.19	0.22	0.19	0.22	0.02	0.02	0	2	0.84	0.86
Satisfaction with Interactions	1.07	0.80	0.43	0.32	0.48	0.36	0.01	0.00	0.02	0.00	0.02	0.00	0.01	0.01	2	4	0.33	1.00
School Satisfaction	2.15	1.98	0.37	0.46	0.44	0.54	0.00	0.08	0.01	0.04	0.01	0.04	0.01	0.02	2	4	0.70	0.60
Attentional Control	4.76	3.99	0.59	0.59	0.71	0.71	0.00	0.01	0.27	0.28	0.27	0.28	0.03	0.03	0	2	0.87	0.76
Civic Engagement	4.61	3.84	0.59	0.59	0.71	0.72	0.00	0.01	0.27	0.29	0.27	0.29	0.03	0.03	0	2	0.87	0.63
Empathy	4.44	3.52	0.59	0.58	0.71	0.69	0.00	0.01	0.27	0.26	0.27	0.26	0.03	0.03	0	2	0.87	0.77
Assertive Communication	4.23	2.18	0.59	0.50	0.71	0.58	0.00	0.00	0.27	0.16	0.27	0.16	0.03	0.02	0	2	0.87	0.81
Leadership	4.17	4.56	0.58	0.60	0.69	0.74	0.00	0.04	0.25	0.30	0.25	0.30	0.03	0.03	0	2	0.90	0.54
Gender	0.68	0.45	0.48	0.38	0.52	0.44	0.00	0.00	0.05	0.02	0.05	0.02	0.01	0.01	1	1	0.33	0.00
Growth Mindset	3.99	2.79	0.52	0.55	0.66	0.66	0.00	0.00	0.25	0.24	0.25	0.24	0.03	0.02	0	2	0.89	0.80
Emotional Awareness	3.39	2.26	0.50	0.47	0.63	0.58	0.00	0.00	0.22	0.18	0.22	0.18	0.02	0.02	0	2	0.92	0.79
Household Composition	0.22		0.34		0.38		0.00		0.01		0.01		0.01		2		0.00	
Teacher Support	1.66	1.64	0.33	0.33	0.38	0.37	0.00	0.00	0.00	0.01	0.00	0.01	0.01	0.01	1	3	1.00	1.00
Environmental Safety	0.24	0.40	0.33	0.32	0.36	0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	1	3	0.00	0.00
Attendance and Punctuality	0.22		0.32		0.34		0.00		0.00		0.00		0.01		1		0.00	
Expectations	0.71		0.40		0.46		0.00		0.06		0.06		0.01		0		1.00	
Emotional Well-Being	4.74	4.65	0.70	0.72	0.78	0.82	0.32	0.33	0.24	0.31	0.24	0.31	0.03	0.04	1	3	0.40	0.41
Max	5.30	6.24	0.72	0.70	0.82	0.82	0.19	0.24	0.32	0.29	12.00	7.00	1.00	1.00	1.00	1.00	3.90	6.89
Min	1.76	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.00	2.00	0.00	0.00	0.00	0.00	1.25	1.25
$\widehat{\mu}(\cdot)$	3.15	3.71	0.41	0.46	0.48	0.54	0.02	0.02	0.13	0.13	6.50	3.44	0.57	0.63	0.40	0.46	2.29	3.79
$\widehat{\sigma}(\cdot)$	1.11	1.36	0.22	0.18	0.26	0.22	0.04	0.05	0.12	0.12	2.38	1.72	0.32	0.29	0.38	0.40	0.95	2.19

Note. All edges are undirected. Coefficients *Closeness centrality*, *Harmonic closeness centrality* and *Betweenness centrality* are normalize centralities [0,1]. All factors are supposed to be latent variables. Coefficients cannot be converted into an odds ratio.

For both networks, color of the nodes represents the *Modularity class*, their size represents *Katz centrality* coefficient, and their label size *Betweenness centrality.* In this study, Modularity is employed to identify densely connected communities of factors, providing structural evidence on whether the resilient group exhibits a highly integrated resource ecology (low modularity, typically indicating that resources operate synergistically) or a fragmented configuration. In parallel, the Clustering coefficient offers a local measure of network density around each node, signalling the presence of tightly knit functional reinforcement among immediate neighbours [Masuda et al., 2024]. Together, these metrics supply the quantitative foundation necessary to advance from mere correlational observations to structural and causal inferences regarding the effectiveness of distinct academic pathways.

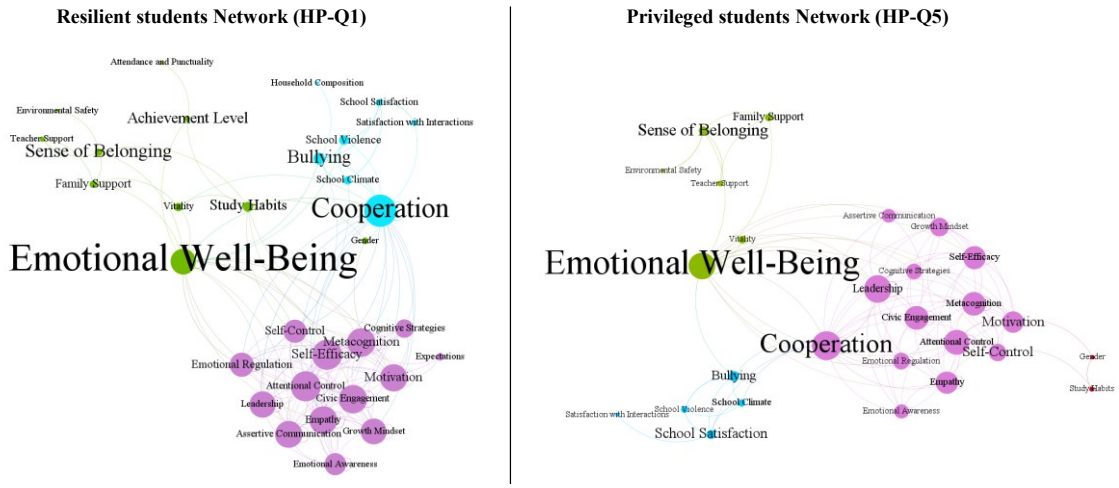


Fig. 1. Academic Factor Correlation Networks. The figure illustrates the structural and topological differences in resource organization between resilient (HP-Q1, *left*) and privileged (HP-Q5, *right*) high-performing students, quantified by node centralities and modular partitioning.

The comparative analysis of network topological properties provides evidence for a mechanism of structural compensation utilized by resilient students. While the privileged network maintains high performance through the stability of diffuse, redundant resources—evidenced by low variability in influence and moderate Clustering—, the resilient topology achieves success through highly optimized structural necessity. Specifically, the HP-Q1 network demonstrates enhanced structural efficiency (Katz centrality $K_{Q1} = 2.86$ vs. $K_{Q5} = 2.02$), where the factor Cooperation exerting the maximum total reach is much bigger than that of the Q5 network (5.04 vs. 3.25). This outcome is likely achieved through the strategic positioning of internal factors: core socio-cognitive skills such as Attentional Control, Self-Efficacy, and Metacognition remain highly stable and are functionally integrated within the central module—reflected by low modularity class counts—thereby forming a cohesive and reliable resource ecosystem [Masuda et al., 2018; 2024]. Furthermore, high *Eigenvector Centrality* scores for these Q1 factors (e.g., *Self-Efficacy* at 4.73 vs. Q5 at 2.94 and *Self-Efficacy* at 4.73 vs. Q5 at 2.76) confirm that resilient students leverage the highest-quality network

connections, substituting the systemic security of wealth with amplified relational influence and local cohesion (Clustering coefficient ≈0.87 for key factors).

On the other hand, resilient students are structurally forced to channel the resource flow through their own cognitive regulation (Self-Efficacy, Metacognition). In this sense, the factor Bullying holds a high bridging position in Q1, suggesting that managing environmental threat is an integrated, active mechanism needed to sustain performance. In privileged networks, factors such as attentional control, self-efficacy, and civic participation appear substantially less influential in achieving comparable cognitive outcomes than within the resilient student group. For Q1, key bridging roles—irrelevant in Q5—are held by factors such as Achievement Level (0.07), Study Habits (0.06), and Metacognition (0.04), indicating a reliance on stability and general contentment rather than on acute self-management strategies.

5. Discussion

This study conducted a systematic examination of the structural interplay between academic and socio-cognitive factors and high achievement across contrasting socioeconomic contexts, introducing a network-based conceptualization of resilience. Departing from traditional approaches that isolate variance in single predictors, this framework models the entire multivariate resource ecosystem, enabling the identification of specific topological mechanisms—rather than simple bivariate associations—that differentiate students who overcome socioeconomic disadvantage from their privileged high-performing peers. These findings represent a substantive shift from linear models, demonstrating how structural reorganization compensates for resource scarcity. Results indicate that high performance, irrespective of socioeconomic background, is fundamentally mediated by the stability of the resource network. This underscores the pronounced structural differences between high-achieving students from disadvantaged and privileged backgrounds, driven by topological substitution mechanisms whereby resilient learners strategically leverage internal resources—such as self-knowledge, metacognition, and study habits—to mediate and compensate for socioeconomic disadvantages.

5.1. Topological invariants of academic success

These new pieces of evidence confirm that resilience is not solely an individual trait but also deeply embedded within relational and structural contexts, suggesting that resilience should not be conceived as a stable personal attribute but rather as a dynamic and adaptive process emerging through the continuous interplay of individual, familial, and institutional factors [Wang, F., 2025]. Within this approach, it might be a deliberate topological reorganization where internal, controllable assets assume maximal functional necessity to maintain performance despite external disadvantage [Borsboom et al., 2021]. From this point of view, phenomenon moves beyond simple trait psychology, and it becomes an observable, structural efficacy of a student's resource network. While the privileged cohort relies on systemic redundancy and broadly distributed influence, the resilient cohort executes a strategy of forced efficiency, where success is

this sense, the resilient topology is primarily defined by the concentration of influence metrics, confirming that the resilient network holds a significantly higher average cumulative systemic reach, demonstrating that the resilient system's factors work harder and connect better on average. This efficiency might be achieved by elevating internal, controllable assets—such as Attentional control and Self-efficacy—to positions of maximum functional necessity, compelling them to act as powerful structural organizers that channel influence across the entire system. This concentrated efficiency also might be supported by Strategic resource integration by exhibiting remarkably low Modularity class counts, signifying a highly integrated and stable functional core that resists fragmentation. In this regard, it was found that compelling structural evidence for a universal functional core is necessary for high performance, irrespective of socioeconomic background, suggesting that factors such as Emotional Well-being (identified as the universal system bottleneck via Betweenness Centrality) and Cooperation (a primary structural bridge) must be highly organized and integrated in both the resilient and privileged groups [Newman, M. E. J., 2018]. Thus, while resources may differ wildly, the structural capacity to regulate emotion and engage collaboratively is the fundamental, non-negotiable topological prerequisite for maintaining any high-achieving academic trajectory, suggesting that no amount of privilege can buy freedom from the necessity of good emotional and relational plumbing.

5.2. Internalized control: cognitive factors as structural mediators of resilience

The most crucial finding resides in the stark contrast between the factors that define the structural integrity of each network, illustrating the core difference between redundancy and adaptation. In the privileged cohort, the secondary structural mediators are predominantly stable, external factors like School satisfaction and Sense of belonging, reflecting a system organized by general contentment and security. Conversely, the resilient cohort structurally substitutes these buffers by dramatically elevating internal, performance-oriented resources—specifically, Attentional control, Self-efficacy, and Metacognition—into high-betweenness positions. This disparity illustrates the mechanism of topological substitution: whereas the network of advantaged students maintains stability passively through established resources, the network of resilient students must actively mobilize controllable cognitive factors as structural linchpins, channelling resource flow through deliberate self-management strategies to compensate for fundamental deficits in socioeconomic and cultural capital [Borsboom et al., 2021].

The resilient student does not just work harder; their internal resources operate in a vastly different, more efficient structural role. High Clustering coefficients around critical factors confirm that skills are not isolated but are tightly reinforced by synergistic local habits, creating reliable internal feedback loops [Masuda et al., 2018]. In essence, the resilient academic trajectory is defined by topological substitution. In this sense, the system effectively replaces the structural security provided by external wealth with the overwhelming functional necessity of internal self-mastery. The result is a specialized, high-intensity structure that is more efficient, strategically organized, and adaptable, demonstrating that resilience is a quantifiable structural phenomenon of optimal organization imposed upon a limited resource budget [Borsboom et al., 2021].

The success mechanism of resilient students appears to involve the structural elevation of socio-emotional factors—particularly Emotional Regulation, Assertive Communication, and Growth Mindset—as compensatory assets within their network topology. Assertive Communication gains notable structural weight (high Katz Centrality) by serving as a key mechanism through which resilient students actively seek, secure, and leverage limited support from external agents such as teachers or peers. Emotional Regulation contributes by mitigating the volatility inherent in low-SES environments, helping to preserve the integrity of the internal resource system. Similarly, Growth Mindset, as a foundational belief, reinforces positive feedback loops within the network, sustaining the levels of Self-Efficacy necessary to maintain effort despite persistent adversity. These factors are not merely peripheral soft skills; rather, they function as an actively managed socio-emotional infrastructure that supports effective resource integration.

Despite the novel insights provided by complex network analysis into the structural basis of academic resilience, several methodological limitations warrant consideration for future work. First, the construction of the network relied on correlation matrices, which, by definition, capture co-occurrence rather than direct causal pathways. While the use of betweenness centrality provides strong evidence for functional control, longitudinal data is imperative to validate the directionality and sequence of influence among these socio-cognitive factors [Ye, W., et al 2021]. Second, the results are derived from a specific population subset (high-performing students in extreme socioeconomic quintiles), limiting the generalizability of the specific compensatory mechanisms identified to the broader population of all students facing adversity. Future research must, therefore, employ Dynamic Network Analysis (DNA) to track network evolution over time, confirming whether the observed topological structures—such as the high Closeness of Attentional Control—are stable adaptive strategies or transient responses to specific stressors. Finally, the role of specific contextual factors (like Bullying and Teacher support), which showed high volatility in their modular assignment, requires multilevel network modeling to properly nest student-level interactions within the school's structural environment. The necessity of strategic resource allocation remains a critical challenge for educational micro-planning and effective policy design.

References

[1] Aleph (2022). Documento técnico interno. Nuevo León Aprende: análisis psicométrico de los instrumentos de evaluación educativa. Agencia Latinoamericana de Evaluación y Política Pública. México.

[2] Azevedo, J. P. (2020). Learning poverty. *World bank policy research working paper, Washington DC: World Bank*.

[3] Barabási, A. L. (2003). Emergence of scaling in complex networks. *Handbook of Graphs and Networks: From the Genome to the Internet. Berlin: Wiley-VCH*.

[4] Borsboom, D., & Molenaar, D. (2015). Psychometrics. In J. Wright (Ed.), International encyclopedia of the social & behavioral sciences. Vol.19 pp.418–422.Amsterdam: Elsevier

[5] Bolt, T. S., Hampton, R. S., Furr, R. M., Fleeson, W., Laurienti, P. J., & Dagenbach, D. (2016). Integrating personality/character neuroscience with network analysis. In Neuroimaging personality, social cognition, and character (pp. 51-69). Academic Press.

[1]Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. Nature Reviews Methods Primers, 1(1), 1–23.
[2]Bracho, T. (2017), Índice de déficit en competencias ¿Avanzamos hacia la garantía del derecho a la educación? en Reformas y Políticas Educativas; No. 4, septiembre-diciembre 2017, FCE, México.
[3]Brass, D. J., & Borgatti, S. P. (Eds.). (2020). *Social networks at work*. New York: Routledge.
[4]Brinkhof, L. P., Ridderinkhof, K. R., de Wit, S., Krugers, H. J., & Murre, J. M. J. (2025). A cross-sectional network analysis of successful aging in a resilience-based framework. *PLoS ONE*, *20*(1), e0315445.
[5]Buser, T., Ahlskog, R., Johannesson, M., Koellinger, P., & Oskarsson, S. (2023). Using genes to explore the effects of cognitive and non-cognitive skills on education and labor market outcomes. Tinbergen Institute Discussion Paper, TI 2021-088/I. Retrieved from https://papers.tinbergen.nl/21088.pdf
[6]Davis, B., & Sumara, D. (2014). Complexity and education: Inquiries into learning, teaching, and research. Routledge.
[7]Deary, I. J., Cox, S. R., & Hill, W. D. (2021). Genetic variation, brain, and intelligence differences. Nature Reviews Genetics, 22(3), 162–173. https://doi.org/10.1038/s41380-021-01027-y
[8]Delprato, M. (2025). Determinants of Latin American students' academic resilience. Insights based on PISA 2022 using an explainable machine learning approach
[9]Hanushek, E. A., & Woessmann, L. (2020). The economics of international differences in educational achievement. Journal of Economic Perspectives, 34(4), 83–108.
[10] Hauke, J., & Kossowski, T. (2011). Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones geographicae*, *30*(2), 87-93.
[11] Johnson, J. H. (2016). Hypernetworks: multidimensional relationships in multilevel systems. The European Physical Journal Special Topics, 225(6), 1037-1052.
[12] Johnson, J. H., Fortune, J., & Bromley, J. (2017). Systems, networks, and policy. In Non-Equilibrium Social Science and Policy (pp. 111-134). Springer, Cham.
[13] Le Gall, J. F. (2022). Measure theory, probability, and stochastic processes (p. 22). Cham: Springer.
[14] Liu, Y. Y., & Barabási, A. L. (2016). Control principles of complex systems. *Reviews of Modern Physics*, *88*(3), 035006.
[15] Liu, C., Ding, J., Song, Y., & Li, Y. (2024, August). Tdnetgen: Empowering complex network resilience prediction with generative augmentation of topology and dynamics. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 1875-1886).
[16] Luthar, S. S., Cicchetti, D., & Becker, B. (2020). The construct of resilience: A critical evaluation and guidelines for future work. *Child Development*, *71*(3), 543–562.
[17] Malliaros, F. D., Rossi, M. E. G., & Vazirgiannis, M. (2016). Locating influential nodes in complex networks. *Scientific reports*, *6*(1), 19307.
[18] Marta, E. M., & Michelle, C. (2023). OECD Skills Studies Gender, Education and Skills. The Persistence of Gender Gaps in Education and Skills. OECD Publishing.
[19] Martin, A. J., & Marsh, H. W. (2006). Academic resilience and its psychological and educational correlates: A construct validity approach. *Psychology in the Schools*, *43*(3), 267-281.
[20] Masuda, N., Sakaki, M., Ezaki, T., & Watanabe, T. (2018). Clustering coefficients for correlation networks. Frontiers in Neuroinformatics, 12, 7.
[21] Masuda, N., Boyd, Z. M., Garlaschelli, D., & Mucha, P. J. (2024). Correlation networks: Interdisciplinary approaches beyond thresholding.

[1]Mishchuk, H., Oliinyk, O., & Bilan, Y. (2024). Brain gain and country's resilience: A dependency analysis exemplified by OECD countries. *Equilibrium (1689-765X)*, *19*(2).
[2]Moll, K., Göbel, S. M., Gooch, D., Landerl, K., & Snowling, M. J. (2016). Cognitive risk factors for specific learning disorder: processing speed, temporal processing, and working memory. Journal of learning disabilities, 49(3), 272-281.
[3]Newman, M. E. J. (2016). Community detection in networks: Modularity optimization and maximum likelihood are equivalent. Physical Review E, 94(5), 052315. https://doi.org/10.1103/PhysRevE.94.052315
[4]Newman, M. E. J. (2018). Networks (2nd ed.). Oxford University Press.
[5]NLA (2025). *Componentes técnicos de la evaluación.* Nuevo León Aprende. Coordinación del ecosistema educativo: "Legado Nuevo León". Nuevo León, México.
[6]OECD (2021). *Education Policy Outlook 2021: Shaping Responsive and Resilient Education in a Changing World*. OECD Publishing. https://doi.org/10.1787/75e40a16-en
[7]OECD (2025), *Education at a Glance 2025: OECD Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/1c0d9c79-en
[8]Oldham, S., Fulcher, B., Parkes, L., Arnatkevičiūtė, A., Suo, C., & Fornito, A. (2019). Consistency and differences between centrality measures across distinct classes of networks. PLoS ONE, 14(7), e0220061.
[9]Plomin, R., & von Stumm, S. (2018). The new genetics of intelligence. Nature Reviews Genetics, 19(3), 148–159. https://doi.org/10.1038/nrg.2017.104 [kclpure.kcl.ac.uk]
[10] Rudd, K. L., & Yates, T. M. (2020). The Health Implications of Resilience. The Wiley Encyclopedia of Health Psychology, 269-277.
[11] Rudd, G., Meissel, K., & Meyer, F. (2021). Measuring academic resilience in quantitative research: A systematic review of the literature. *Educational Research Review, 34*, 100402. https://doi.org/10.1016/j.edurev.2021.100402
[12] Rudd, R. E., Wingo, J., & Bynum, L. (2021). Resilience: Adaptive and proactive coping in the process of developing resilience. In R. E. Rudd, J. Wingo, & L. Bynum (Eds.), *Research in occupational stress and well being* (Vol. 19, pp. 19–46). Emerald Publ. Limited.
[13] Rutter, M. (2012). Resilience as a dynamic concept. *Development and Psychopathology*, *24*(2), 335–344.
[14] Sánchez-Restrepo, H., & Louçã, J. (2019). Topological Properties of Inequality and Deprivation in an Educational System: Unveiling the Key-Drivers Through Complex Network Analysis. In International Conference on Human Systems Engineering and Design: Future Trends and Applications (pp. 469-475). Springer, Cham.
[15] Sánchez-Restrepo, H. S., & Carmona-Soto, M. B. (2023). Aprendizajes imprescindibles en la educación básica: Avances y desafíos durante 2023. En: Secretaría de Educación de Nuevo León, México. https://doi.org/10.5281/zenodo.15205761
[16] Sánchez-Restrepo, H., Carmona-Soto, M. B., Morales-Garza, S. & Almeida-Díaz, L. G. (2025). Cuatro ciclos escolares mejorando los resultados de aprendizaje. Evolución de los aprendizajes de los estudiantes de educación básica. Informe estatal 2022 - 2025. Nuevo León Aprende. Coordinación del ecosistema educativo: "Legado Nuevo León". Nuevo León, México.
[17] Saqr, M., Dever, D., López-Pernas, S., Gernigon, C., Marchand, G., & Kaplan, A. (2024). *Complex dynamic systems in education: Beyond the static, the linear and the causal reductionism*. In M. Saqr & S. López-Pernas (Eds.), *Advanced Learning Analytics Methods: AI, Precision and Complexity* (Chap. 12). Springer. [Forthcoming].
[18] Senior, C., & Sahlberg, P. (2025). The evolution of the OECD's position on equity in global education. *International Journal of Educational Development*, *114*, 103241.

[1]Serrano, M. A., Boguná, M., Pastor-Satorras, R., & Vespignani, A. (2007). *Correlations in complex networks*. Large scale structure and dynamics of complex networks: From information technology to finance and natural sciences, 35-66.
[2]Schober, P., Boer C, Schwarte LA. (2018). Correlation Coefficients: Appropriate Use and Interpretation. Anesth Analg. 2018 May;126(5):1763-1768.
[3]Siew, C. S. Q. (2020). Applications of Network Science to Education Research: Quantifying Knowledge and the Development of Expertise through Network Analysis. *Education Sciences, 10*(4), 101. https://doi.org/10.3390/educsci10040101
[4]Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*(3), 417–453.
[5]Sykes, C., Baştuğ, S. I., & Fiasconaro, S. (2025). A metareview of research on educational inequality and socioeconomic disadvantage. *Education Sciences, 15*(6), 740. https://doi.org/10.3390/educsci15060740
[6]Tonegawa, Y. (2023). Education in SDGs: What is inclusive and equitable quality education? In Sustainable Development Disciplines for Humanity (pp. 55–70). Springer. UNESCO. (n.d.). SDG 4 – Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all. Motl, T. C., Multon, K. D., & Zhao, F. (2018). Persistence at a tribal university: Factors associated with second year enrollment. Journal of Diversity in Higher Educ, 11(1), 51.
[7]Tracey, B., & Florian, K. (Eds.). (2016). Educational Research and Innovation Governing Education in a Complex World. OECD Publishing.
[8]UNESCO Institute for Statistics (UIS). (2024). *World Education Statistics, 2024*. Paris: UNESCO.
[9]UNESCO (2015). Education 2030: Incheon Declaration and Framework for Action Towards Inclusive and Equitable Quality Education and Lifelong Learning for All.
[10] UNESCO. (2024). Monitoring SDG 4: 2024/5 Global Education Monitoring Report. Paris: UNESCO.
[11] van der Maas, H., Kan, K., MarSSan, M., & Stevenson, C. (2017). Network models for cognitive development and intelligence. Journal of Intelligence, 5(2), 1–17. doi: 10.3390/jintelligence5020016
[12] Wang, F. (2025). Fostering resilience among university students: The role of resilience as a dynamic process of positive adaptation when facing adversity. *Higher Education Research & Development*. https://doi.org/10.1007/s10734-025-01484-2
[13] World Bank, UNESCO, & UNICEF. (2021). *The state of the global education crisis: A path to recovery.* Washington, DC: World Bank.
[14] World Bank, UNESCO, & UNICEF. (2022). *The State of Global Learning Poverty: 2022 Update*. Washington, DC: World Bank.
[15] Xenofontos, C., & Mouroutsou, S. (2023). Resilience in mathematics education research: A systematic review of empirical studies. *Scandinavian Journal of Educational Research*, *67*(7), 1041-1055.
[16] Yan, Y., & Gai, X. (2022). High achievers from low family socioeconomic status families: Protective factors for academically resilient students. *International Journal of Environmental Research and Public Health, 19*(23), 15882. https://doi.org/10.3390/ijerph192315882
[17] Ye, W., Strietholt, R., & Blömeke, S. (2021). Academic resilience: Underlying norms and validity of definitions. Educational Assessment, Evaluation and Account, 33(1), 169-202.
[18] Zhang, L., & Peixoto, T. P. (2020). Statistical inference of assortative community structures. Physical Review Research, 2(4), 043271.
[19] Zhu, J., Dai, F., Zhao, F., & Guo, W. (2023). Integrating node importance and network topological properties for link prediction in complex network. Symmetry, 15(8), 1492.

CHAPTER 8

Topological Patterns of Academic Success: Insights from Complex Network Analysis

First Autor: Harvey Spencer Sánchez Restrepo
Harvey.Restrepo@iscte-iul.pt
Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL
1649-026, Lisbon, Portugal

Coauthor: Jorge Louçã
jorge.l@iscte-iul.pt
Information Sciences, Technologies and Architecture Research Center (ISTA), ISCTE-IUL
1649-026, Lisbon, Portugal

Running head: **Topological Patterns of Academic Success**

Article accepted and already in press in: *Proceedings of the 6th World Conference on Complex Systems*. **Springer Nature**.

Topological Patterns of Academic Success: Insights from Complex Network Analysis

Harvey Spencer Sánchez-Restrepo [1[0000-0001-7311-8622]], Jorge Louçã [2[0000-0003-4766-5627]]

Martha Belén Carmona-Soto [3[0000−0001−7834−6608]], Sofialeticia Morales Garza [4]

[1,2] Information Sciences, Technologies and Architecture Research Center. ISCTE-IUL, 1649-013, Lisbon, Portugal
Harvey.Restrepo@iscte-iul.pt,jorge.l@iscte-iul.pt
[3] Research Center for Psychological Science. Universidade de Lisboa, 1649-026, Lisbon, Portugal
martha.belen@edu.ulisboa.pt
[4] Coordinación del Ecosistema Educativo Legado Nuevo León. 64000. Monterrey, México
sofialeticia.morales@nuevoleon.gob.mx

Abstract. One of the biggest challenges in education policy is understanding how some students manage to avoid learning deprivation despite persistent adversity. This ability—often described as academic resilience—tends to rely on foundational socio-cognitive and socioemotional skills, such as cooperation, self-efficacy, metacognition, and emotional regulation. This study conducted a secondary analysis of an official large-scale assessment database to systematically examine the structural interactions among socio-cognitive, socioemotional, and contextual factors and to develop a network-based conceptualisation of resilience. By overcoming traditional approaches that isolate variance in single predictors to estimate covariance between factors and learning outcomes, this framework models the entire multivariate resource ecosystem, enabling the identification of specific topological mechanisms—rather than simple bivariate associations—that differentiate students who overcome socioeconomic disadvantage from their privileged, high-performing peers. Our findings constitute a substantive shift from linear models, specifying how structural reorganization of individual and institutional factors can compensate for the impact of familial, cultural, and social sources of scarcity. Results indicate that high levels of achievement, irrespective of socioeconomic status, are intrinsically mediated by network features and specific factors that serve as bridges to success. This discovery highlights deep structural differences between deprived and privileged students, showing that resilient learners strategically develop topological substitution mechanisms to leverage internal resources—such as self-knowledge, metacognition, and study habits—to mediate and compensate for socioeconomic disadvantages.

Keywords: Resilience, Complex Network, Socioemotional skills, Large-scale assessments, Educational Policy.

1 Introduction

The 2030 Agenda's Sustainable Development Goal 4 (SDG 4) commits 193 countries to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" (UNESCO, 2015). Every government recognises that Education is not only a fundamental human right but also a key generator of better opportunities for development and prosperity—reducing poverty, advancing health, promoting gender equality, and strengthening peace and stability [4][5][6][7]. Despite this general agreement, most education systems persistently fail to design evidence-based policies to provide the foundational knowledge and competencies that every student requires to thrive academically [8][9][10][11]. A key strategy to begin avoiding this management error is based developing valid and reliable data on educational outcomes based on Large-Scale Assessments (LSAs) [12]. For increasing its power and capacity to explain complex phenomena, these LSAs are usually accompanied by a pool of surveys that investigate a large number of academic and non-academic aspects of the students, including some factors and opinions about families, teachers and even the principal of their schools, commonly referred to as Factors Associated with Learning (FAL), which use to provide deeper insights into the conditions that influence student performance, satisfaction and engagement with both learning processes and outcomes [13].

The primary objective of LSAs is to furnish high-quality, quantitative evidence to measure educational dimensions for detecting and measuring statistical associations between FAL and learning outcomes across different time scales and territorial domains [14][15]. Historically, most of the analytical tools for estimating such relationships are based in correlations and the main models come from neoclassical economics, which have become the dominant paradigm for studying educational phenomena, postulating that simple linear covariations between learning outcomes and individual factors can account for educational gaps [14][18]. However, the methodological efficacy of LSAs in informing policy has come under increasing scrutiny [15][16][18]. This traditional approach is increasingly seen as valuable but insufficient [16][19][20]. Main arguments show that the explained variance used to be proportional to the number of directly correlated factors, and that this approach relegates factor interactions to the margins, excluding them from central discussion [16], as well as its near to exclusive reliance on linear regression models use to fragment the interrelated educational ecosystem into artificial units far from explaining academic failure or success. Consequently, it fails to capture complex educational phenomena [21].

However, these simplified models often overlook how inequality in learning outcomes relates to socioeconomic disadvantages and the resulting social, cultural, and emotional deprivation [20][22]. Linear models also struggle to capture these dynamics directly. This methodological limitation—the inability to model complexity, nonlinearity, and interactions—is likely one of the main reasons for the lack of explanatory power in most models and it may also account for the correspondingly low levels of accountability [19].

1.1 Learning Poverty

The construct Learning Poverty (LP) refers to "The ability to read with comprehension [as] a foundational skill that every education system around the world strives to impart by late in primary school […] as it is essential for all future learning and development" [11]. The indicator is a composite measure combining the proportion of primary-aged children who are out of school (Schooling Deprived, SD) with the proportion of students in school who perform below minimum proficiency levels in reading (Learning Deprived, LD). In the context of SDG 4, according to the Global Alliance to Monitor Learning (GAML), learning deprivation is defined as "the share of children at the end of primary school who read at below the minimum proficiency level" [23]. This combined measure manages to capture in a parsimonious way the double imperative of any educational policy: to ensure "greater schooling" (access) and to guarantee "better learning" (quality), providing the opportunity to evaluate to what extent the time that students spend in classrooms translates into essential skills and abilities.

Current global estimates for the 2019–2025 period indicate that nearly 60% of all 10-year-olds worldwide—equivalent to more than 390 million children—experience learning poverty [10]. As with most social indicators, learning poverty is considerably deeper and more widespread in low-and-middle-income countries, where more than 70% of children lack basic reading comprehension skills—a situation exacerbated during the COVID-19 pandemic and its impact on prolonged school closures, as well as the near total absence of educational services throughout that period [6][7][11]. Thus, in Latin America and the Caribbean, learning poverty again exceeds 79% of the population, erasing the modest improvements observed during two decades of sustained educational progress. In Mexico, the most recent data from [6] estimate that around 77% of children cannot read or understand a simple text.

1.2 Academic Resilience

Academic performance is widely regarded as a primary indicator of educational success and has been extensively used in the academic resilience literature [24][25][26][27]. In educational research, the Socioeconomic Status (SES) of the students' family has been identified as one of the most significant factors affecting learning and educational experiences. According to [28], approximately 15% of the variance in academic achievement can be explained by family SES. Low SES is often associated with challenging learning conditions and reduced motivation, which contribute to lower academic achievement. Students achieving better-than-expected educational outcomes were once defined as high-level academic performers [27][29]. To a greater or lesser extent, students from disadvantaged socioeconomic backgrounds in all countries persistently face structural, social, cultural, and pedagogical barriers that restrict access to education, reflected in lower learning outcomes and lower expectations of study [16][30]. However, this trajectory, which accumulates the effects of poverty in an almost deterministic way, is interrupted with some regularity by a small but statistically significant subset of students from highly disadvantaged socioeconomic backgrounds, often described as academically resilient (hereafter: ARL), who

manage to achieve high levels of academic success despite facing severe social and economic adversities [31][32]. Nevertheless, much of the existing literature continues to rely on linear modelling approaches that obscure the systemic interdependencies underlying the resilience processes.

The concept of resilience appears to have its conceptual roots in social psychology, emerging as an explanation for why some people achieve positive outcomes in their development of emotional or social stability, despite continuous exposure to significant stress or adversity [24][26][31][37]. In its earliest formulations, resilience is described as a stable personal trait, a characteristic possessed by specific individuals who achieve good performance under pressure [1][24]. In quantitative research, academic resilience can also be understood as "the greater likelihood of success in school and other life achievements, despite environmental adversities caused by early traits, conditions, and experiences" [27]. The previous definition also implies that resilient students possess a set of abilities to recover from adversity and maintain a commitment to learning despite challenging circumstances, which most likely integrates individual qualities, fostered by family and society, such as persistence, adaptability, and self-regulation [26]. In some cases, other socioemotional skills have been identified in research conducted in highly complex educational contexts, which has revealed that students' resilience strategies often emerge through leveraging social support, recognising institutional possibilities, and reframing identity narratives around persistence [26].

One of the most common approaches to measuring resilience involves asking participants to complete self-report questionnaires about traits often associated with resilience (also known as the latent-construct approach). One strategy for detecting academic resilience is to use a survey to capture students' perceptions of their ability to overcome educational challenges [24][25]. Another approach, generally much more objective, uses two predetermined criteria to measure academic resilience (also known as the definition-driven approach (DDA) [31]). Its main criterion is the presence of chronic adversity that hinders students' academic success. The DDA integrates students' relative positions in family SES and academic achievement relative to their peers, a measure widely used in previous studies [26][27][31][38]. Common analyses of educational outcomes at country-level uses a DDA to define and operationalize the concept of resilience [36]. First, it identifies students who belong to the lowest segment of the SES distribution, then selects those who achieve academic results within the highest performance percentiles in their respective country or region, thereby establishing a high-contrast group.

From a global assessment perspective, this methodological framework substantially enhances cross-national comparability by standardizing resilience metrics across heterogeneous educational contexts. In general, this approach is considered particularly rigorous [25] because it is based on standardized learning indicators measured externally by an LSA, rather than basing selection on metrics based on self-reported information, further reducing the proclivity to incorporate selection bias due to socially desirable responses or other cultural dimensions, which in turn contributes to a more robust and complete understanding of the dimensions associated with structural resilience within educational systems.

2 Methods

2.1 Material

This study draws on official datasets from the state of Nuevo León, Mexico, collected as the core of the official LSA at the end of the cycle 2025 [39]. The official dataset NLA25-D is derived from a census-based process that encompasses 83 800 ninth-grade students' records from 1 066 (97%) schools, both public and private, ensuring comprehensive coverage of all municipalities and all strata [39]. This dataset also includes high-quality linked microdata, such as responses from the associated multivariate factors survey and indexes for students' abilities in Literacy, Numeracy and Science. According to official reports, the LSA adheres to international standards to validate the scales, establish construct validity, ensure measurement invariance over time, and achieve high reliability coefficients $(\alpha, \omega > 0.80)$ [40]. The estimation of skills and other psychometric analyses are carried out using Item Response Theory (IRT), specifically, a two-parameter logistic model (2PL), where higher learning levels correspond to higher ability estimates $\theta_i \in (-\infty, \infty)$ [40]. To identify low-level learners (LLL) and high-level learners (HLL), all abilities $\{\theta_i\}$ were classified into m achievement levels (L_m) throughout a Bookmarking process [40]. Students performing at Level 1 (L_1) exhibit significant learning deficits that hinder continued educational growth, situating them in conditions of learning deprivation and, potentially, social vulnerability. Conversely, students performing at Level 3 (L_3) belong to the highest level, corresponding to consolidated knowledge that enables the acquisition of more complex competencies.

To contextualize student achievement within broader societal differences, and to address critical questions regarding educational opportunity and inequalities in learning outcomes, this LSA incorporates a composite index, identified as the Index of Economic, Social, and Cultural Status (ISEC), estimated using a standard structural equation model [LLECE-UNESCO, 2024] based on three key components: (1) parental educational attainment and occupational status; (2) home possessions (as and book count); and (3) home services (including dwelling materials, access to sanitation, waste collection, internet and telephony). The estimates of the ISEC score and its corresponding socioeconomic status (SES) deciles $(D_j), j = 1, \ldots, 10$ are directly employed in this study. School weights were applied to correct sampling errors and enhance the representativeness of the findings. Complete methodological documentation and raw datasets are available through the official webpage [39]. Data analyses were conducted using R (version 4.4.3), with the 'corrplot' package employed to estimate and visualise the correlation matrix. Network analysis was performed using Gephi (version 0.10). Statistical inference followed the algorithms proposed by [42], while betweenness centrality based on correlation adjacency matrices was computed using the approach described by [2]. This database offers a unique opportunity to examine associations among a wide range of variables at the end of the lower secondary school in Mexico.

2.2 Learning deprivation and socioeconomic status

Analysis of learning deprivation across socioeconomic deciles revealed a pronounced and systematic gradient for both women and men. The central role of socioeconomic status in learning deprivation rates for both population groups demonstrates the effects of the second force mentioned earlier. Another relevant finding is the persistent gender gap, which shows a continued disadvantage for men in terms of their learning outcomes, a disadvantage that increases in the middle deciles. Table 1 presents descriptive statistics for these indicators of deprivation. At the student level, the Index of Socioeconomic Context (ISEC) was negatively correlated with absolute learning deprivation for both males (r = —0.233, $p < 0.001$) and females (r = —0.262, $p < 0.001$). Among students already deprived, ISEC also showed a negative correlation with the intensity of deprivation (r = —0.125 for males; r = —0.135 for females; $p < 0.001$). This monotonic inverse relationship indicates that educational risk peaks at the lowest end of the socioeconomic spectrum. To estimate the trend and monotony of the RAD with respect to the SES, it is treated as an ordinal index (1-10). Spearman's correlation confirmed the strength of these trends (ρ = —0.96 for men and ρ= —0.98 for women, p<0.001), providing a solid basis for measuring the gaps between favoured and disadvantaged students as an indicator of systemic failures to ensure essential learning.

For instance, among men, RAD decreased from 0.689 in the poorest decile (D01) to 0.304 in the richest decile (D10), representing a 55.9% reduction. For women, the decrease was even more pronounced, from 0.673 to 0.235 (—65.1%). The regression slopes estimate that each upward decile change is associated with a decrease of approximately 0.040 in RAD for men and 0.044 for women. Consistent with other findings [43], the estimates shown in Table 1 confirm that gender acts as a secondary but persistent axis of inequality. Regarding the learning deprivation dimension, men consistently exhibit higher rates. However, the absolute gap is considerably smaller in D01 (0.016) and increases toward the middle deciles, peaking in D04 (0.088). On the other hand, the male-to-female ratio increased from 1.02 in D01 to 1.29 in D10, suggesting that improvements in socioeconomic status benefit both sexes, but that the male disadvantage persists and is accentuated in higher socioeconomic deciles.

Table 1. Rates of Absolute Deprivation for SES deciles and gender.

SES	Male		Female	
	RAD	SE	RAD	SE
D01	0.689	0.008	0.673	0.007
D02	0.658	0.008	0.628	0.007
D03	0.639	0.007	0.588	0.008
D04	0.628	0.007	0.54	0.008
D05	0.579	0.008	0.529	0.008
D06	0.57	0.008	0.49	0.008
D07	0.541	0.008	0.456	0.008
D08	0.462	0.008	0.401	0.008
D09	0.387	0.007	0.323	0.007
D10	0.304	0.007	0.235	0.007

2.3 Resilient and Privileged Networks

After confirming the relationship between RAD and SES and that the indices behave with statistical regularity, it is possible to use DDA to analyse the topology of the networks of factors associated with learning. Thus, the HLL(L_3) dataset can be built by selecting students from the two lowest and two highest socioeconomic deciles. Subsequently, students within L_3 can be classified into two groups: privileged $\left(L_3^{Q5}\right)$ and resilient $\left(L_3^{Q1}\right)$. From the initial population of 83 600 students, 16 760 (20%) were selected from 859 schools (80.6%), with female students representing 53.7% of the group. Among these low-SES students, 2 673 (16.1%) were identified as ARL, coming from 603 (56.6%) schools, where 55.2% of them are females. To avoid bias from potential exceptional conditions associated with private schools, 113 students (0.7%) were excluded from the analysis.

To estimate the entries in the adjacency matrix **R**, we used Spearman's rank correlation, a standard nonparametric method that provides estimates based on observations to make a ranking on ordinal variables rather than directly comparing data values by assuming hypotheses about their statistical distribution [44]. This way of identifying structured dependencies in complex datasets seems to be more accurate in capturing the monotonic association between variables—reflecting consistent directional changes, whether increasing or decreasing, across attributes [45][46], a property especially suitable for exploring nonlinear or irregular patterns that still preserve ordered relationships. Before the first iteration, the matrix for the whole dataset had 114 variables and had 6,555 entries, including autocorrelations. As expected, the statistical significance of the covariances among factors varies across subpopulations, which helps explain why it is highly informative about the underlying structure of both resilient and privileged students.

To analyse both groups, resilient $\mathbf{R}\left(L_3^{Q1}\right)$ and privileged $\mathbf{R}\left(L_3^{Q5}\right)$ matrices were calculated; this process is equivalent to controlling for achievement level on Q1 and Q5 and then grouping by SES to have the contrast. This kind of segmentation is highly informative about the group's underlying structure, as the statistical significance of the covariances among factors depends on the hidden structure of each dataset [46]. At the end of this phase, for removing noise and excess of misinformation, both matrices $\mathbf{R}\left(L_3^{Q1}\right)$ and $\mathbf{R}\left(L_3^{Q5}\right)$ were reduced for being just positive defined matrices. To ensure that each correlation participating in the network analysis is relevant and maintains the same directionality of the system, only those correlations between factors satisfying the condition $\rho_{i,j} > 0.2$ $(p < 0.001)$ were considered. The final adjacency matrices are $\mathbf{R}^{+}\left(L_3^{Q1}\right)$ and $\mathbf{R}^{+}\left(L_3^{Q5}\right)$ had 967 (14.75%) and 783 (11.95%) positive entries, respectively. (see Fig. 1).

3 Results

3.1 Asymmetries in network parameters

For the comparative analysis, in both networks Q1 (Fig. 1) and Q5 (Fig. 2), the network topologies were systematically configured to visualise structural functionality, and the same algorithms and parameters were implemented, using a fixed resolution of

1.0 and edge weighting via a randomised process. The layout was built throughout Yifan Hu's proportional algorithm [47]. Node colour reflects the Modularity Class, which delineates densely interconnected communities and provides structural evidence of resource ecology—indicating whether the system operates as a highly integrated (synergistic) network or exhibits a fragmented configuration. Regarding node size, it is scaled by the Betweenness Centrality coefficient for revealing hierarchical patterns of structural control within the system. For node label sizes, Katz centrality was used to rank how far and wide a node's influence extends, a handy metric that gives a stronger force to direct connections, holding indirect connections' contribution to influence as attenuated degree functional reinforcement among immediate neighbours [48]. Results for Modularity on both networks find three communities for Q1 and five for Q5 with $M_{Q1} = 0.236$ and $M_{Q5} = 0.236$. In our research, this parameter is identified with densely connected communities of factors, providing structural evidence on whether the resilient group exhibits a highly integrated resource ecology (low modularity, typically indicating that resources operate synergistically) or a fragmented configuration [48].

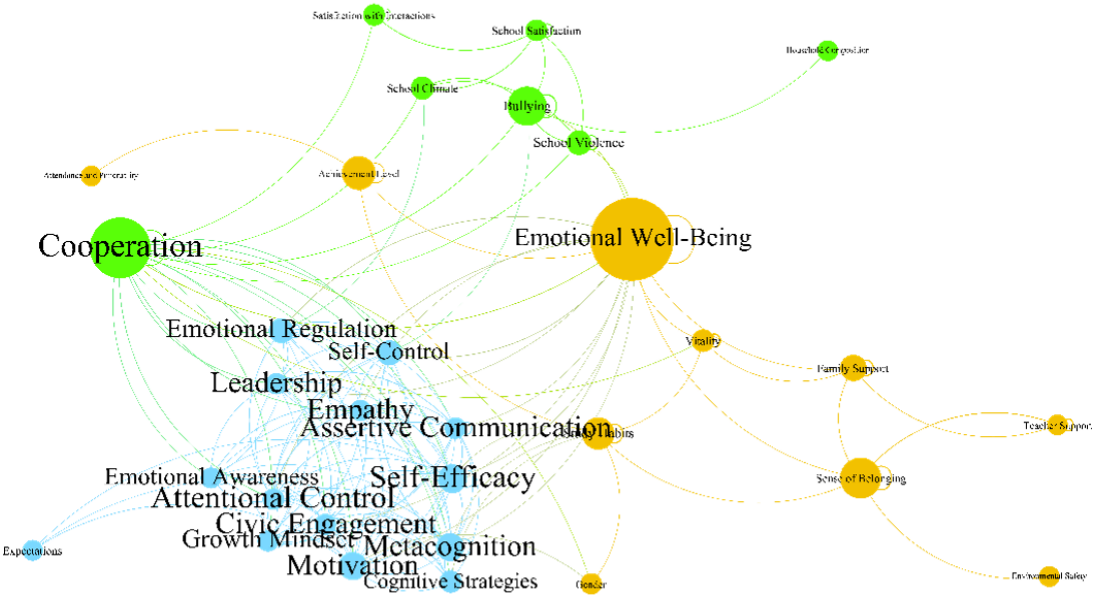


Fig. 1. Factor's Network for Resilient Students showing Cooperation and Emotional Well-Being as central factors for high-achievement levels.

As a parallel metric, the Clustering coefficient offers a local measure of network density around each node, signalling the presence of tightly knit functional reinforcement among immediate neighbours. When we put these metrics together, we might obtain quantitative information to advance from mere correlational observations to structural inferences regarding the effectiveness of distinct academic pathways.

3.2 Topological substitution for Success

The comparative assessment of network topological attributes reveals a compensatory structural mechanism employed by resilient students. According to the parameters in Table 2, the privileged network shows low influence variability and moderate cluster-

ing, suggesting that its high performance is based on the stability of diffuse and redundant resources. On the one hand, the resilient configuration succeeds by optimising structural requirements. In this sense, the HP-Q1 network exhibits superior structural efficiency, for instance, Katz centrality: $K_{Q1} = 2.86$ versus $K_{Q5} = 2.02$), with the factor *Cooperation* exerting the most significant overall reach, substantially exceeding that of the Q5 network (5.04 versus 3.25). This outcome is likely due to the structural positioning of internal factors: core socio-cognitive skills such as *Attentional Control*, *Self-Efficacy*, and *Metacognition* remain highly stable and are functionally integrated within the central module—reflected by low modularity class counts—thereby forming a cohesive and reliable resource ecosystem [48][49]. Furthermore, high Eigenvector Centrality scores for Attentional Control (Q1: 4.73 vs. Q5: 2.94) and Self-Efficacy (Q1: 4.73 vs. Q5: 2.76) confirm that resilient students leverage the highest-quality network connections, exhibiting amplified relational influence and local cohesion, in contrast to the systemic security observed in the privileged network (Clustering coefficient ≈0.87 for key factors). This integration forms the structural signature of effective self-regulated learning (SRL) in adverse contexts. In this sense, it is necessary to highlight that emotional well-being seems to be the universal gatekeeper (see Table 2), but without it, nothing flows; no matter the background, if emotional stability is missing in the system, the entire resource network stops functioning [50]. On the other hand, resilient students reveal a remarkable capacity for structural adaptation, probably because, lacking redundancy in the relevant factors that would allow them to achieve the high performance of privileged students, they require a compensatory optimization mechanism when they have access to any of them through what [52] describe as topological substitution: repositioning internal strengths—such as metacognition and study habits—into pivotal bridging roles that sustain network connectivity [41].

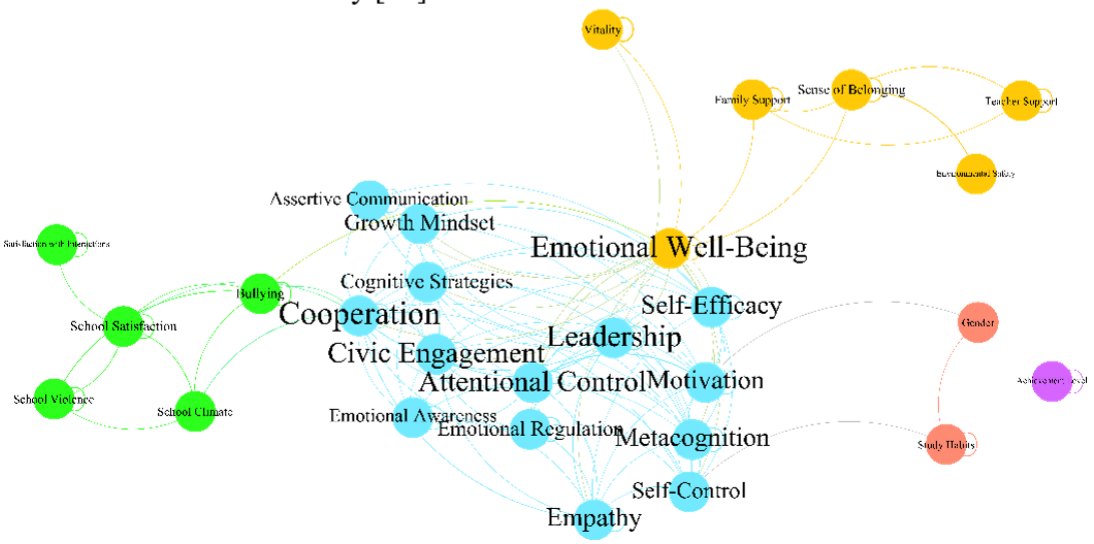


Fig. 2. Factor's Network for Privileged students showing how Cooperation, Leadership and Emotional Well- Being are the central factors on their structure.

Table 2. Parameter Academic Factor Networks.

Factor	Closeness centrality		Harmonic closeness centrality		Betweenness centrality		Katz Centrality		Weighted Degree		Eigen centrality		Modularity class	
	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5
Emot. Well-Being	0.70	0.72	0.78	0.82	0.32	0.33	4.11	3.06	4.74	4.65	0.84	0.97	1	3
Cooperation	0.70	0.70	0.82	0.81	0.21	0.23	5.04	3.25	5.87	4.92	1.00	1.00	2	2
Sense of Belonging	0.48	0.46	0.54	0.53	0.10	0.10	1.67	1.56	2.29	2.20	0.10	0.11	1	3
Bullying	0.52	0.51	0.59	0.57	0.09	0.06	2.21	1.68	2.85	2.15	0.24	0.22	2	4
Achiev. Level	0.45	0.00	0.50	0.00	0.07	0.00	1.39	1.06	1.30	0.58	0.09	0.00	1	0
Study Habits	0.49	0.38	0.58	0.43	0.06	0.00	2.02	1.24	2.55	1.18	0.23	0.07	1	1
Self-Efficacy	0.61	0.59	0.74	0.71	0.04	0.01	4.73	2.76	5.04	3.66	0.95	0.86	0	2
Motivation	0.61	0.58	0.74	0.69	0.04	0.07	4.36	2.62	4.54	3.41	0.95	0.75	0	2
Metacognition	0.60	0.59	0.72	0.71	0.04	0.01	4.50	2.79	4.82	3.78	0.89	0.87	0	2
Family Support	0.45	0.46	0.51	0.51	0.03	0.03	1.63	1.44	2.22	1.80	0.09	0.11	1	3
Self-Control	0.60	0.57	0.70	0.67	0.02	0.07	3.79	2.27	3.25	2.35	0.79	0.63	0	2
Emotional Reg.	0.59	0.52	0.69	0.60	0.02	0.00	3.91	2.23	4.00	2.72	0.83	0.55	0	2
School Violence	0.50	0.36	0.56	0.42	0.02	0.00	1.91	1.38	2.29	1.70	0.19	0.05	2	4
School Climate	0.45	0.46	0.53	0.53	0.01	0.01	1.73	1.42	1.28	1.02	0.17	0.12	2	4
Vitality	0.49	0.44	0.54	0.48	0.01	0.00	1.79	1.48	2.21	1.83	0.17	0.17	1	3
Cognitive Strat	0.54	0.54	0.63	0.64	0.01	0.00	3.16	2.14	2.50	2.06	0.61	0.63	0	2
Satisfaction	0.43	0.32	0.48	0.36	0.01	0.00	1.39	1.14	1.07	0.80	0.08	0.02	2	4
School Satisfaction	0.37	0.46	0.44	0.54	0.00	0.08	1.57	1.56	2.15	1.98	0.06	0.14	2	4
Attentional Control	0.59	0.59	0.71	0.71	0.00	0.01	4.73	2.94	4.76	3.99	0.92	0.86	0	2
Civic Engagement	0.59	0.59	0.71	0.72	0.00	0.01	4.52	2.84	4.61	3.84	0.92	0.90	0	2
Empathy	0.59	0.58	0.71	0.69	0.00	0.01	4.49	2.70	4.44	3.52	0.92	0.80	0	2
Assertive Comm.	0.59	0.50	0.71	0.58	0.00	0.00	4.31	1.97	4.23	2.18	0.92	0.49	0	2
Leadership	0.58	0.60	0.69	0.74	0.00	0.04	4.23	3.11	4.17	4.56	0.88	0.91	0	2
Gender	0.48	0.38	0.52	0.44	0.00	0.00	1.52	1.17	0.68	0.45	0.16	0.07	1	1
Growth Mindset	0.52	0.55	0.66	0.66	0.00	0.00	4.19	2.38	3.99	2.79	0.86	0.74	0	2
Emotional Aware	0.50	0.47	0.63	0.58	0.00	0.00	3.63	2.03	3.39	2.26	0.76	0.54	0	2
Household	0.34		0.38		0.00		1.10		0.22		0.02		2	
Teacher Support	0.33	0.33	0.38	0.37	0.00	0.00	1.34	1.31	1.66	1.64	0.02	0.03	1	3
Environmental Safe	0.33	0.32	0.36	0.35	0.00	0.00	1.08	1.12	0.24	0.40	0.01	0.01	1	3
Attendance	0.32		0.34		0.00		1.06		0.22		0.01		1	
Expectations	0.40		0.46		0.00		1.64		0.71		0.20		0.00	
μ	0.51	0.49	0.60	0.57	0.04	0.04	2.90	2.02	2.92	2.44	0.49	0.45	0.80	2.43
σ	0.10	0.14	0.13	0.17	0.07	0.08	1.43	0.71	1.63	1.30	0.40	0.37	0.81	1.00
Max	0.70	0.72	0.82	0.82	0.32	0.33	5.04	3.25	5.87	4.92	1.00	1.00	2.00	4.00
Min	0.32	0.00	0.34	0.00	0.00	0.00	1.06	1.06	0.22	0.40	0.01	0.00	0.00	0.00

The comparison of centrality metrics across the resilient (Q1) and privileged (Q5) high-performing networks suggests a fundamental divergence in structural organization. The data in Table 2 demonstrate that *Emotional Well-Being* acts as the single most critical bridge (Q1: 0.32, Q5: 0.33). This affirms its role as the primary structural prerequisite for mediating resource flow regardless of socioeconomic status [51]. The network parameters reveal further contrasts that demonstrate topological substitution mechanisms by forcing the mediation of key internal assets. For instance, the students' ability to predict their Achievement Level (Q1: 0.07 vs. Q5: 0.00), a highly relevant metacognitive process that highly correlates with *Study Habits*, which demonstrate their relevance through substantial *Betweenness Centrality* for Q1 while in Q5 they remain completely peripheral and practically irrelevant (Q1: 0.06 vs. Q5: 0.00). This supports the idea that the Q1 network must structurally rely on internal effort and outcomes to bridge and sustain network integrity, a functional necessity absent in Q5. Furthermore, active self-management factors like *Emotional Regulation* (Q1: 0.02 vs. Q5: 0.00) and *Vitality* (Q1: 0.01 vs. Q5: 0.00) become active mediators only in the resilient network, distinguishing it from the privileged structure, which tolerates more resource isolation.

4 Discussion

Our main finding is the marked contrasts in structural integrity between the network of privileged and resilient students. This contrast helps distinguish how inequality manifests in each group. Characterising these properties via modularity and centrality is not trivial. It requires distinguishing students by family SES—groups that differ in psychological and behavioural traits reflected in learning—while also identifying differences among high-achieving students. This sheds light on essential processes of redundancy and adaptation documented in conventional educational frameworks [1][2]. For instance, in the privileged cohort, the secondary structural mediators are predominantly external factors that tend to be stable, such as school satisfaction and a sense of belonging, reflecting a system organised around scenarios in which satisfaction and security are generalised. Conversely, the resilient cohort reveals a different mechanism. Internal resources, such as attentional control, self-efficacy, and metacognition, replace external modellers as structural bridges. These are core components of self-regulated learning (SRL). Their elevated centrality aligns with recent resilience literature [1][2][3].

Therefore, studying how to transform these factors into structural axes is very likely a pedagogical objective to help students from low socioeconomic backgrounds boost the flow of support resources through self-regulation strategies. This strategy might help to compensate for the inherited socioeconomic and cultural capital deficits of more affluent groups [40]. From this perspective, it can also be inferred that adaptive mechanisms for academic success are not automatic and that resilient students not only work harder, but their internal resources occupy structurally distinct positions, suggesting a different mode of operation [2][41]. A plausible hypothesis for future research is that resilient students focus their efforts on intermediation, building bridg-

es between their limited access to or enjoyment of the factors that drive learning, while privileged students benefit from widely distributed intermediation [51]. In this sense, the distinction lies not only in the metrics but also in the network design: one system thrives on redundancy, a clear expression of learning system efficacy; the other, through various optimisation mechanisms, is more focused on efficiency, with both demonstrating their effectiveness in achieving academic success.

Based on the observed contrasts, we provide clear, simple, and robust evidence that supports the idea that resilience should not be conceived solely as a personal attribute, but as a dynamic and adaptive process arising from the continuous interaction of individual, family, and institutional factors [2]. This dynamic process is operationalized through self-regulated learning, which enables students to orchestrate internal and external resources, showing how network theory offers a combination of advantages for analyzing multivariate relationships: (1) a robust analytical framework capable of capturing nonlinear relationships between a wide range of variables, strata, and contexts, (2) a series of formal tools for large-scale empirical analysis, providing enormous capabilities for understanding the dimensions involved in student learning and enabling them to succeed in adverse situations [33][34], providing fertile ground for advancing equity-oriented educational research and informing more context-adapted policies and practices [35][36].

Despite the new knowledge that this network analysis provides on the structural basis of academic resilience, several methodological limitations arise. First, the network construction was based on positive-definite correlation matrices, which implies the analysis of adverse factors using correlation matrices with negative magnitudes, thus requiring the inclusion of low-performing students to carry out the comparisons. Secondly, given that betweenness centrality provides evidence of functional control, it is essential to conduct longitudinal panel data to estimate parametric invariance across related sociocognitive and socioemotional factors [38]. Third, to test the stability of the factors involved in topological substitution, it might be useful to implement a dynamic network analysis (DNA) to assess whether the observed phenomena are consistent and to confirm that the key drivers of resilience remain the same. Finally, to continue with future research, student-level interactions within the schools' structural environment should be segmented by territorial characteristics, given that they are relevant for defining social and cultural aspects that can influence not only different practices but also educational expectations and academic performance.

5 Acknowledgments

The authors gratefully acknowledge the Government of Nuevo León for providing the comprehensive dataset, the participating teachers and students for their invaluable contributions, and the Latin American Agency of Evaluation and Policy (Aleph) for its essential support in advancing this research. Carmona-Soto, M. B. acknowledges financial support from the *Fundação para a Ciência e a Tecnologia*, I.P. (FCT), Portugal, under grant 2025.06357.BD.

References

1. Weed, K., Keogh, D., & Borkowski, J. (2006). Stability of resilience in children of adolescent mothers. Journal of Applied Developmental Psychology, 27(1), 60-77.
2. Wang, F. (2025). Fostering resilience among university students: The role of resilience as a dynamic process of positive adaptation when facing adversity. Higher Education Research & Development. https://doi.org/10.1007/s10734-025-01484-2
3. Mishchuk, H., Oliinyk, O., & Bilan, Y. (2024). Brain gain and country's resilience: A dependency analysis exemplified by OECD countries. Equilibrium (1689-765X), 19(2).
4. Tonegawa, Y. (2023). Education in SDGs: What is inclusive and equitable quality education? In Sustainable Development Disciplines for Humanity (pp. 55–70). Springer.
5. UNESCO (2015). Education 2030: Incheon Declaration and Framework for Action Towards Inclusive and Equitable Quality Education and Lifelong Learning for All.
6. UNESCO Institute for Statistics (UIS). (2024). World Education Statistics, 2024. Paris: UNESCO.
7. UNESCO. (2024). Monitoring SDG 4: 2024/5 Global Education Monitoring Report. Paris: UNESCO.
8. Hanushek, E. A., & Woessmann, L. (2020). The economics of international differences in educational achievement. Journal of Economic Perspectives, 34(4), 83–108.
9. Mishchuk, H., Oliinyk, O., & Bilan, Y. (2024). Brain gain and country's resilience: A dependency analysis exemplified by OECD countries. Equilibrium (1689-765X), 19(2).
10. World Bank, UNESCO, & UNICEF. (2021). The state of the global education crisis: A path to recovery. Washington, DC: World Bank.
11. World Bank, UNESCO, & UNICEF. (2022). The State of Global Learning Poverty: 2022 Update. Washington, DC: World Bank.
12. Senior, C., & Sahlberg, P. (2025). The evolution of the OECD's position on equity in global education. International Journal of Educational Development, 114, 103241.
13. Motl, T. C., Multon, K. D., & Zhao, F. (2018). Persistence at a tribal university: Factors associated with second year enrollment. Journal of Diversity in Higher Educ, 11(1), 51.
14. Tracey, B., & Florian, K. (Eds.). (2016). Educational Research and Innovation Governing Education in a Complex World. OECD Publishing.
15. Davis, B., & Sumara, D. (2014). Complexity and education: Inquiries into learning, teaching, and research. Routledge.
16. Brinkhof, L. P., Ridderinkhof, K. R., de Wit, S., Krugers, H. J., & Murre, J. M. J. (2025). A cross-sectional network analysis of successful aging in a resilience-based framework. *PLoS ONE*, *20*(1), e0315445.
17. Moll, K., Göbel, S. M., Gooch, D., Landerl, K., & Snowling, M. J. (2016). Cognitive risk factors for specific learning disorder: processing speed, temporal processing, and working memory. Journal of learning disabilities, 49(3), 272-281.
18. Sykes, C., Baştuğ, S. I., & Fiasconaro, S. (2025). A metareview of research on educational inequality and socioeconomic disadvantage. *Education Sciences, 15*(6), 740. https://doi.org/10.3390/educsci15060740
19. Johnson, J. H. (2016). Hypernetworks: multidimensional relationships in multilevel systems. The European Physical Journal Special Topics, 225(6), 1037-1052.
20. Johnson, J. H., Fortune, J., & Bromley, J. (2017). Systems, networks, and policy. In Non-Equilibrium Social Science and Policy (pp. 111-134). Springer, Cham.
21. van der Maas, H., Kan, K., MarSSan, M., & Stevenson, C. (2017). Network models for cognitive development and intelligence. Journal of Intelligence, 5(2), 1–17. doi: 10.3390/jintelligence5020016

22. Sánchez-Restrepo, H., & Louçã, J. (2019). Topological Properties of Inequality and Deprivation in an Educational System: Unveiling the Key-Drivers Through Complex Network Analysis. In International Conference on Human Systems Engineering and Design: Future Trends and Applications (pp. 469-475). Springer, Cham.
23. Azevedo, J. P. (2020). Learning poverty. World Bank Policy Research working paper, Washington DC: World Bank.
24. Martin, A. J., & Marsh, H. W. (2006). Academic resilience and its psychological and educational correlates: A construct validity approach. *Psychology in the Schools*, *43*(3), 267-281.
25. Rudd, K. L., & Yates, T. M. (2020). The Health Implications of Resilience. The Wiley Encyclopedia of Health Psychology, 269-277.
26. Rudd, G., Meissel, K., & Meyer, F. (2021). Measuring academic resilience in quantitative research: A systematic review of the literature. *Educational Research Review, 34*, 100402. https://doi.org/10.1016/j.edurev.2021.100402
27. Wang, F. (2025). Fostering resilience among university students: The role of resilience as a dynamic process of positive adaptation when facing adversity. *Higher Education Research & Development*. https://doi.org/10.1007/s10734-025-01484-2
28. Marta, E. M., & Michelle, C. (2023). OECD Skills Studies Gender, Education and Skills. The Persistence of Gender Gaps in Education and Skills. OECD Publishing.
29. Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*(3), 417–453.
30. Luthar, S. S., Cicchetti, D., & Becker, B. (2020). The construct of resilience: A critical evaluation and guidelines for future work. *Child Development*, *71*(3), 543–562.
31. Rudd, R. E., Wingo, J., & Bynum, L. (2021). Resilience: Adaptive and proactive coping in the process of developing resilience. In R. E. Rudd, J. Wingo, & L. Bynum (Eds.), *Research in occupational stress and well being* (Vol. 19, pp. 19–46). Emerald Publ. Limited.
32. Yan, Y., & Gai, X. (2022). High achievers from low family socioeconomic status families: Protective factors for academically resilient students. *International Journal of Environmental Research and Public Health, 19*(23), 15882. https://doi.org/10.3390/ijerph192315882
33. Siew, C. S. Q. (2020). Applications of Network Science to Education Research: Quantifying Knowledge and the Development of Expertise through Network Analysis. *Education Sciences, 10*(4), 101. https://doi.org/10.3390/educsci10040101
34. Saqr, M., Dever, D., López-Pernas, S., Gernigon, C., Marchand, G., & Kaplan, A. (2024). Complex dynamic systems in education: Beyond the static, the linear and the causal reductionism. In M. Saqr & S. López-Pernas (Eds.), Advanced Learning Analytics Methods: AI, Precision and Complexity (Chap. 12). Springer
35. OECD (2021). *Education Policy Outlook 2021: Shaping Responsive and Resilient Education in a Changing World*. OECD Publishing. https://doi.org/10.1787/75e40a16-en
36. OECD (2025), *Education at a Glance 2025: OECD Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/1c0d9c79-en
37. Xenofontos, C., & Mouroutsou, S. (2023). Resilience in mathematics education research: A systematic review of empirical studies. *Scandinavian Journal of Educational Research*, *67*(7), 1041-1055.
38. Ye, W., Strietholt, R., & Blömeke, S. (2021). Academic resilience: Underlying norms and validity of definitions. Educational Assessment, Evaluation and Account, 33(1), 169-202.
39. NLA (2025). *Características psicométricas de la evaluación Nuevo León Aprende.* Coordinación del ecosistema educativo: "Legado Nuevo León". Nuevo León, México.

40. Borsboom, D., & Molenaar, D. (2015). Psychometrics. In J. Wright (Ed.), International encyclopedia of the social & behavioral sciences. Vol.19 pp.418–422.Amsterdam: Elsevier
41. Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. Nature Reviews Methods Primers, 1(1), 1–23.
42. Zhang, L., & Peixoto, T. P. (2020). Statistical inference of assortative community structures. Physical Review Research, 2(4), 043271.
43. Hanushek, E. A., & Woessmann, L. (2020). The economics of international differences in educational achievement. Journal of Economic Perspectives, 34(4), 83–108.
44. Zhu, J., Dai, F., Zhao, F., & Guo, W. (2023). Integrating node importance and network topological properties for link prediction in complex network. Symmetry, 15(8), 1492.
45. Hauke, J., & Kossowski, T. (2011). Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones geographicae*, *30*(2), 87-93.
46. Schober, P., Boer C, Schwarte LA. (2018). Correlation Coefficients: Appropriate Use and Interpretation. Anesth Analg. 2018 May;126(5):1763-1768.
47. Hu, Y., & Shi, L. (2015). Visualizing large graphs. *Wiley Interdisciplinary Reviews: Computational Statistics*, *7*(2), 115-136.
48. Masuda, N., Boyd, Z. M., Garlaschelli, D., & Mucha, P. J. (2024). Correlation networks: Interdisciplinary approaches beyond thresholding.
49. Masuda, N., Sakaki, M., Ezaki, T., & Watanabe, T. (2018). Clustering coefficients for correlation networks. Frontiers in Neuroinformatics, 12, 7.
50. Newman, M. E. J. (2016). Community detection in networks: Modularity optimization and maximum likelihood are equivalent. Physical Review E, 94(5), 052315. https://doi.org/10.1103/PhysRevE.94.052315
51. Newman, M. E. J. (2018). Networks (2nd ed.). Oxford University Press.
52. Liu, C., Ding, J., Song, Y., & Li, Y. (2024, August). Tdnetgen: Empowering complex network resilience prediction with generative augmentation of topology and dynamics. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 1875-188)

CHAPTER 9

Conclusion and Perspectives

8.1. A three-phase project

From the beginning of this doctoral thesis, the core of this research has been that educational inequality cannot be fully understood through the lens of determinism or linear statistical or econometrical models. As argued in Chapter 1, the "general linear reality" [Abbott, 1988] has allowed researchers to describe achievement gaps but has often struggled to explain the mechanisms of their persistence. By adopting the ontology of Complex systems and Network science, we propose a foundational change in the educational paradigm for quantitative research, claiming that this approach offers powerful concepts for building new constructs and hierarchies to understand educational phenomena, as well as for analyzing inequality in learning outcomes. As can be expected, the transition from linear to complex models does not eliminate uncertainty; rather, it shifts the focus from attribute uncertainty to interaction uncertainty.

During this research project, three well-defined phases can be recognized. The first one, reflected in chapters 2, 3, and 4, showed how educational systems can be described as networks and that their parameters and topological features might offer highly valuable information for characterizing population groups and other spatial or artificial partitions of students and schools, as well as providing insights and thresholds for contrasting hypotheses. One aspect to consider is that the data used in these chapters are largely drawn from one or more Latin American countries during a specific assessment window. While our models successfully unveil some factors related to structural inequality, it is important to keep in mind that they capture a static snapshot of a process that changes slowly but remains dynamic, potentially giving a "frozen" view of the topology and other features that should be investigated in the coming years to understand their evolution. In this sense, while we identified that non-educational factors— SES or ethnicity—act as rigid topological constraints, the mechanism by which a factor evolves from a peripheral node to a central hub—or devolves into isolation—, a significant aspect for policy, remains as an open question that our current methodology has not answered yet and should be tested across the temporal evolution of these constraints. There is no doubt that future research should go in this direction to estimate transition probabilities to

different states—providing order parameters for the system— rather than just measuring the current state—as classical theory does through statistical expectations.

The second phase focused on extending the analysis to multiple systems and on comparing gaps and structural inequality across 15 Latin American countries (Chapters 5 and 6). In this part of the research, we found a striking "regional isomorphism" in the structure of deprivation across them. The data suggest that, regardless of a country's wealth, students' family socioeconomic status is a determinant for suffering learning deprivation. Furthermore, while it is true that the intensity of inequality in learning emerges consistently across all scales, the variation in the quality, quantity, and granularity of available official data varies from country to country, including that reported to international organizations, which introduces a layer of noise that our current models may not filter out. There is a small risk of overgeneralizing the "Latin American" results by smoothing over local events or how they have managed the COVID-19 pandemic. When comparing countries, it is unclear, for instance, whether the observed structural similarities are due to shared educational foundations or simply to having a similar economic development stage. The tension between the discovered regional structure and the specific "local modulations" remains a fertile ground for investigation, especially with UNESCO's forthcoming results available at the end of 2026.

The third phase consolidates results and opens new opportunities for the project. Departing from traditional approaches based on maximizing additive predictors—or to obtain an additive specification—our model examines the multivariate resource ecosystem and its interactions, analyzing the structural interplay between academic and socio-cognitive factors that drive high achievers across diverse socioeconomic settings. Moreover, incorporates a network-based definition of resilience, enabling the identification of specific topological mechanisms—rather than simple bivariate associations—that differentiate students who overcome socioeconomic disadvantage from their privileged high-performing peers. These findings are crucial because they are impossible to extract from econometric models, as they demonstrate how structural reorganization is implemented to compensate for resource scarcity, and indicate that high academic achievement, regardless of socioeconomic context, is fundamentally mediated by the stability of the resource network. Once again, the stark structural differences between high-achieving students from disadvantaged and privileged backgrounds, driven by topological substitution mechanisms whereby resilient learners strategically leverage internal resources—such as Self-knowledge, Metacognition, and Study habits. These mechanisms mediate and compensate for socioeconomic disadvantages, providing the concept of *topological substitution*, a significant theoretical advancement with the potential to be incorporated into the

jargon of educational research, in which resilient students rewire their networks to rely on internal nodes such as Self-Efficacy and Attentional control.

8.2. Methodological Constraints and Opportunities

The validity of the conclusions presented in this compendium is bound by specific constraints that define the scope for future inquiry:

1. Most of the networks constructed in this thesis are based on cross-sectional data from specific years. Educational systems are adaptive; they change in response to policy interventions, economic crises, and migration patterns, especially after pandemics such as COVID-19. A longitudinal network analysis based on a longitudinal data panel is necessary to go beyond and develop stronger, more plausible simulation scenarios.
2. In translating social phenomena into nodes and edges, we rely on proxies, e.g., using the level of studies or the kind of job of the parents as a proxy for cultural capital. While standard in the field, these proxies are imperfect measures of the rich, qualitative reality of a student's life. However, regarding the map is an approach to the territory, and that our topological map is an abstraction that inevitably loses some sociological nuance, the big numbers offer stable insights and conclusions in most cases.
3. While network science allows for strong inferences about structural dependencies, it does not inherently prove causality. We can assert that resilient students have a specific network structure. Still, we cannot definitively prove that the structure caused resilience without experimental or quasi-experimental designs, which might be feasible within the longitudinal data panel mentioned above.

Yet, this finding invites critical scrutiny of the nature of academic resilience and other positive educational phenomena, given that the network analysis identifies that this substitution occurs but does not explain its cost. In this sense, it might be plausible that the “rewiring process” imposes a high cognitive, emotional, or social tax on students from the most deprived backgrounds, potentially leading to burnout or “John Henryism”—when people face prolonged adversity, such as inequality or discrimination, the effort required to thrive will harm their health due to stress later in life. Seen this way, topological substitution should not be celebrated as a mechanism of agency or an individual's triumph but rather conceived as a survival strategy that highlights the system's failure. In this sense, it might be helpful to remember that LSA's

data used to be limited to educational outcomes, far away from potential negative externalities. In the future, it would be desirable to conduct more research using longitudinal panel data that integrates psychological and emotional aspects of students, with the aim of better understanding whether topological strategies are sustainable throughout a person's life.

8.3. Policy Perspectives: From Models to Action

The use of the results of this research for the design of public policy is one of the main social motivations for carrying it out, not only because they could be quite significant, but because they have already been considered or being implemented by some Latin American countries; however, they must be implemented with supervision and humility, since school communities are often immersed in inertias that need to be considered for policymakers get the best impact of the resource allocation. Overall, because a simulation is a controlled environment that assumes agents' behavior and the fluid transfer of resources. Real-world educational policies in Latin America are characterized by friction, bureaucracy, resistance, and, sadly, corruption. Therefore, the knowledge derived from this thesis is not a prescription for a technocratic takeover of education. Instead, it tries to offer a more heuristic tool for developing evidence-based policy. The value of mapping the “topology of inequality of deprivation” lies in exposing the hidden barriers that well-intentioned but linear policies often miss. We claim for relational models not just because they might be more valuable and realistic than the classic recipes, but because in a region with a culture often resistant to evaluation and accountability, making the structure visible is the first step toward challenging it.

In conclusion, this thesis confirms that the tools of complexity science are essential for reframing the analysis of educational inequality and most of the large-scale educational phenomena. Over these years, we have shown that the “achievement gap” functions are topological traps, and that resilience involves a structural substitution of resources. Yet these findings are still opening statements—as most findings in educational science—not final verdicts. Finally, we would like to state that the structural isomorphism across Latin America—and the mechanisms of individual agency presented here—can serve as a rigorous baseline for continuing this research. The main challenge for the future is to test the temporal stability of these structures and to deepen our understanding of the human cost of resilience, moving from the statistical physics of education to a human-centered science of complexity.

References

Azevedo, J. P. (2020). Learning poverty. World Bank Policy Research working paper, Washington DC: World Bank.

Barabási, A. L. (2003). Emergence of scaling in complex networks. Handbook of Graphs and Networks: From the Genome to the Internet. Berlin: Wiley-VCH.

Bloome, D., Dyer, S., & Zhou, X. (2018). Educational inequality, educational expansion, and intergenerational income persistence in the United States. *American Sociological Review*, 83(6), 1215-1253.

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, 1(1), 1–23.

Bourdieu, P. (1986). The forms of capital. En I. Szeman & T. Kaposy (Eds.), Cultural theory: An anthology (pp. 81–93). *Wiley-Blackwell*

Davis, B., & Sumara, D. (2014). Complexity and education: Inquiries into learning, teaching, and research. *Routledge*.

Deguchi, T., Takahashi, K., Takayasu, H., & Takayasu, M. (2014). Hubs and authorities in the world trade network using a weighted HITS algorithm. PloS one, 9(7), e100338.

Espinosa, A., & Walker, J. (2017). Complexity approach to sustainability, a: theory and application (Vol. 5). *World Scientific*.

Gao, J., Barzel, B., & Barabási, A. L. (2016). Universal resilience patterns in complex networks. *Nature*, 530(7590), 307-312.

Hanushek, E. A., & Woessmann, L. (2020). The economics of international differences in educational achievement. *Journal of Economic Perspectives*, 34(4), 83–108.

Hazy, J. K., & Ashley, A. (2011). Unfolding the future: Bifurcation in organizing form and emergence in social systems. *Emergence: Complexity & Organization*, 13(3), 57-79.

Hébert-Dufresne, L., Allard, A., Garland, J., Hobson, E. A., & Zaman, L. (2024). The path of complexity. npj Complexity, 1(1), 4.

Johnson, J. H. (2016). Hypernetworks: multidimensional relationships in multilevel systems. The European *Physical Journal Special Topics*, 225(6), 1037-1052.

Johnson, J. H., Fortune, J., & Bromley, J. (2017). Systems, networks, and policy. In *Non-Equilibrium Social Science and Policy* (pp. 111-134). Springer, Cham.

Liu, Y. Y., & Barabási, A. L. (2016). Control principles of complex systems. *Reviews of Modern Physics*, 88(3), 035006.

Marta, E. M., & Michelle, C. (2023). OECD Skills Studies Gender, Education and Skills. The Persistence of Gender Gaps in Education and Skills. OECD Publishing.

Mason, M. (2008). Complexity theory and the philosophy of education. *Educational philosophy and theory*, 40(1), 4-18.

Masuda, N., Sakaki, M., Ezaki, T., & Watanabe, T. (2018). Clustering coefficients for correlation networks. *Frontiers in Neuroinformatics*, 12, 7.

Masuda, N., Boyd, Z. M., Garlaschelli, D., & Mucha, P. J. (2024). Correlation networks: *Interdisciplinary approaches beyond thresholding*.

Morin, E. (1992). From the concept of system to the paradigm of complexity. *Journal of social and evolutionary systems*, 15(4), 371-385.

Morin, E. (2007). Complejidad restringida y Complejidad generalizada o las complejidades de la Complejidad. Utopía y praxis latinoamericana, 12(38), 107-119.

Mueller, J. S., Wakslak, C. J., & Krishnan, V. (2014). Construing creativity: The how and why of recognizing creative ideas. *Journal of Experimental Social Psychology*, 51, 81-87.

Newman, M. E. J. (2016). Community detection in networks: Modularity optimization and maximum likelihood are equivalent. *Physical Review E*, 94(5), 052315. https://doi.org/10.1103/PhysRevE.94.052315

Newman, M. E. J. (2018). Networks (2nd ed.). Oxford University Press.

Nowosad, I. (2021). Educational policy in an integrated model of intervention at the macro, meso and micro levels. Designing and implementing public policy of contemporary polish society: Selec. problems, 115-126.

Piketty, T. (2020). Capital and ideology. Harvard University Press.

Pinsky, M., & Karlin, S. (2010). An introduction to stochastic modeling. Academic press.

Pósfai, M., & Barabási, A. L. (2016). Network science (Vol. 3). Cambridge, UK. Cambridge University Press.

Rudd, G., Meissel, K., & Meyer, F. (2021). Measuring academic resilience in quantitative research: A systematic review of the literature. *Educational Research Review*, 34. https://doi.org/10.1016/j.edurev.2021.100402

Sanchez-Restrepo, H. (2018). Autonomía de los organismos de evaluación. *Estado Abierto. Revista sobre el Estado, la administración y las políticas públicas*, 2(3), 15-45.

Sánchez-Restrepo, H., & Louçã, J. (2019). Topological Properties of Inequality and Deprivation in an Educational System: Unveiling the Key-Drivers Through Complex Network Analysis. In International Conference on Human Systems Engineering and Design: Future Trends and Applications (p. 469-475). *Advances in Intelligent Systems and Computing*. Springer, Cham. https://doi.org/10.1007/978-3-030-27928-8_71

Sánchez-Restrepo H., Louçã J. (2020). Inequality in Learning Outcomes: Unveiling Educational Deprivation Through Complex Network Analysis. In: Cherifi H., Gaito S., Mendes J., Moro E., Rocha L. (eds) Complex Networks and Their Applications VIII. COMPLEX NETWORKS 2019. *Studies in Computational Intelligence*, vol 882. Springer, Cham. https://doi.org/10.1007/978-3-030-36683-4_27

Sanchez-Restrepo H., Louça J. (2020). A Complex Network Approach to Structural Inequality of Educational Deprivation in a Latin American Country. In: Barbosa H., Gomez-Gardenes J., Gonçalves B., Mangioni G., Menezes R., Oliveira M. (eds) Complex Networks XI. *Springer Proceedings in Complexity*. Springer, Cham. https://doi.org/10.1007/978-3-030-40943-2_29

Sánchez-Restrepo H., Louçã J. (2020). Educational Deprivation in Latin America: Structural Inequality Beyond Borders. In: Nazir S., Ahram T., Karwowski W. (eds) Advances in Human Factors in Training, Education, and Learning Sciences. AHFE 2020. *Advances in Intelligent Systems and Computing*, vol 1211. Springer, Cham. https://doi.org/10.1007/978-3-030-50896-8_14

Sánchez-Restrepo H. and Jorge Louçã (2021). Educational Gaps in Latin American childhood: Regional Inequality in Learning Deprivation with Local Modulations. *Journal of Physics*: Conf. Ser. **1828** 012135. https://doi.org/10.1088/1742-6596/1828/1/012135

Sánchez-Restrepo, H., & Carmona-Soto, M. B. (2023). Aprendizajes imprescindibles en la educación básica: Avances y desafíos durante 2023. En: Secretaría de Educación de Nuevo León, México. https://zenodo.org/records/15205761

Sánchez-Restrepo, H., Carmona-Soto, M. B., Morales-Garza, S. & Almeida-Díaz, L. G. (2025). Cuatro ciclos escolares mejorando los resultados de aprendizaje. Evolución de los aprendizajes de los estudiantes de educación básica. Informe estatal 2022 - 2025. Nuevo León Aprende. Coordinación del ecosistema educativo: "Legado Nuevo León". Nuevo León, México.

Sánchez-Restrepo, H. (in press, 2026). Topological substitution: Mapping the structural mechanisms of academic resilience via complex network analysis. In Socioeconomic studies of today's Mexico: A mathematical overview (Springer Series). Springer Nature.

Sánchez-Restrepo, H., Louçã, Jorge, Carmona-Soto, M. B. & Morales-Garza, S. (2026). Topological Patterns of Academic Success Insights from Complex Network Analysis. *Springer Proceedings in Complexity*. Springer Nature

Schober, P., Boer C, Schwarte LA. (2018). Correlation Coefficients: Appropriate Use and Interpretation. Anesth Analg. 2018 May;126(5):1763-1768.

Sen, A. (1982). Poverty and famines: an essay on entitlement and deprivation. Oxford university press.

Solé, R. V. (2011). Phase transitions (Vol. 16). Princeton University Press.

Tracey, B., & Florian, K. (Eds.). (2016). Educational Research and Innovation Governing Education in a Complex World. OECD Publishing.

UNESCO (2015). Education 2030: Incheon Declaration and Framework for Action Towards Inclusive and Equitable Quality Education and Lifelong Learning for All.

UNESCO (2017). UNESCO moving forward the 2030 Agenda for Sustainable Development (Document code BSP-2017/WS/1). UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000247785

UNESCO (2022). Global Education Monitoring Report 2022: Gender report – Deepening the debate on those still left behind (Informe GEM). https://doi.org/10.54676/RCZB6329

UNESCO (2024). Monitoring SDG 4: 2024/5 Global Education Monitoring Report. Paris: UNESCO.

UNESCO (2026). Global Education Monitoring Report 2026: Access and equity – Countdown to 2030. Paris, UNESCO.

UNESCO & Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI) (2025). Global Education Monitoring Report 2025 on Leadership in Education – Regional Edition for Latin America: Lead for democracy (Informe de seguimiento de la educación en el mundo 2025). UNESCO. https://doi.org/10.54676/FKZA2207

UNESCO Institute for Statistics (2024). World Education Statistics, 2024. Paris: UNESCO.

UNESCO Institute for Statistics (2024a). National SDG 4 benchmarks: Challenges and solutions forward (Programme and meeting document No. UIS/EDS/9). UNESCO. https://unesdoc.unesco.org/ark:/48223/pf000038

UNESCO Institute for Statistics. (2022). Global Alliance to Monitor Learning (GAML): Report on learning assessment and monitoring frameworks. UNESCO. Disponible en: https://gaml.uis.unesco.org

van der Maas, H., Kan, K., MarSSan, M., & Stevenson, C. (2017). Network models for cognitive development and intelligence. Journal of Intelligence, 5(2), 1–17. doi: 10.3390/jintelligence5020016

Van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. Educational research review, 5(2), 134-150.

Wang, F. (2025). Fostering resilience among university students: The role of resilience as a dynamic process of positive adaptation when facing adversity. Higher Education Research & Development. https://doi.org/10.1007/s10734-025-01484-2