ARTICLE

Conceptual evolution and methodological challenges of value-added in educational assessment: A scoping review 1971–2024

Kelly Alice Barbosa Macêdo [a]
Maria Eugénia Ferrão [b]
Marcel de Toledo Vieira [c]

Abstract

Value-added (VA) models are widely used to evaluate educational effectiveness, yet their conceptual and methodological development remains insufficiently systematized. This study conducts a scoping review, complemented by an exploratory bibliometric analysis of studies published between 1971 and 2024, focusing on the seminal contributions of Hanushek (1971) in Education economics and Bryk and Weisberg (1976) in educational statistics. The findings reveal persistent gaps in the literature, notably the limited evidence from developing countries and the underexplored roles of school leadership and causal inference within VA frameworks. The review highlights a growing diversification of methodological approaches alongside an increasing alignment with accountability policies. From a policy perspective, the results underscore the need for context-sensitive VA applications, stronger school leadership, improved longitudinal data systems, and closer collaboration between researchers and policymakers to support more effective and equitable educational interventions.

Keywords: Educational Evaluation. Educational Effectiveness. School Effectiveness. Social Equity. Scoping Review.

[a] Federal University of Juiz de Fora, Juiz de Fora, MG, Brazil.
[b] University of Beira Interior, Covilhã; Cemapre and ISEG Research, Lisboa, Portugal.
[c] Federal University of Juiz de Fora, Juiz de Fora, MG, Brazil.

Received: Dec 18, 2024
Accepted: Feb 09, 2026

1 Introduction

The term "value-added" is commonly used in economics to describe the value created throughout a production process, reflecting the contribution of each stage to the generation of new value. Eric Hanushek was a pioneer in transferring this concept to the educational context, and in his seminal 1971 study, "Teacher Characteristics and Gains in Student Achievement", he examined how teacher and classroom characteristics influence student achievement over time, establishing Value-Added (VA) analysis as a foundational approach in educational evaluation.

> Families obviously have considerable impact on education through physical conditions, attitude formation and direct involvement in the educational process. Since these factors tend to be highly correlated with socioeconomic status of the family, this aspect of education is proxied by father's occupation and family structure. The influence of peers is much the same as that from the family, and, thus, this aspect is proxied by aggregate measures of the socioeconomic status of individuals in a given class or school. While innate abilities are included in the conceptual model, there is no direct measure of this aspect. However, there is reason to believe that biases in the school parameters due to this missing variable are minimal. First, the model with initial achievement measures the "value added" of various inputs and biases will occur only if the missing portion of innate abilities is correlated with the rate of learning (as opposes to the level). Second, at least for whites, it is reasonable to assume that this factor is captured fairly well in the family background variables. This is the case if innate abilities tend to be hereditary and if social mobility is highly correlated with ability. Severe problems, at least in the school portion of the model, do not arise unless there is a mechanism which leads to the correlation of the 'nonhereditary' part of innate abilities and specific school resources (Hanushek, 1971, p. 281-282).

Hanushek (1971) introduced a regression-based framework to estimate VA in Education by assessing the contributions of schools and teachers while controlling for prior achievement, family background, socioeconomic status, and peer effects. By conditioning current performance on past outcomes, the VA approach shifts evaluation from absolute achievement levels to incremental learning gains attributable to teachers, classrooms, and schools. Rather than a specific statistical technique, VA constitutes a conceptual framework in which estimated coefficients

are interpreted as contributions to students' learning trajectories, net of expected performance based on prior achievement and background characteristics.

Within this VA framework, Bryk and Weisberg (1976) made a seminal contribution to the formalization and application of the VA concept in Education. In "Value-Added Analysis: A Dynamic Approach to the Estimation of Treatment Effects", they proposed a methodology to evaluate educational interventions when randomized experiments are not feasible, modeling expected growth attributable to the intervention to estimate its VA. This approach enabled direct assessment of program effects and advanced both the theoretical development and practical application of VA in educational evaluation.

> Our alternative view of educational processes focuses on individual growth. In the ideal situation, we would like to have available complete descriptions of growth under "natural" conditions. Thus for each subject, prior to the implementation of the program, the expected outcome could be predicted. The observed value could then be compared with the predicted value. The difference, or value-added, represents the additional benefit attributable to the program, and is thus a natural measure of treatment effect. (Bryk; Weisberg, 1976, p. 131-132).

Together, Hanushek (1971) and Bryk and Weisberg (1976) contributed to establishing the logic of VA approaches. Their work shifted educational evaluation from static achievement levels to estimating the incremental effects of schools, teachers, and interventions. This framework underpinned subsequent research on educational effectiveness, conceptualized as the capacity of schools and teachers to generate learning gains beyond students' initial achievement.

This conceptualization of educational effectiveness is particularly relevant considering the documented disparities in educational outcomes among groups categorized by socioeconomic status, race/ethnicity, or gender, which reflect persistent challenges regarding social and educational equity (Ferrão, 2022a, 2022b; Ferrão; Alves, 2023). In response to such disparities, public policies and educational programs must be actively designed to promote social and educational equity. Research has consistently shown association between students' sociodemographic attributes and educational outcomes. Still, some recent evidence suggests that the gaps are narrowing (Armor; Marks; Malatinszky, 2018; Ferrão, 2022b; Ferrão; Alves, 2023; Hanushek *et al*., 2022; Marks, 2017), providing important insights into how public policy can be mobilized in the service of equity in Education.

The literature on VA models (Amrein-Beardsley; Pivovarova; Geiger, 2016; Everson, 2017; Ferrão; Couto, 2013; Levy *et al*., 2019; Saunders, 1999; 2000) has addressed the relationship between students' Socioeconomic Status (SES) and achievement showing inspiring results that demonstrate the effectiveness of some schools (Albernaz; Ferreira; Franco, 2002; Alves; Franco, 2008; Brooke *et al*., 2014; Charalambous; Kyriakides; Creemers, 2018; Kyriakides *et al*., 2019; Sammons, 2007; Sammons; Hillman; Mortimore, 1995; Soares *et al*., 2017; Thomas; Peng; Gray, 2007), and in particular their differential effectiveness (Palardy, 2008; Strand, 2010; 2012; Twersky, 2022) in closing the gaps. However, concerns on VA methodological issues, purposes and uses are still on the agenda of scientific research (AERA - American Educational Research Association, 2015; Morganstein; Wasserstein, 2014). This evolution is particularly relevant for contexts like Brazil, where scholars have applied VA models to examine educational effectiveness amid persistent social inequalities (Bartolucci *et al*., 2023; Brooke *et al*., 2014; Ferrão, 2022a, 2022b; Soares *et al*., 2017).

In this regard, mapping the literature citing Hanushek (1971) and Bryk and Weisberg (1976) enables an understanding of how the VA concept has evolved within educational assessment and been applied across diverse empirical contexts. Based on a scoping review of publications from 1971 to 2024 identified through systematic searches of the Web of Science and Scopus databases, and complemented by an exploratory bibliometric analysis, this study analyzes these citing works to identify recurring patterns, emerging trends, shifts in research priorities, and persistent gaps in the literature. By clarifying how VA has been operationalized across different research strands, the study contributes to methodological and policy-oriented debates on educational effectiveness, accountability, and equity.

This study is organized into three sections. The following section outlines the methodology and the study selection process. The third section presents the results and discussion, and the final section concludes the study.

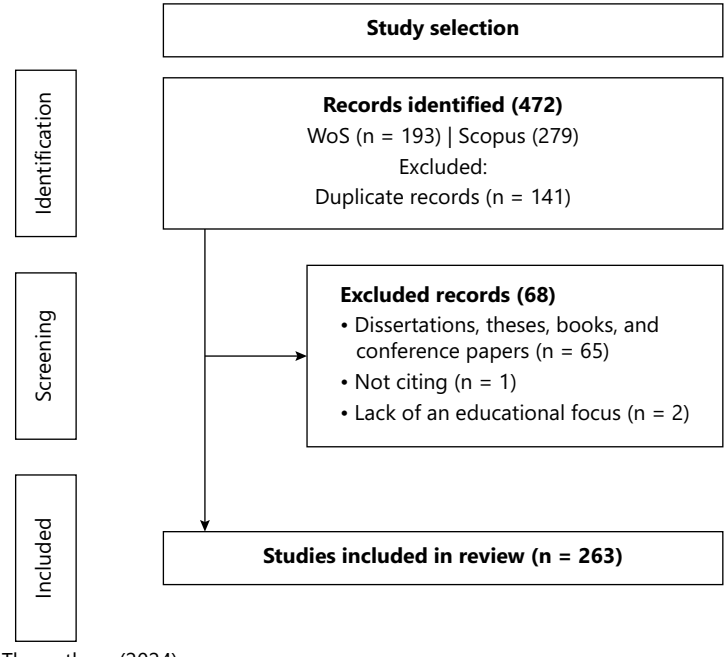
2 Methodology

In line with the objective of mapping the development of VA research in Education, this study addresses the following research question: What are the main conceptual, methodological, and thematic developments in VA research in Education since its introduction by Hanushek (1971) and Bryk and Weisberg (1976)? To address this research question, this study adopts a scoping review approach, which allows for an exploratory mapping of the existing literature to identify trends, recurring patterns, and gaps without critically evaluating individual studies. Compared to systematic reviews, this method encompasses a broader range of study designs and methodologies and is conducted under a transparent and rigorous research

protocol (Arksey; O'Malley, 2005; Braunack-Mayer *et al*., 2020; Pham *et al*., 2014). The review draws on studies citing Hanushek (1971) and Bryk and Weisberg (1976) indexed in Scopus and Web of Science (WoS), ensuring comprehensive and interdisciplinary coverage.

Data were collected on July 5, 2024, and an exploratory bibliometric analysis was conducted as the first stage of the scoping review using the Bibliometrix package in R and the Biblioshiny interface (Aria; Cuccurullo, 2017). This stage mapped the structure, evolution, and thematic organization of the literature and guided the subsequent selection and in-depth analysis of studies. To ensure that only studies relevant to the research objectives were considered, the inclusion criteria required: (1) articles in Portuguese, English, and Spanish; (2) specific citation of Hanushek (1971) between 1971 and 2024, or Bryk and Weisberg (1976) between 1976 and 2024, to capture the influence of these studies; and (3) educational focus. To avoid including records that could compromise the focus of the review, the exclusion criteria were applied: (1) dissertations, theses, books, and conference papers; (2) studies not citing any of the specified works; and (3) studies lacking an educational focus.

Figure 1 – Identification and selection of studies citing Hanushek (1971) and Bryk and Weisberg (1976)



Source: The authors (2024)

A search of the Scopus and WoS databases identified 472 articles citing Hanushek (1971) or Bryk and Weisberg (1976), of which 263 met the inclusion criteria after deduplication and screening (Figure 1). Screening and data extraction were conducted by a single author. The analysis classified studies by publication year, objectives, citation context, methodology, main findings, and journal impact factor, with the aim of providing a comprehensive overview of the development of the VA concept in the Education literature citing Hanushek (1971) and Bryk and Weisberg (1976). Building on this classification, the following section reviews these studies to examine the evolution and key characteristics of the VA concept, identify patterns, trends, and research gaps across diverse methodological approaches and educational subfields, and analyze collaboration networks among authors to reveal broader theoretical shifts and the growing policy relevance of VA[1].

3 Results and discussion

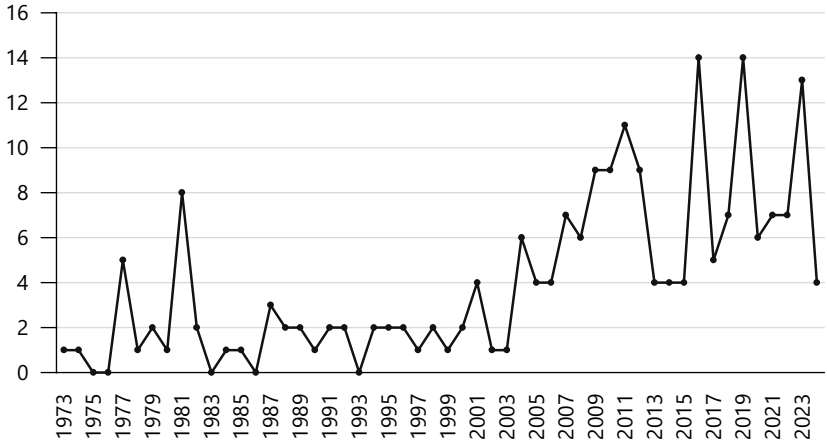
3.1 Analysis of studies citing Hanushek (1971)

As part of the scoping review, a bibliometric analysis of collaboration networks reveals that influential authors citing Hanushek's work are concentrated in high-income countries, particularly the United States, Germany, the United Kingdom, the Netherlands, and Italy, highlighting persistent geographic disparities in research capacity and data access and reinforcing the need to expand studies in developing regions (Armor *et al*., 2018; Ferrão, 2022b; Hanushek *et al*., 2022). Despite their diverse educational contexts and policy challenges, regions such as Ibero-America remain underrepresented in the literature (Larrea; Rojas; Hernández, 2024; Martínez-Abad; Gamazo; Rodríguez-Conde, 2020; Santelices *et al*., 2017). At the same time, the number of studies citing Hanushek has grown markedly over the past two decades (Figure 2), with academic output concentrated in economics (38%) and education (32%), and published in both leading economics journals, such as American Economic Review and Econometrica, and specialized education journals (Table A1, Appendix A[2]), underscoring the cross-disciplinary relevance of this body of research.

[1] In the reference list, studies marked with an asterisk (*) indicate studies that form part of the scoping review. A complete list of all included studies is available in Tables A2 and A3 in Appendix A.

[2] Appendix A is available online at: https://doi.org/10.5281/zenodo.18665963

Figure 2 – Annual scientific output of articles citing Hanushek (1971)

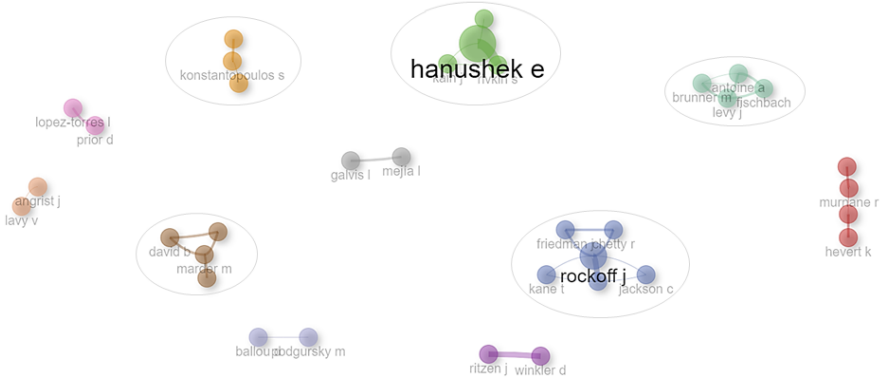


Source: The authors (2024)

3.1.1 Analysis of the authors' collaboration network and influential clusters

Influential clusters surrounding Hanushek's work document both empirical advances in VA estimation and their growing impact on policy-relevant interpretations of teacher and school effectiveness. Based on the collaboration network, Figure 3, which informed the selection of articles in Table A2 (Appendix A), five prominent clusters are identified, Hanushek, Rockoff, Marder, Konstantopoulos, and Levy, reflecting co-authorship patterns and thematic specialization. Analysis of the clusters reveals central themes: Hanushek focuses on educational policy and economic performance; Rockoff examines school practice and management; Marder evaluates educational programs with a focus on STEM teachers; Konstantopoulos studies educational inputs and teacher effectiveness; and Levy explores VA models in education with a focus on accuracy and methodology. This thematic distribution is consistent with broader reviews indicating that VA research spans policy analysis, classroom practices, and methodological development (Amrein-Beardsley *et al*., 2013; Ferrão; Couto, 2013; Hanushek, 2019; Muñoz; Prem, 2024; Reynolds *et al*., 2014), highlighting how the field has progressively diversified to address both methodological rigor and the practical challenges of improving teacher and school performance.

Figure 3 – Collaboration network among authors citing Hanushek (1971), 1971–2024



Source: The authors (2024)

Hanushek (1971) laid the groundwork for understanding the relationship between school inputs and student performance by introducing the concept of VA. His findings on resource effects prompted further investigation, including the meta-analysis by Hedges, Laine and Greenwald (1994), which highlighted the significant role of school investment in student achievement. Subsequent studies, including Nye, Konstantopoulos and Hedges (2004) and Rivkin, Hanushek and Kain (2005), applied multilevel and fixed-effects models to demonstrate that teacher experience and school quality substantially influence student outcomes. More recent research by Chetty, Friedman and Rockoff (2014a; 2014b), Jackson, Rockoff and Staiger (2014) and Levy *et al*. (2019) reflects methodological advances in VA estimation, incorporating strategies to account for unobserved student characteristics, refine model specifications, and improve transparency. These advances enable the analysis of more complex and contemporary issues, such as differential school effectiveness and the capacity of schools to mitigate social inequalities (Ferrão, 2022a, 2022b; Reynolds *et al*., 2014; Strand, 2016). Table A2 in Appendix A summarizes the main characteristics of studies within this cluster.

Building on these methodological advances, some of the studies identified in this scoping review have examined the substantive implications of VA models for educational outcomes and public policy. Research by Chetty, Friedman and Rockoff (2014a; 2014b), Levy *et al*. (2020, 2023), and Emslander *et al*. (2022) demonstrates that VA estimates are associated with positive long-term outcomes, including higher college enrolment rates and future earnings. These patterns are consistent with broader methodological discussions in the

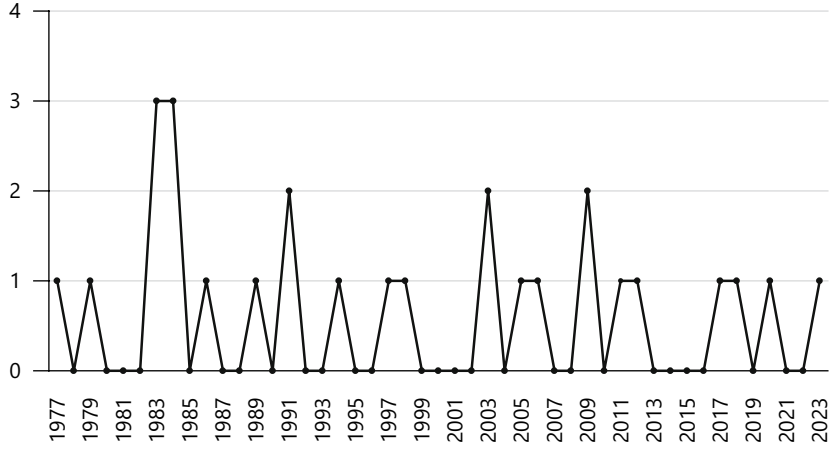
literature, which emphasize the relevance of incorporating contextual covariates and advanced statistical techniques to enhance the reliability of VA measures (Amrein-Beardsley *et al*., 2013; Sirin, 2005). Within the reviewed corpus, studies on teacher effectiveness consistently highlight the persistence of teacher effects and show that teacher turnover undermines both student achievement and the stability of VA scores, a concern also widely discussed in the methodological literature (Braun, 2013; Koedel; Mihaly; Rockoff, 2015). At a broader level, Hanushek and Woessmann (2011) extend the VA framework to the macroeconomic sphere, suggesting that improvements in educational effectiveness contribute to long-term economic growth in Organisation for Economic Co-operation and Development (OECD) countries.

However, despite these potential benefits, such as increased test scores among disadvantaged students, the impact of VA estimates is contingent on how they are implemented. In high-stakes accountability contexts without adequate institutional support, these same metrics can contribute to adverse outcomes, including increased teacher turnover and widening school-level disparities. Evidence from large-scale reforms such as the *No Child Left Behind Act* (Hanushek; Rivkin, 2010), and from isolated instruments of fragmented policies, like merit pay, suggest that the use of VA estimates for high-stakes decisions may produce adverse outcomes such as teacher shortages. These findings align with broader methodological concerns regarding the instability and specification sensitivity of VA models and are especially relevant in unequal and institutionally heterogeneous contexts such as Brazil, where accountability policies grounded in VA metrics but detached from local conditions may exacerbate teacher turnover and deepen school-level disparities (Amrein-Beardsley, 2023).

Studies citing Hanushek (1971) predominantly use VA to assess educational policies by estimating the contributions of schools and teachers to student achievement, with influential author clusters advancing both methodological and policy-relevant insights. However, this literature remains concentrated in high-income countries, and the growing use of VA indicators in teacher evaluation, resource allocation, and accountability highlights the need for caution in high-stakes applications. These concerns are especially relevant in contexts such as Brazil, where institutional capacity constraints and persistent inequalities may magnify associated risks. Improving educational outcomes therefore requires attention to teacher training and retention, as even methodologically robust VA models can produce unintended effects when institutional design and managerial capacity are weak.

3.2 Analysis of studies citing Bryk and Weisberg (1976)

Figure 4 – Annual scientific output of articles citing Bryk and Weisberg (1976)



Source: The authors (2024)

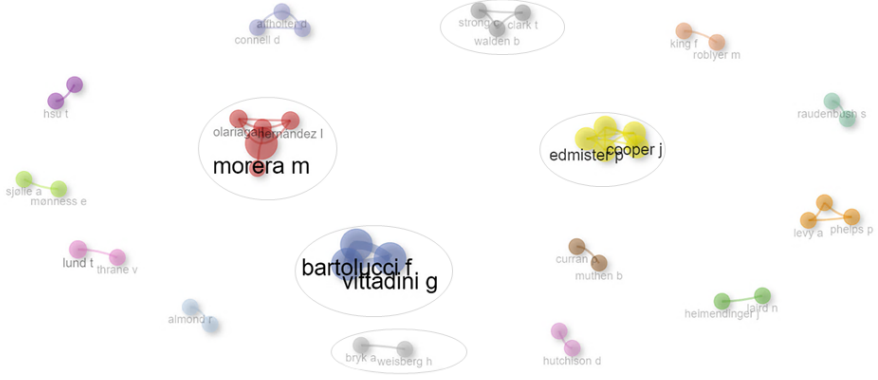
Building on early VA conceptualizations, Bryk and Weisberg (1976) introduced key methodological innovations for estimating educational effects in non-experimental settings. Between 1976 and 2024 (Figure 4), 27 studies citing their work were identified, with network analysis revealing a geographic concentration of influential authors in the United States, like that observed for studies citing Hanushek (1971). Although less frequently cited, the continued use of Bryk and Weisberg's (1976) framework reflects its sustained relevance in specific research contexts. Given the fragmented citation pattern, the in-depth analysis focused on seven studies forming significant collaborative clusters, representing consolidated methodological lineages. These studies adopt a multidisciplinary approach across Education, psychology, and applied statistics and are published in high-impact journals such as Psychological Methods and Psychological Bulletin (Table A1, Appendix A).

3.2.1 Analysis of the authors' collaboration network and influential clusters

The collaboration network analysis of studies citing Bryk and Weisberg (1976) reveals a small number of influential author clusters. Three main clusters are led by Bartolucci, Cooper, and Morera, with two additional clusters associated

with Muthen and Bryk, as shown in Figure 5 and summarized in Table A3 (Appendix A). With the exception of the Bartolucci and Morera clusters, most clusters consist of a single article, indicating a fragmented pattern of citation diffusion within this literature.

Figure 5 – Collaboration network of authors citing Bryk and Weisberg (1976), 1976–2024



Source: The authors (2024)

The identified clusters provide substantive insights into the evolution of statistical modeling in educational evaluation. Bryk and Weisberg (1977) critically assessed traditional cross-sectional approaches and proposed linear growth models based on pre- and post-test data to more accurately capture learning trajectories over time, a perspective later developed and operationalized in the broader literature on VA and growth curve literature (Palardy, 2008; Primi; Ferrão; Almeida, 2010). This approach also contributed to the development of subsequent longitudinal and multilevel modeling frameworks, which extend VA analyses to more complex educational data structures (Goldstein, 2003; Hox; Moerbeek; Van de Schoot, 2017; Singer, 1998).

Building on this framework, researchers such as Muthen and Curran (1997), Morera *et al*. (2009), and Bartolucci, Pennoni and Vittadini (2011; 2023) advanced educational evaluation through latent variable models, longitudinal designs, and causal extensions that account for school- and classroom-level effects. These innovations aimed to enhance the precision of intervention effect estimates and the modeling of dynamic educational processes, providing a

stronger foundation for assessing educational effectiveness in complex and heterogeneous contexts.

In summary, the studies citing Bryk and Weisberg (1976) primarily use VA to evaluate educational interventions, with methodological innovations that support longitudinal, multilevel, and latent variable approaches. These advances enable more precise estimation of program effects and capture the dynamic nature of learning, informing debates on educational effectiveness in contexts where randomized experiments are not feasible, including settings that rely on large-scale administrative data such as Brazil. At the same time, the uneven adoption of these methods highlights that methodological advances must be supported by adequate research infrastructure, funding, and data access to generate robust results capable of informing public policy.

3.3 Thematic analysis of studies

While Sections 3.1 and 3.2 examined the evolution of VA research through distinct citation lineages, the following thematic synthesis integrates these trajectories to identify broader patterns and implications for educational evaluation. Rather than merely indicating keyword frequency, thematic maps[3] are interpreted here as signals of how the VA field has expanded.

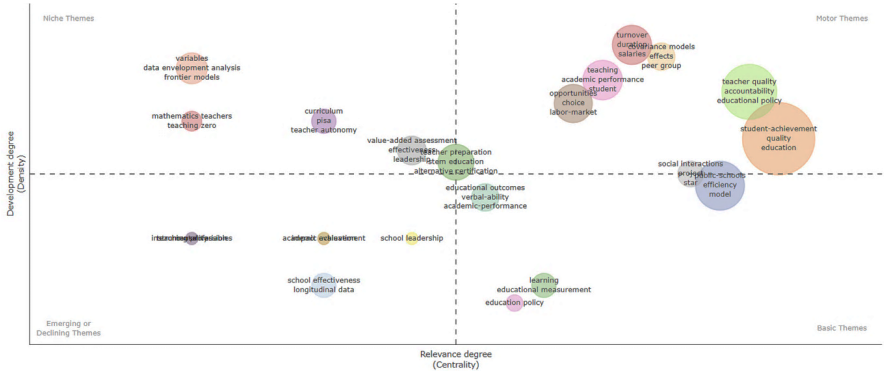
This analysis highlights themes likely to be central to future educational research, as shown in the thematic map in Figure 6, which organizes clusters along dimensions of centrality and density. Quadrant 1 (motor themes) includes topics that are both central and well developed; Quadrant 2 (niche themes) comprises specialized but less influential topics; Quadrant 3 (emerging or declining themes) captures underdeveloped or evolving areas; and Quadrant 4 (foundational themes) contains well-established and widely integrated knowledge in the field.

Figure 6 maps the central themes of studies citing Hanushek (1971), highlighting a dominant focus on school factors, teacher effectiveness, and resource efficiency in public Education. Well-developed driver themes in Quadrant 1, such as student achievement, teacher quality, and covariance modeling, have influenced educational evaluation research (Hanushek, 2011; Hanushek; Rivkin; Schiman, 2016; Levy *et al*., 2023; Rivkin; Hanushek; Kain, 2005). Emerging themes in Quadrant 3, including school leadership, school effectiveness, and longitudinal data, are at an early stage, given the increased frequency of these terms in recent

[3] Thematic maps were created using all studies from each database without applying any filters, as the aim is to identify the general directions of each group of studies.

years. Although these terms have been present in the educational literature, their integration and development in discussions of educational assessment and VA are still in their early stages, indicating the need for more in-depth exploration of these concepts in the current context (Kuzmanic *et al*., 2024; Levy *et al*., 2020; Valle; Lillejord, 2023).

Figure 6 – Thematic map of studies citing Hanushek (1971), 1971–2024



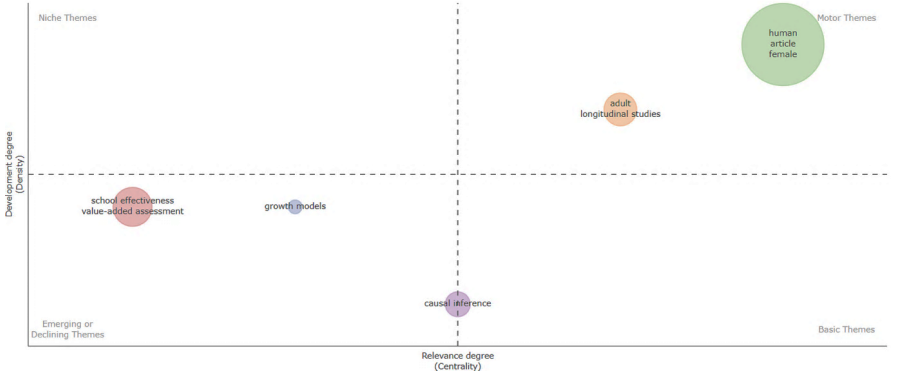
Source: The authors (2024)

The shift toward systemic processes and temporal dynamics identified in Quadrant 3 reflects growing evidence that robust, longitudinal models are essential for capturing educational complexity and ensuring stable VA estimates (Ferrão; Couto, 2013; Morera *et al*., 2009). Recent research also highlights the increasing role of school leadership in student outcomes (Muñoz; Prem, 2024), demonstrating how managerial effectiveness is associated with student outcomes. Together, these trends suggest that future research will increasingly emphasize school-level factors, managerial effectiveness, and longitudinal analyses to advance educational evaluation and VA modeling.

The thematic analysis of studies citing Bryk and Weisberg (1976) identifies both established and emerging research areas. Well-developed themes in Quadrant 1 (Figure 7) emphasize human factors and longitudinal analysis (Bartolucci *et al*., 2011, 2023), while causal inference occupies an intermediate position between Quadrants 3 and 4, reflecting its growing prominence (Raudenbush; Schwartz, 2020). Additional emerging themes in Quadrant 3, such as school effectiveness and VA assessment, remain important and evolving areas of research despite their extensive prior discussion (Sammons, 2007; Santelices *et al*., 2017;

Soares *et al*., 2017; Strand, 2010, 2016). However, they continue to evolve, incorporating new approaches and perspectives that warrant further exploration and refinement.

Figure 7 – Thematic map of studies citing Bryk and Weisberg (1976), 1976–2024



Source: The authors (2024)

The thematic analysis of studies citing Hanushek (1971) and Bryk & Weisberg (1976) reveals a field predominantly focused on traditional factors such as teacher effectiveness, while also showing increasing attention to emerging areas, including school leadership, longitudinal data, and causal inference. This shift toward a more comprehensive and methodologically sophisticated understanding of school effectiveness is particularly relevant in contexts like Brazil, where longitudinal approaches are essential for producing stable and equitable VA measures (Brooke *et al*., 2014; Ferrão; Couto, 2013; Soares *et al*., 2017). At the same time, the literature highlights persistent limitations related to the stability, interpretability, and high-stakes application of VA models, emphasizing the need to complement quantitative analyses with qualitative evidence and engagement with school actors, such as teachers, students and families. Together, these trends underscore the potential of VA as a contextualized tool for educational evaluation, while informing educational policies (Amrein-Beardsley *et al*., 2013; Braun, 2013).

4 Conclusion

In response to the research question, the findings show that the literature citing Hanushek (1971) and Bryk & Weisberg (1976) has matured substantially, revealing

a diversification of methodological approaches, emerging research trends, and persistent gaps. By combining bibliometric analysis, collaboration network mapping, and thematic synthesis of studies published between 1971 and 2024 that cite Hanushek (1971) and Bryk and Weisberg (1976), this review traces the evolution of VA models, documenting key methodological developments and providing evidence to inform educational policy.

Despite methodological advances, significant gaps remain in the VA literature, including the limited representation of studies from developing countries and the need to explore new thematic areas such as school leadership, longitudinal data, school effectiveness, and causal inference. A more detailed examination of the contextual variables that influence educational effectiveness has also been identified as necessary to adapt interventions to different contexts. The integration of quantitative VA measures with qualitative and stakeholder-based evidence (including teachers, students, and families) remains underdeveloped, as do analyses of teacher training, retention, and equity in resource allocation.

Addressing these gaps constitutes a key agenda for future research aimed at advancing both the methodological rigor and the policy relevance of VA analyses. As a scoping review, this study is limited by its emphasis on breadth rather than depth, the absence of formal quality appraisal, reliance on seminal citations, and single-author screening, which warrants caution in generalizing the findings. In this context, we emphasize the importance of conducting a Systematic Literature Review (SLR) to consolidate evidence and strengthen the methodological and policy relevance of VA analyses.

Based on the patterns identified in the literature, these findings contribute to educational policy and practice by emphasizing the critical role of school leadership in school culture and improving student outcomes. It highlights the value of collecting and analyzing longitudinal data to monitor academic progress and inform evidence-based policy adjustments. The results also underscore the importance of using VA assessments to guide practice, enhance educational effectiveness, and promote equity, including the allocation of resources to underserved schools and the implementation of targeted programs. In contexts such as Brazil, where persistent social inequalities and institutional constraints affect educational outcomes, these insights are particularly relevant for designing interventions that are both equitable and effective. At the same time, generating robust evidence capable of informing public policy requires adequate research infrastructure, funding, and access to data. Collaboration between researchers and

policymakers is encouraged to bridge research and practice, and the training and retention of qualified teachers is identified as essential for improving instructional quality. We hope this work encourages research and discussion to foster more equitable and effective Education systems.

Evolução conceitual e desafios metodológicos do valor agregado na avaliação educacional: uma revisão de escopo 1971–2024

Resumo

Modelos de Valor Agregado (VA) são amplamente utilizados para avaliar a eficácia educacional; contudo, seu desenvolvimento conceitual e metodológico ainda não foi sistematizado de forma satisfatória. Este estudo conduz uma revisão de escopo, complementada por uma análise bibliométrica exploratória de trabalhos publicados entre 1971 e 2024, com foco nas contribuições fundacionais de Hanushek (1971), na economia da educação, e de Bryk e Weisberg (1976), na estatística educacional. Os resultados evidenciam lacunas persistentes na literatura, especialmente a escassez de evidências provenientes de países em desenvolvimento e a limitada atenção dedicada ao papel da liderança escolar e da inferência causal nos modelos de VA. Ademais, a revisão indica uma progressiva diversificação das abordagens metodológicas, acompanhada de um maior vínculo com políticas de responsabilização. Do ponto de vista das políticas públicas, os achados apontam para a necessidade de aplicar modelos de VA de forma sensível ao contexto, fortalecer a liderança escolar, aprimorar os sistemas de dados longitudinais e promover maior articulação entre pesquisadores e formuladores de políticas, com vistas ao desenvolvimento de intervenções educacionais mais eficazes e equitativas.

Palavras-chave: *Avaliação Educacional. Eficácia Educacional. Eficácia Escolar. Equidade Social. Revisão de Escopo.*

Evolución conceptual y desafíos metodológicos del valor añadido en la evaluación educativa: una revisión de alcance 1971–2024

Resumen

Los modelos de valor añadido (VA) se utilizan ampliamente para evaluar la eficacia educativa; no obstante, su desarrollo conceptual y metodológico aún no ha sido sistematizado de forma suficiente. Este estudio realiza una revisión de alcance, complementada con un análisis bibliométrico exploratorio de trabajos publicados entre 1971 y 2024, centrado en las contribuciones fundacionales de Hanushek (1971) en economía de la educación y de Bryk y Weisberg (1976) en estadística educativa. Los resultados evidencian lagunas persistentes en la literatura, especialmente la escasa evidencia procedente de países en desarrollo y la limitada atención prestada al liderazgo escolar y a la inferencia causal en los marcos de los modelos VA. Asimismo, la revisión muestra una progresiva diversificación de los enfoques metodológicos, acompañada de una mayor vinculación con las políticas de rendición de cuentas. Desde una perspectiva de política pública, los hallazgos ponen de relieve la necesidad de aplicar los modelos

VA de manera sensible al contexto, reforzar el liderazgo escolar, mejorar los sistemas de datos longitudinales y promover una colaboración más estrecha entre investigadores y responsables públicos, con el objetivo de impulsar intervenciones educativas más eficaces y equitativas.

Palabras clave: *Evaluación Educativa. Eficacia Educativa. Eficacia Escolar. Equidad Social. Revisión de Alcance.*

References

AERA. Statement on Use of Value-Added Models (VAM) for the Evaluation of Educators and Educator Preparation Programs. *Educational Researcher,* [S. l.], v. 44, n. 8, p. 448–452, nov. 2015.

ALBERNAZ, Ângela; FERREIRA, Francisco H. G.; FRANCO, Creso. Qualidade e eqüidade no ensino fundamental brasileiro. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 32, n. 3, dez. 2002.

ALVES, Maria Teresa Gonzaga; FRANCO, Creso. A pesquisa em eficácia escolar no Brasil: evidências sobre o efeito das escolas e fatores associados à eficácia escolar. *In*: BROOKE, N.; SOARES, José Francisco (orgs.). *Pesquisa em eficácia escolar*: origem e trajetórias. Belo Horizonte: Editora UFMG, 2008. v. 22 p. 482–500.

AMREIN-BEARDSLEY, Audrey *et al.* Value-Added Model (VAM): research for educational policy: framing the issue. *Education Policy Analysis Archives*, [S. l.], v. 21, p. 4, 27 jan. 2013.

AMREIN-BEARDSLEY, Audrey. A Florida Teacher wrongfully terminated? Alleged (mis)uses of Value-Added Model (VAM) estimates for high-stakes teacher evaluation decisions. *Educational Assessment, Evaluation and Accountability*, [S. l.], v. 35, n. 3, p. 313–352, ago. 2023.

AMREIN-BEARDSLEY, Audrey; LAWTON, Kerry; RONAN, Katherine. An elusive policy imperative: Data and methodological challenges when using growth in student achievement to evaluate teacher education programs' "value-added". *Teaching Education*, [S. l.], v. 28, n. 3, p. 296–316, 2017.

AMREIN-BEARDSLEY, Audrey; PIVOVAROVA, Margarita; GEIGER, Tray J. Value-added models: What the experts say. *Phi Delta Kappan*, [S. l.], v. 98, n. 2, p. 35–40, out. 2016.

ARIA, Massimo; CUCCURULLO, Corrado. Bibliometrix: an R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, [S. l.], v. 11, n. 4, p. 959–975, nov. 2017.

ARKSEY, Hilary; O'MALLEY, Lisa. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, [S. l.], v. 8, n. 1, p. 19–32, fev. 2005.

ARMOR, David J.; MARKS, Gary N.; MALATINSZKY, Aron. The impact of school SES on student achievement: evidence from U.S. statewide achievement data. *Educational Evaluation and Policy Analysis*, [S. l.], v. 40, n. 4, p. 613–630, dez. 2018.

*BARTOLUCCI, Francesco; PENNONI, Fulvia; VITTADINI, Giorgio. Assessment of school performance through a multilevel latent Markov Rasch model. *Journal of Educational and Behavioral Statistics*, [S. l.], v. 36, n. 4, p. 491–522, ago. 2011.

*BARTOLUCCI, Francesco; PENNONI, Fulvia; VITTADINI, Giorgio. A causal latent transition model with multivariate outcomes and unobserved heterogeneity: application to human capital development. *Journal of Educational and Behavioral Statistics*, [S. l.], v. 48, n. 4, p. 387–419, ago. 2023.

BERTONI, E. *et al.* Teacher selection instruments and teacher value-added: evidence from Peru. *Educational Evaluation and Policy Analysis*, [S. l.], v. 46, n. 1, p. 53–81, 2024.

BRAUN, Henry. Value-added modeling and the power of magical thinking. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 21, n. 78, p. 115–130, mar. 2013.

BRAUNACK-MAYER, Annette J. *et al.* Student and staff perspectives on the use of big data in the tertiary education sector: a scoping review and reflection on the ethical issues. *Review of Educational Research*, [S. l.], v. 90, n. 6, p. 788–823, dez. 2020.

BROOKE, N. *et al.* Modeling of the growth of learning in the early years with longitudinal data of GERES research. *Educação e Pesquisa*, Rio de Janeiro, v. 40, n. 1, p. 77–94, 2014.

*BRYK, Anthony S.; WEISBERG, Herbert I. Value-added analysis: A dynamic approach to the estimation of treatment effects. *Journal of Educational Statistics*, [S. l.], v. 1, n. 2, p. 127, 1976.

*BRYK, Anthony S.; WEISBERG, Herbert I. Use of the nonequivalent control group design when subjects are growing. *Psychological Bulletin*, [S. l.], v. 84, n. 5, 1977.

CHARALAMBOUS, Evi; KYRIAKIDES, Leonidas; CREEMERS, Bert P. M. Promoting quality and equity in socially disadvantaged schools: A group- randomisation study. *Studies in Educational Evaluation*, [S. l.], v. 57, p. 42–52, jun. 2018.

*CHETTY, Raj; FRIEDMAN, John N.; ROCKOFF, Jonah E. Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, [S. l.], v. 104, n. 9, p. 2593–2632, 1 set. 2014a.

*CHETTY, Raj; FRIEDMAN, John N.; ROCKOFF, Jonah E. Measuring the impacts of teachers II: teacher value-added and student outcomes in adulthood. *American Economic Review*, [S. l.], v. 104, n. 9, p. 2633–2679, 1 set. 2014b.

*EMSLANDER, Valentin *et al.* Value-added scores show limited stability over time in primary school. *PLOS ONE*, [S. l.], v. 17, n. 12, p. e0279255, 28 dez. 2022.

EVERSON, Kimberlee C. Value-added modeling and educational accountability: Are we answering the real questions?. *Review of Educational Research*, [S. l.], v. 87, n. 1, p. 35–70, fev. 2017.

*FERRÃO, M. E.; COUTO, A. Value-added indicator and topics on consistency and stability: An application to Brazil. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 21, n. 78, p. 131–164, 2013.

FERRÃO, M. E.; COUTO, A. The use of a school value-added model for educational improvement: a case study from the Portuguese primary education system. *School Effectiveness and School Improvement*, [S. l.], v. 25, n. 1, p. 174–190, 2014.

*FERRÃO, Maria Eugénia. Estudo longitudinal sobre eficácia diferencial e equidade social no Brasil. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, Madrid, v. 20, n. 1, 1 jan. 2022a.

*FERRÃO, Maria Eugénia. The evaluation of students' progression in lower secondary education in Brazil: Exploring the path for equity. *Studies in Educational Evaluation*, [S. l.], v. 75, dez. 2022b.

FERRÃO, Maria Eugénia; ALVES, Maria Teresa Gonzaga. The relationship between students' socio-demographics and the probability of grade repetition in Brazilian primary education: is it decreasing over time? *Large-scale Assessments in Education*, [S. l.], v. 11, n. 1, p. 12, 4 abr. 2023.

GOLDSTEIN, Harvey. Multilevel modelling of educational data. *In*: COURGEAU, Daniel (org.). *Methodology and Epistemology of Multilevel Analysis*. Dordrecht: Springer Netherlands, 2003. p. 25–42.

*HANUSHEK, Eric. Teacher characteristics and gains in student achievement: estimation using micro data. *The American Economic Review*, [S. l.], v. 61, n. 2, p. 280–288, 1971.

HANUSHEK, Eric A. The economic value of higher teacher quality. *Economics of Education Review*, [S. l.], v. 30, n. 3, p. 466–479, jun. 2011.

HANUSHEK, Eric A. Testing, accountability, and the american economy. *The Annals of the American Academy of Political and Social Science*, [S. l.], v. 683, n. 1, p. 110–128, maio 2019.

HANUSHEK, Eric A. *et al.* Long-run trends in the US SES: achievement gap. *Education Finance and Policy*, [S. l.], v. 17, n. 4, p. 608–640, 2022.

HANUSHEK, Eric A.; RIVKIN, Steven G. The quality and distribution of teachers under the No Child Left Behind Act. *Journal of Economic Perspectives*, [S. l.], v. 24, n. 3, p. 133–150, 1 ago. 2010.

*HANUSHEK, Eric A.; RIVKIN, Steven G.; SCHIMAN, Jeffrey C. Dynamic effects of teacher turnover on the quality of instruction. *Economics of Education Review*, [S. l.], v. 55, p. 132–148, dez. 2016.

*HANUSHEK, Eric A.; WOESSMANN, Ludger. How much do educational outcomes matter in OECD countries? *Economic Policy*, [S. l.], v. 26, n. 67, p. 427–491, jul. 2011.

*HEDGES, Larry V.; LAINE, Richard D.; GREENWALD, Rob. An exchange: Part I: does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher*, [S. l.], v. 23, n. 3, p. 5, abr. 1994.

HOX, Joop; MOERBEEK, Mirjam; VAN DE SCHOOT, Rens. *Multilevel analysis: Techniques and applications*. [S. l.]: Routledge, 2017.

*JACKSON, C. Kirabo; ROCKOFF, Jonah E.; STAIGER, Douglas O. Teacher effects and teacher-related policies. *Annual Review of Economics*, [S. l.], v. 6, n. 1, p. 801–825, 1 ago. 2014.

KOEDEL, Cory; MIHALY, Kata; ROCKOFF, Jonah E. Value-added modeling: A review. *Economics of Education Review*, [S. l.], v. 47, p. 180–195, ago. 2015.

KUZMANIC, Danilo *et al.* Inequalities in learning loss during the COVID-19 pandemic in Chile: the significance of school effectiveness. *School Effectiveness and School Improvement*, [S. l.], v. 35, n. 2, p. 95–115, 2 abr. 2024.

KYRIAKIDES, Leonidas *et al.* Improving quality and equity in schools in socially disadvantaged areas. *Educational Research*, [S. l.], v. 61, n. 3, p. 274–301, 3 jul. 2019.

LARREA, Verónica Azpillaga; ROJAS, Isabel Bartau; HERNÁNDEZ, Luis Lizasoain. Mejora escolar y coeducación en centros de secundaria de la comunidad autónoma vasca: Implicaciones para la orientación educativa. *REOP - Revista Española de Orientación y Psicopedagogía*, [S. l.], v. 35, n. 1, p. 45–62, 23 abr. 2024.

*LEVY, Jessica *et al.* Methodological issues in value-added modeling: An international review from 26 countries. *Educational Assessment, Evaluation and Accountability*, [S. l.], v. 31, n. 3, p. 257–287, ago. 2019.

*LEVY, Jessica *et al.* Contrasting classical and machine learning approaches in the estimation of value-added scores in large-scale educational data. *Frontiers in Psychology*, [S. l.], v. 11, p. 2190, 21 ago. 2020.

*LEVY, Jessica *et al.* How sensitive are the evaluations of a school's effectiveness to the selection of covariates in the applied value-added model? *Educational Assessment, Evaluation and Accountability*, [S. l.], v. 35, n. 1, p. 129–164, fev. 2023.

MARDER, Michael. Measuring teacher quality with value-added modeling. *Kappa Delta Pi Record*, [S. l.], v. 48, n. 4, p. 156–161, 2012.

MARKS, Gary N. Is SES really that important for educational outcomes in Australia? A review and some recent evidence. *The Australian Educational Researcher*, [S. l.], v. 44, n. 2, p. 191–211, 2017.

MARTÍNEZ-ABAD, Fernando; GAMAZO, Adriana; RODRÍGUEZ-CONDE, María-José. Educational data mining: Identification of factors associated with school effectiveness in PISA assessment. *Studies in Educational Evaluation*, [S. l.], v. 66, p. 100875, set. 2020.

*MORERA, María *et al.* Selección de escuelas de alta y baja eficacia en Baja California (México). *Revista Electrónica de Investigación Educativa*, [S. l.], v. 19, n. 2, p. 38–53, 2009.

MORGANSTEIN, David; WASSERSTEIN, Ron. ASA statement on value-added models. *Statistics and Public Policy*, [S. l.], v. 1, n. 1, p. 108–110, 22 dez. 2014.

MUÑOZ, P.; PREM, M. Managers' productivity and recruitment in the public sector. *American Economic Journal: Economic Policy*, [S. l.], v. 16, n. 4, p. 223–253, 2024.

*MUTHEN, Bengt O.; CURRAN, Patrick J. General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological methods*, [S. l.], v. 2, n. 4, 1997.

NEWTON, Xiaoxia A. *et al.* Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Education Policy Analysis Archives*, [S. l.], v. 18, n. 23, set. 2010.

*NYE, Barbara; KONSTANTOPOULOS, Spyros; HEDGES, Larry V. How large are teacher effects? *Educational Evaluation and Policy Analysis*, [S. l.], v. 26, n. 3, p. 237–257, set. 2004.

PALARDY, Gregory J. Differential school effects among low, middle, and high social class composition schools: a multiple group, multilevel latent growth curve analysis. *School Effectiveness and School Improvement*, [S. l.], v. 19, n. 1, p. 21–49, mar. 2008.

PHAM, Mai T. *et al.* A scoping review of scoping reviews: advancing the approach and enhancing the consistency. *Research Synthesis Methods*, [S. l.], v. 5, n. 4, p. 371–385, dez. 2014.

PRIMI, Ricardo; FERRÃO, Maria Eugénia; ALMEIDA, Leandro S. Fluid intelligence as a predictor of learning: A longitudinal multilevel approach applied to math. *Learning and Individual Differences*, [S. l.], v. 20, n. 5, p. 446–451, out. 2010.

RAUDENBUSH, Stephen W.; SCHWARTZ, Daniel. Randomized experiments in education, with implications for multilevel causal inference. *Annual Review of Statistics and Its Application*, [S. l.], v. 7, n. 1, p. 177–208, 9 mar. 2020.

REYNOLDS, David *et al.* Educational Effectiveness Research (EER): A state-of-the-art review. *School Effectiveness and School Improvement*, [S. l.], v. 25, n. 2, p. 197–230, 3 abr. 2014.

*RIVKIN, Steven G.; HANUSHEK, Eric A.; KAIN, John F. Teachers, schools, and academic achievement. *Econometrica*, [S. l.], v. 73, n. 2, p. 417–458, mar. 2005.

SAMMONS, Pam. School effectiveness and equity: Making connections. *Reading: CfBT*, p. 1–22, 2007.

SAMMONS, Pamela; HILLMAN, Josh; MORTIMORE, Peter. *Key characteristics of effective schools*: a review of school effectiveness research. [S. l.]: B & MBC Distribution Services, 1995.

SANTELICES, Maria Veronica *et al.* Two teacher quality measures and the role of context: Evidence from Chile. *Educational Assessment, Evaluation and Accountability*, [S. l.], v. 29, n. 2, p. 111–146, maio 2017.

SAUNDERS, Lesley. A brief history of educational "value added": How did we get to where we are? *School Effectiveness and School Improvement*, [S. l.], v. 10, n. 2, p. 233–256, 1 jun. 1999.

SAUNDERS, Lesley. Understanding schools' use of 'value added' data: the psychology and sociology of numbers. *Research Papers in Education*, [S. l.], v. 15, n. 3, p. 241–258, jan. 2000.

SINGER, Judith D. Using SAS PROC MIXED to Fit multilevel models, hierarchical models, and Individual growth models. *Journal of Educational and Behavioral Statistics*, [S. l.], v. 23, n. 4, p. 323–355, dez. 1998.

SIRIN, Selcuk R. Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, [S. l.], v. 75, n. 3, p. 417–453, set. 2005.

SOARES, T. M. *et al.* Value-added models for the measurement of the effectiveness of GERES schools. *Ensaio*, Rio de Janeiro, v. 25, n. 94, p. 59–89, 2017.

STRAND, Steve. Do some schools narrow the gap? Differential school effectiveness by ethnicity, gender, poverty, and prior achievement. *School Effectiveness and School Improvement*, [S. l.], v. 21, n. 3, p. 289–314, set. 2010.

STRAND, Steve. The white british-black caribbean achievement gap: Tests, tiers and teacher expectations. *British Educational Research Journal*, [S. l.], v. 38, n. 1, p. 75–101, fev. 2012.

STRAND, Steve. Do some schools narrow the gap? Differential school effectiveness revisited. *Review of Education*, [S. l.], v. 4, n. 2, p. 107–144, jun. 2016.

THOMAS, Sally; PENG, Wen Jung; GRAY, John. Modelling patterns of improvement over time: value added trends in english secondary school performance across ten cohorts. *Oxford Review of Education*, [S. l.], v. 33, n. 3, p. 261–295, jun. 2007.

TWERSKY, Sylvia E. Do state laws reduce uptake of Medicaid/CHIP by U.S. citizen children in immigrant families: evaluating evidence for a chilling effect. *International Journal for Equity in Health*, [S. l.], v. 21, n. 1, p. 50, dez. 2022.

VALLE, Roald; LILLEJORD, Sølvi. A new regime of understanding. School leadership in Norwegian education policy (1990–2017). *Nordic Journal of Studies in Educational Policy*, [S. l.], v. 9, n. 2, p. 144–160, 4 maio 2023.

CC BY

Information about the authors

Kelly Alice Barbosa Macêdo: PhD in Economics, Graduate Program in Economics, Federal University of Juiz de Fora, Juiz de Fora, MG, Brazil. Contact: kellyabmacedo@gmail.com

Maria Eugénia Ferrão: Assistant Professor of Faculty of Sciences, University of Beira Interior, Covilhã, Portugal; and CEMAPRE and ISEG Research, Lisboa, Portugal. Contact: meferrao@ubi.pt

Marcel de Toledo Vieira: Full Professor of Department of Statistics and Research Programs in Applied Economics and in Public Education Governance and Evaluation, Federal University of Juiz de Fora, Juiz de Fora, MG, Brazil. Contact: marcel.vieira@ufjf.br

Authors' contribution: Kelly Alice Barbosa Macêdo: Conceptualization, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing and Funding acquisition. Maria Eugénia Ferrão: Conceptualization, Writing - Original Draft, Writing - Review & Editing, Supervision and Funding acquisition. Marcel de Toledo Vieira: Conceptualization, Writing - Original Draft, Writing - Review & Editing, Supervision and Funding acquisition.

Data: The dataset supporting the findings of this study has been published as open access and is available on Zenodo at https://doi.org/10.5281/zenodo.18665963

Conflicts of interest: The authors declare that they have no commercial or associative interests that could be construed as a conflict of interest in relation to this manuscript.

Funding: The first author is financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) - Brazil - Finance Code 001. The second author is funded by national funds through FCT- Fundação para a Ciência e a Tecnologia, I.P., in the framework of the ISEG Research project UID/06522/2025, and in the framework of FCT project 2023.18332.ICDT. The third author would like to thank the National Council for Scientific and Technological Development of Brazil for the support (process number 311445/2022-5).

Use of AI: During the preparation of this manuscript, ChatGPT (GPT-5 mini) was used for grammatical revision of the text. Following its use, the authors carefully reviewed and edited the content as necessary and assumed full responsibility for the final version of the manuscript.

Editors responsible for this article:
Alvaro Chrispino (Associate Editor)
Sidney Mello (Associate Editor)