Culture, Education, and Future
Volume 3, Number 1, 2025, 3-25
DOI: 10.70116/2980274131

SYMPHONY
ISSN: 2980-2741

Article

OPEN ACCESS

Interdependent versus independent research: An overdue shift in perspective

Russ Marion [1], **Xiaorang Ma** [2]

[1] Clemson University, College of Education, Clemson, United States
[2] Langfang Normal University, Department of Education, Langfang, Hebei, China

Abstract

Traditionally, research in the social sciences has focused on the role of individual attitudes, skills, and dyadic relationships in shaping educational outcomes. However, less attention has been paid to the influence of broader patterns of social interaction, particularly within school contexts. This study demonstrated how interaction-based data could be generated and analyzed to better capture these dynamics. Drawing on data collected from 2,682 students and 118 teachers across 10 schools, we applied an AI-driven machine learning algorithm to examine the effects of interactive dynamics on student achievement. Results indicate that while socioeconomic status (SES) remains a consistent predictor of student test scores, the most significant effects stem from interactions related to social capital, the diversity of information each person has access to, and to the degree of effort one invests in network dynamics. These findings highlight the value of incorporating social network structures into educational research and suggest that interactive dynamics within school communities may play a pivotal role in shaping student achievement.

Article History
Received 23.08.2024
Accepted 01.04.2025

Keywords
Postmodern analysis; network analysis; AI; machine learning; student achievement

Group Level Rather Than Individual or Dyadic Research: An Overdue Shift on Perspective

Educational research has traditionally focused on individual-level factors and dyadic relationships in understanding student outcomes. However, this individualistic approach may be limiting our ability to capture the complex, interactive dynamics that actually drive educational success. Given these limitations, we propose a fundamental paradigm shift toward examining collective behaviors and network-level interactions within educational settings. Two converging developments make this shift both timely and necessary. First, organizational behavior research over the past two decades has increasingly demonstrated the superiority of collective over independent approaches to understanding workplace dynamics. Major journals including *The Leadership Quarterly* (Cullen-Lester & Yammarino, 2016) and *Organization Science* (Bailey et al., 2022) have devoted special issues to this emerging perspective. Second, the rapid advancement of artificial intelligence and machine learning technologies in recent years provides powerful new analytical tools for examining complex, interactive systems (von Krogh et al., 2023). These developments suggest that educational

Corresponding Author Russ Marion marion2@clemson.edu Clemson University, College of Education, 414D Tillman Hall, Clemson, SC 9634, United States

research is ready for a methodological transformation—one that moves beyond traditional assumptions about independent individual behaviors toward a more holistic understanding of how educational outcomes emerge from interactions within school communities.

We offer, therefore, a modest suggestion: Let's begin to reframe current assumptions which claim that attitudes, preferences, and behaviors of educators and others are the individual's independent property, largely free of external influence, and thus subject to the reduction of social behaviors into component parts to be analyzed. This perspective aligns with the foundational principles of post-positivism, which emphasizes objective measurement and hypothesis testing (Popper et al., 1961). Postmodernism, as we use it, transitions instead to an assumption that attitudes, preferences, and such are influenced by a complex of people's interactions and that they should be understood holistically. Teacher's attitudes about their principal, policy, school culture, even safety are shaped by complex interactions among, students, parents, and others (Hendrickx, 2012; Morinaj et al., 2023; Supovitz et al., 2010).

This new research epistemology will need new research methodologies. Several possibilities are currently available, but we will utilize social network analysis (SNA), which analyzes networks of interactive, adaptive agents, and scores agents on the nature of their interactions (Borgatti et al., 2009; Jackson et al., 2017). We demonstrate in this paper how the measures produced by SNA can be analyzed with boosted tree, a statistical methodology used in artificial intelligence decision making. Boosted tree sorts through datasets of agents and identifies those that best distinguish the absence or presence of a target outcome (Ponomareva et al., 2017). Boosted tree analysis was selected because it is highly accurate, robust with complex datasets, and effectively handles curvilinear relationships (Bastos, 2022; Gabidolla et al., 2025; Hastie et al., 2017). In a dataset of measures and achievement test scores from elementary-age children and their teachers, we sought to determine whether there were networked measures of interactions among teachers that could identify which measures distinguish among teachers whose students had high test scores and those that did not. Essentially, it's similar to creating a checklist of network traits that identify teachers likely to produce high test scores. The more important goal, however, is to illustrate how we can reconceptualize the nature of statistical research and assumptions, and to use modern tools on those reconceptualized research designs, thus moving assumptions about research forward. The goal, then, is to explore strategies for examining research questions from different perspectives than are currently used.

Accordingly, there are two elements of this research: The first is the generation of interaction-level data, or mechanisms, and the second is evaluation of the effects of holistic mechanisms on independent test scores using machine learning, an AI-related statistic. More broadly, we argue that the two epistemologies underlying these elements, postmodernism, which underlies the generation of mechanisms, and statistical positivism, which informs the evaluation of test scores, or variables. can be reconciled and used together in research studies. Our ultimate intent is to demonstrate how researchers can expand their research knowledge using holistic, interactive system and emerging strategies for evaluating such systems. The study we present relative to this goal is driven by the following research question:

> RQ: Do interactive-level mechanisms among elementary school teachers influence their students' standardized math achievement scores?

Research on Student Achievement

Despite decades of research, understanding what drives student academic achievement remains a complex and evolving challenge in the field of education. Prior studies have established socioeconomic status (SES) and family background as dominant predictors of academic success. Seminal works by Coleman et al. (1966) and Jencks (1972) showed that students from higher SES backgrounds consistently outperform their lower SES peers, even after accounting for school-level factors. Subsequent research, such as Alspaugh (1996), reaffirmed that external factors tied to SES, like parental education and income, explain more variance in achievement than school-based interventions. Contemporary studies confirm the enduring weight of SES across national contexts: Yeung et al. (2022), for example, showed that family income and parental education levels significantly shaped reading achievement through mediators Mover, Su-Russell and Russell (2021) demonstrate that SES-related parenting practices indirectly influence children's persistence and early academic competencies.

From a related perspective, the school effectiveness research has demonstrated that school-level factors play a significant role in shaping student outcomes, beyond individual or family background. Since the 1990s, numerous macro-level contextual factors, such as school socioeconomic composition, school size, and urban–rural location, have been identified as important influences on school effectiveness (Teddlie & Stringfield, 1993; Virgilio et al., 1991). Empirical studies have shown that student achievement is shaped not only by socioeconomic background but also by school-level conditions such as principal leadership and teacher quality.

A central insight from this research is that strong principal leadership is a defining characteristic of high-performing schools serving disadvantaged students (Bossert et al., 1982; Edmonds, 1979; Sammons et al., 1995). Subsequent research further revealed that leadership effects are largely indirect, operating through their influence on teacher collaboration, professional culture, and student engagement (Hallinger & Heck, 1998). More recent scholarship emphasizes a shift toward leadership that is context-sensitive, combining transformational and instructional elements, and, in some cases, participatory, and distributed (e.g., Liu et al., 2020; Ma & Marion, 2024, 2025; Spillane, 2005). In addition, the role of school climate, particularly perceptions of safety, academic press, and relational support, has been shown to be crucial for fostering teacher morale and promoting equitable learning outcomes (Thapa et al., 2013; Yeung et al., 2022).

Beyond structural and organizational determinants, research increasingly recognizes the role of individual agency and psychosocial processes in shaping achievement. Theoretical advances in academic motivation (Eccles & Wigfield, 2002), self-regulated learning (Zimmerman, 2002), and achievement emotions (Pekrun, 2006) have reframed academic success as the outcome of dynamic interactions among beliefs, behaviors, and emotional capacities, rather than simple function of cognitive ability. Psychological constructs such as growth mindset, academic self-concept, and goal orientation have been identified as proximal drivers of student engagement and persistence (e.g., Bartels & Magun-Jackson, 2021; Bouchet & Kizilcec, 2020; Karaman & Watson, 2020; Olivier et al., 2019).

Recent studies extend these frameworks by revealing how internal dispositions operate in tandem with social context. For example, perceived academic competence and school connectedness have been shown to mediate long-term achievement, even after accounting for

SES disparities (Nunes et al., 2023). Similarly, interpersonal relationships within the school context, such as supportive teacher–student interactions and positive peer relationships, have been associated with increased student engagement, motivation, and, ultimately, improved academic achievement (Hughes et al., 2008; Wang & Eccles, 2012; Wentzel et al., 2016).

Yet finding an appreciable link between the actions of educational professionals and testing outcomes, after accounting for contextual variables such as socio-economic status and ethnicity, has been vexing. Indeed, such studies typically explain less than 12% of the variation in test scores (see Louis et al., 2010, Table 1.1.5, for example). School leaders have tried to overwhelm this vexation with instructional leadership initiatives such as time-on-task and high expectations (Hallinger, 2007; Klapp et al., 2024) or leadership strategies like transformational leadership (Esposito & Bauer, 2022; Leithwood & Jantzi, 2005, 2006) and leader-member exchange (LMX) relationships (Somech, 2010; Somech & Wenderow, 2006; Yeung et al., 2022), but, still, a compelling link between schooling characteristics and de-contextualized test scores remains elusive.

Social Network Dynamics and Academic Achievement

We propose that researchers experience difficulties defining a consistent relationship between schooling variables and student achievement because the assumptions underlying their analyses are often limiting. The prevailing assumption, one which has dominated organizational studies generally over the past century, is that productivity follows individuals who possess exceptional skills, knowledge, and attitudes; this is a human capital assumption (Pil & Leana, 2009; Tan, 2014). The assumption we pursue is that outcomes are influenced more by social capital and group dynamics (Louis & Marks, 1998; Marks & Printy, 2003; Pearce et al., 2008; Pil & Leana, 2009; Vescio et al., 2008).

We frame this assumption with complexity theory. Complexity theorists examine the degree to which people engage in, and are influenced by, the quantity, frequency, nature and speed of information flow across networks; how leaders enable such flow; and how complex dynamics, which emerge from interactive processes, enhance the productive capacities of groups and individuals (Bedeian & Hunt, 2006; Friedkin & Slater, 1994; Hunt & Dodge, 2000; Moolenaar et al., 2012). Interactive pathways and interaction facilitate access to resources (Dess & Shaw, 2001; Frank et al., 2004; Li, 2013) and enable collective dynamics (Marion & Gonzales, 2013) that convert information flow into productive outcomes. These assumptions raise the possibility that student test scores may be influenced by interactional and group-level processes. In light of this, understanding the determinants of academic achievement requires moving beyond static, individual traits to include the dynamic structure of relationships within school communities. The following sections examine how network configurations, such as the flow of information and patterns of social connection, may shape educational outcomes.

Centrality and Academic Success

Centrality, an SNA measure of the number of links people (called agents in network analysis) experience, has been consistently identified as a crucial determinant of students' academic achievement. Specifically, students occupying central positions in peer networks, indicated by high in-degree and closeness centralities, often demonstrate improved academic outcomes (López-Sánchez et al., 2023; Williams et al., 2019). Research suggests that central students have increased access to informational and emotional resources, academic support, and guidance, which collectively contribute to higher academic performance. Furthermore, such positions enhance students' social visibility, facilitating the diffusion of effective learning behaviors within the network and thus creating a positive reinforcement cycle (Bruun & Brewe, 2013). However, Bond et al. (2017) present an alternative perspective, indicating that while higher-achieving students frequently occupy central positions, centrality itself might not necessarily lead to enhanced achievement. Instead, network centrality could reflect social recognition of pre-existing academic success, underscoring a reciprocal relationship rather than unidirectional causality. In addition, agent-level informal leadership measured within affective task-related networks has shown mixed effects on student test scores, with affective network measures typically demonstrating stronger predictive power compared to purely informational network measures (Briley, 2016; Friedkin & Slater, 1994).

Network Structure and Cohesion: Beyond Centrality

Beyond individual centrality, broader network characteristics such as density and related structures have also been associated with academic performance. Research findings suggest that denser and more reciprocally interconnected student networks can foster collaborative learning, peer support, and shared academic resources, which are linked to enhanced achievement. Specifically, teacher collaboration networks that exhibit high connectivity and centralization have been linked to improved perceptions of collective efficacy among educators, potentially enhancing student achievement through better instructional coordination (Moolenaar et al., 2012). Furthermore, the emotional quality of teacher-student relationships within cohesive networks plays an important mediating role in student engagement and academic success (Roorda et al., 2011).

Network-level measures, especially when examined alongside human capital factors like teacher experience and demographic attributes, show significant effects on student test scores (Moolenaar et al., 2012; Ronfeldt et al., 2015). Additionally, studies have identified curvilinear effects of clique engagement and the influence of Simmelian ties (Marion et al., 2016; Tortoriello & Krackhardt, 2010). Marion et al. (2016) looked at cliques and information flow (interactions), and found a curvilinear relationship between the extent of an agent's engagement in both dynamics and productive capacity of an organization. Similarly, Krackhardt and colleagues found that Simmelian ties—triadic, reciprocal relationships representing emerging cliques—were significantly associated with academic outcomes, highlighting the importance of localized, cohesive substructures within educational networks (Krackhardt, 1998; Tortoriello & Krackhardt, 2010).

Although we found only a handful of studies of network analyses on student achievement, there are quite a few that productively analyzed other topics, including innovation climate (Moolenaar & Sleegers, 2010; Moolenaar et al., 2010; Obstfeld, 2005), productive capacity (Marion et al., 2016), actual productivity (Mehra et al., 2006; Pil & Leana, 2009), change (Kezar, 2014), adaptability (Schreiber & Carley, 2008), and learning (Schreiber & Carley, 2008), and teacher beliefs (Siciliano, 2016). Most of these examined network flow measures and found various significances for affective and task-related measures. Marion et al. (2016) looked at cliques in addition to information flow; they found curvilinear significance for the degree to which an agent is engaged in clusters. Krackhardt and colleagues similarly reported significant effects for Simmelian ties (three-way reciprocal relationships, or nascent cliques) (Krackhardt, 1998; Tortoriello & Krackhardt, 2010).

Peer Selection and Norm Reinforcement in School Networks

Peer selection processes and associated normative dynamics substantially influence academic performance. Adolescents commonly engage because of academic homophily, establishing friendships predominantly with peers of similar academic standing, thereby forming achievement-based clusters (Gremmen, 2018; Gremmen et al., 2017). These clusters simultaneously reflect and reinforce prevailing academic norms, shaping students' educational expectations, aspirations, and behaviors. Notably, the influence of peer networks on academic achievement appears asymmetric: embeddedness in low-achieving networks significantly decreases students' academic outcomes, whereas affiliation with high-achieving peers does not necessarily guarantee improvements (Wang et al., 2018). This indicates that negative peer norms exert particularly potent effects, potentially outweighing positive peer influences. Additionally, negative social positions such as rejection or friendlessness significantly impair students' academic performance by restricting access to beneficial network resources and diminishing school engagement (Gremmen, 2018). These findings suggest that efforts to support academic development may benefit from attention to both positive and negative normative dynamics within peer networks, including considerations of how such structures are formed and maintained, an area in which empirical investigations remain relatively sparse.

Boosted Trees Methodology and Application

Gradient Boosted Decision Trees (GBDT) are a powerful ensemble of learning technique that constructs predictive models through an additive, stage-wise process, where each new decision tree corrects the errors of the combined ensemble from prior iterations. Unlike bagging-based methods like random forests, GBDT employs functional gradient descent to minimize a specified loss function, allowing for high flexibility and strong predictive performance, particularly in the presence of non-linearities and complex feature interactions (Elith et al., 2008; Friedman, 2001). Notably, GBDT requires minimal data preprocessing and handles heterogeneous feature types and missing values with robustness. Recent algorithmic innovations, such as XGBoost's sparsity-aware regularization (Chen & Guestrin, 2016), CatBoost's efficient treatment of categorical variables (Dorogush et al., 2018), and LightGBM's histogram-based tree growth (Ke et al., 2017), have further enhanced its scalability and computational efficiency. These advancements have cemented GBDT as a preferred approach across structured data domains, including ecological modeling, biomedical analytics, and emerging applications in social and behavioral sciences (Subramani et al., 2023).

Compared to traditional statistical approaches such as linear regression or generalized linear models, Boosted Trees provide distinct advantages in flexibility, robustness, and predictive accuracy. Unlike parametric models, they do not rely on assumptions of linearity, normality, or homoscedasticity, allowing them to perform well in the presence of multicollinearity, non-linear relationships, and missing data (Elith et al., 2008; Ridgeway, 2024). Boosted Trees also capture complex interactions and non-additive effects among predictors without requiring pre-specification, making them particularly useful in exploratory or high-dimensional contexts. As a data-driven method, they prioritize empirical performance over model interpretability, often yielding higher predictive accuracy than generalized additive models (GAMs) or stepwise regressions (Gabidolla et al., 2025; Natekin & Knoll, 2013). Furthermore, the ability of algorithms such as XGBoost and CatBoost to incorporate regularization, control overfitting, and adapt to different loss functions enhances their reliability in noisy or heterogeneous data environments (Chen & Guestrin, 2016). These qualities make Boosted Trees an increasingly preferred modeling tool in empirical research across diverse applied fields.

In educational research, Boosted Trees have gained traction for modeling multifactorial outcomes such as student achievement, behavioral engagement, and teacher effectiveness. Their capacity to accommodate large, noisy, and heterogeneous datasets—while detecting non-linearities and complex interactions—makes them particularly useful in analyzing educational phenomena where traditional assumptions often fall short (Elith et al., 2008; Subramani et al., 2023). Recent applications demonstrate their utility in predicting academic risk, identifying underperforming groups, and modeling relational data such as peer and teacher networks (Butt et al., 2023; Deniz, 2024). For example, Deniz (2024) applied several machine learning models, including Gradient Boosting and XGBoost, to predict student achievement using demographic, psychological, and institutional variables, finding that tree-based methods outperformed linear models in predictive accuracy and were particularly effective in capturing non-linear, multifactorial influences on academic performance. While the explicit integration of Boosted Trees with classical SNA remains limited, their effectiveness in handling structured inputs derived from network metrics underscores their potential for future applications in educational network research (Subramani et al., 2023).

Epistemological Rationale

There are two epistemologically different elements of the research design we are proposing: The first involves the generation of interaction-level data, or mechanisms, and the second is the evaluation of test scores. The design mixes post-positivism (Popper, 1961) and postmodern (Lyotard, 1984) epistemologies. Positivists perceive reality as independent realities; like the science of TV transmission or Darwinian evolution, positivists envision are preexisting phenomena awaiting discovery. Postmodernists paint society as constantly changing and lacking consensus about reality, thus there can be no science. Mechanisms reflect the postmodern idea of constant change. They are defined con are dynamic processes that describe how phenomena such as change emerge (Hedström & Swedberg, 1998). Mechanisms are nominally unsuitable for statistical evaluation because of their change characteristics.

We will use mechanisms in statistical analyses but first must overcome the objections of postmodernists. This change epistemology offers two arguments against using mechanisms: First, mechanisms display constant change, thus statistical analyses are invalidated almost as soon as they are completed. Statistical logic assumes that truth is an independent reality to be discovered, thus what was true yesterday is also true today and will be true tomorrow (Crotty, 1998). Mechanisms, by contrast, change and are thus unsuitable for statistical analysis. Second, the environment in which mechanisms exist fail to show consensus about reality. In the extreme, reality is defined by each individual in a society. So, whose reality, whose perception of meaning, does one measure with statistics? There are no fixed truths to be discovered.

We propose, with Cilliers (1998) and with Boisot and McKelvey (2010), that agents in interactive networks share information with one another (Prigogine, 1997), thus both relative stability and sluggish change (rather than volatile change) are inevitable and measurable. Further, the sharing of information in interactive groups fosters common perceptions of reality, thus science can exist. Consequently, while we perceive causation as an evolutionary, changing process, we argue that evolving mechanisms are sufficiently stable to be processed statistically. During times of extreme change, this would not be true, of course.

Method

Data Description

Student data were collected from all ten elementary schools in one school district in the southeast United States. The school district provided standardized achievement test scores, student lunch status, ethnicity, gender, and school and teacher assignments for Grades 3, 4, and 5 students. Student names were coded to protect their identity. Additional teacher-level data were collected from a researcher-distributed survey. SNA data were solicited with a survey which asked, "Who would you go to for task-related advice?" (plus, other questions not used in this analysis). The teacher sample was bounded (Scott, 2000) to include only personnel whose responsibility affected the classroom function. This included teachers, paraprofessionals, and administrators. A link to the online survey was delivered electronically to participants' mailboxes one hour prior to the staff meetings at their schools, and participants completed the survey during the meeting. The principal was present but did not participate in a way that would influence responses or coerce participation. Although participation was voluntary, this strategy enabled healthy return rates.

Achievement Test Scores

Math scale scores from the ACT Aspire criterion referenced test was the outcome variable in the boosted tree analysis. Cronbach's internal consistency alphas for the 2014 sub-tests in ACT Inspire are reported in Table 1 for grades 3 - 5 (*Technical bulletin #2: Norms, scoring, scaling, and psychometrics*, 2014).

Table 1. Cronbach's Alphas by grade level*

Grade 3	0.79-.079
Grade 4	0.67-0.68
Grade 5	0.67-0.71

**Technical bulletin #2: Norms, scoring, scaling, and psychometrics*, 2014

Achievement test data is converted to best Linear Unbiased Predictors (BLUPs), as described in the section labeled, Controlling for Contextual Variation, below. BLUPs are partial scores that control for certain contextual variables that may affect the accuracy of results.

Mechanisms

Social network analysis (SNA) describes how agents process information. It identifies multiple mechanisms, defined by Hedström and Swedberg (1998) as dynamics by which outcomes (such as change) are generated. SNA yields individual and group- and individual-level coefficients that rate the degree of activity exhibited by a given mechanism. Taken together, pertinent mechanisms describe the capacity of a system to produce change events. Many SNA mechanisms measure centrality, or numbers of links connecting each agent. Centrality identifies, among other things, frequently chosen agents (degree centrality), who are highly connected (e.g., authority centrality) and agents who are connected to powerful people (Katz centrality). Other SNA measures evaluate how rapidly information passes through a system (speed), cliques, or interaction only with an agent's direct contacts (called ego networks, e.g., structural holes-efficiency). Since there are numerous mechanisms, we will only provide definitions for those identified as significant predictors by **the** boosted tree evaluations (Table 2). Definitions of SNA measures are available from Altman et al. (2017), Borgatti et al. (2013), and others.

On a related issue, boosted tree analyses will return a list of mechanisms in order of their predictive importance. This list also includes mechanisms that do not affect the outcome (SS = 0). Because our sample size is relatively small, we will remove unimportant variables and re-run the analysis with the impactful mechanisms. We found no study of procedures for determining the reliability of network data, but certain information provides clues about data reliability. Reliability is logically related to return rates; low rates omit important data about interactions. In a private communication, Carley (personal communication, June 30, 2016) recommended 90% return rates for organizational data. VisibleNetworklabs (2023) recommends 80%, calling it a gold standard. We adopt the 80% rate as our goal. The plan is to calculate mechanisms from matrix data using SNA analysis, then select pertinent mechanisms based on the reduction criteria presented above. These mechanisms will be examined using boosted tree methodology to determine their effects on student test scores. The results will be reduced to only those mechanisms exhibiting significant results.

Data Analysis

Social Network Analysis

Two important characteristics are important in this overview of SNA: First, matrices, or networks, of agents exhibit the movement of, and interactions among, ideas and information. Human entities, in a sense, do not interact; their ideas, information, preferences, personalities, and backgrounds interact and change. Absent this caveat, it is difficult to argue that networks can display creativity, knowledge, and productive capacity, just as individuals do. If creativity, knowledge, and such are solely individualistic, then interactive arguments are without merit. But it is easy to defend the macro perspective and debunk the micro perspective (individual features): One need only observe that people influence one another; they share information that they store in network-level memes and cultural truths.

Second, because human networks are dynamically interactive, they change. Contrary to positivistic logic, networks are not the same yesterday as tomorrow. Networks and the agents that comprise them change and are pressured to do so by internal interactions at least as much as by external work (the pressures of a manager, for example, (Will, 2016). Positivistic logic envisions change as products of external pressure while a postmodern perspective must see change as maneuvering internal mechanisms. SNA, using graph theory logic, acts on matrices of relationships. Matrices can represent advice relationships (from questions such as, "Who do you go to for advice on task-related issues?"), social networks ("Who you prefer to interact with is social situations?"), trust ("Who would you trust with confidential information?"), and others. In this analysis, we used data from an advice matrix.

SNA yields networks that represent interactions among agents in a respective matrix, as exemplified in In Figure 1, circles denote people, or agents. Lines denote links between people, or ties. As observed in the data section above, SNA generates numerous statistics from these representations.

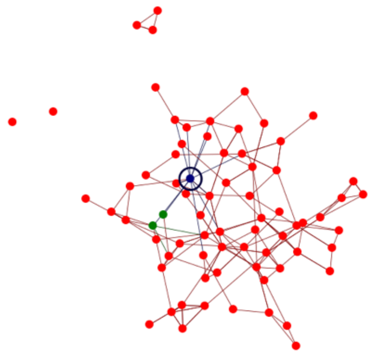

Figure 1. Sample network generated by SNA*
**SNA was performed using ORA software.*

Boosted Tree Methodology

Boosted tree, or more precisely, gradient boosted tree, is a procedure that searches through a dataset searching for a model that best predicts a desired outcome. It is used for classification (to determine, for example, whether an image is a stop sign or not), regression (finding the best set of predictors for a desired outcome), ant other related purposes (Hastie et al., 2017). Boosted tree methodology initially regresses the weakest predictor in a set of possible predictors on an outcome, then the next best predictor, and so on. Each subsequent split is regressed on the residuals from the previous split, this it learns to improve its predictability (Taboga, 2021). The process continues until a predetermined limit is reached. Boosted tree methodology is recommended on datasets greater than 50; its goal is to reduce over-fitting and N less than 50 is not typically a problem for this (Taboga, 2021). The analysis produces, among other things, two overall tests of a predictive model, one for the training data and one for validation data, and a list of mechanisms that contribute to the prediction. Boosted tree and related techniques (random trees, bootstrap trees) are known for their accuracy of prediction, robustness, the ability to deal effectively with nonlinear data, and outliers (Hastie et al., 2017).

The two overall statistics, the training data and the validation data, operate to improve the accuracy of the results. Eighty percent of a data is allocated to the training conditions and 20% to validation procedures. The results for the training data at each split is compared to the validation data and adjustments made to the training data, accordingly, producing R^2 statistics (variation accounted for). The result is not just a more an accurate prediction equation but one that is likely to produce similar results with databases it has never seen before (the validation set is likewise unseen by the training dataset until it is cross validated). The literature we have reviewed is unclear about the R^2 —R^2 or training or R^2 for validation—should be reported when judging the results. Since the goal is to create an optimal result, we have decided to use the test R^2. R^2 of 0.80 seems to be a consensus critical value for significance.

The analysis in this study will adjust for pertinent violations of assumptions regarding regression, particularly collinearity and serial autocorrelation (two or more cases have largely the same patterns of responses). Other violations are dealt with effectively by boosted tree procedures. The outcome variable for the analysis is the math BLUP scores, which were derived from the Aspire Achievement Test scale score. Readers desiring more information are encouraged to read appropriate chapters from Hastie et al. (2017), the help documents from JMP or other statistical packages are useful, or search Google for boosted tree (especially useful online reports are listed in our reference list).

Controlling for Contextual Variation

The nature of the dataset for this study will require one additional analysis after the SNA and before boosted tree: hierarchical linear modeling (HLM). The data in the study was multilevel: level 1 was student data, level 2 was teacher data, and level 3was school data. HLM was performed to partial out levels 1 and 3 variations from the level 2 data and to control problematic contractual variables. Before performing the HLM, data were mean centered across the three grade levels to control for maturity. The HLM model included a random term for school (level 3) to control for differences across school. At level 2, differences by teacher, teacher gender, and teacher race were controlled. We also added two interaction terms for level 2: teacher by student race and teacher by student gender. Finally, fixed terms at the student level (level 1) controlled for student race and gender.

The specific form of the model was:

$$y_{ijklm} = \mu + S_i + T(S)_{ij} + G_k + R_l + G*T(S)_{ijk} + R*T(S)_{ijl} + e_{ijklm}$$

where y_{ijklm} is the end-of-year centered test score for student m of gender k and ethnicity l in school i with teacher j; μ is the overall mean of the end-of-year test scores; S_i is the random effect of school i; $T(S)_{ij}$ is the random effect of teacher j within school i; G_k is the effect of gender k; R_l is the effect of ethnicity l; $G*T(S)_{ijk}$ is the random interaction effect of gender k and teacher j within school i; $R*T(S)_{ijl}$ is the random interaction effect of ethnicity l and teacher j within school i; and e_{ijklm} is random error. The data were analyzed using JMP (v 17) statistical software. JMP, using restricted Maximum Likelihood procedures, generated “best linear unbiased predictors” (BLUPs) or partial test scores by teacher after controlling for the above fixed, interaction, and random effects. The BLUP of teacher j in school i was defined as the teacher average ($\overline{y}_{ij...}$) minus the overall average ($\overline{y}_{.....}$), multiplied by the ratio of the teacher variation to the total variation.

$$BLUP_{ij} = (\overline{y}_{ij...} - \overline{y}_{.....}) * \frac{\sigma^2_{T(S)}}{\sigma^2_{Total}}$$

BLUP is sometimes referred to as “shrinkage” estimates since the smaller the ratio of the teacher variation to the total variation, the closer the BLUP estimate of teacher effect is to the overall average. BLUP is used as the dependent variable for subsequent analyses.

Results

Teacher data for the network analysis and test scores were summarized by teachers in the boosted tree analysis. There were 118 math teachers in the 10 schools; 111 were white and 107 were female (Table 2). No violations of regression assumptions were identified.

Table 2. Demographics for professional personnel

Characteristic	Category	Number
Race	Black	7
	White	111
Gender	Female	107
	Male	9
Years Teaching	Mean	9,6

N=118, Missing = 0

Two thousand six hundred eighty-two students were enrolled in the ten schools, and all but seven took the Aspire end-of-year test. Most students were African American, followed by Whites, Hispanics, Mixed, and then others. There were 375 males and 207 females; 1,412 received free meals, 134 were reduced, and 963 meals were full pay (Table 3).

Table 3. Demographics for students

Characteristic	Category	Count	Probability
Ethnicity	Black	844	0.31469
	White	1426	0.53169
	Hispanic	202	0.07532
	Mixed	160	0.05966
	Asian	46	0.01715
	Other	4	0.00037
	Total	**2682**	**1.00000**
Gender	Male	1375	-
	Female	1307	-
Meal Status	Free	1412	-
	Reduced	963	-
	Paid	134	-

N = 2682, Missing = 7 (0.26%)

HLM Results

Results of the hierarchical linear model analysis with math achievement test scores as outcomes are found in Table 5. The analysis revealed that the significant contextual variables were free-reduced lunch status, ethnicity, and school attended, but that student gender had no significant impact. The total R^2 was 0.34, which is impressive.

Table 4. Results from HLM with math achievement tet scores as outcome

	Variable	F
Context	School	6.37**
	Student Gender	1.38
	Student Lunch Status	68.70**
	Student Ethnicity	28.76**
	R^2	0.34
	R^2Adjusted	0.34

Boosted Trees Analysis

Results of the Boosted Tree analysis (outcome was math achievement) are reported in Table 6. Data were distributed into two sets: a training set and a validation set (see Overall Statistics in Table 6). Eighty percent of the data were randomly selected for attaining and the rest for validation. The validation set was compared against the training set to ascertain accuracy. The R^2 for training was 0.20, which is weak; boosted tree analysis expects an R^2 of at least 0.80. The overall model would not successfully discriminate between teachers with successful math achievement scores (BLUPs) and those without.

The Column Contributions in Table 6 exhibit mechanisms that contributed to the model's effect. The greatest contributions were from clustering coefficient (34%; see the portion column) and authority centrality (24%). Clustering coefficient identifies the density of an agent's ego network, the people to whom agents are directly tied. High clustering indicates a tightly interacting group. Authority centrality indicates the degree to which an agent is tied to people who are "in-the-know". Katz centrality (17%), hub centrality (16%), and cognitive demand (8%) have smaller R^2 coefficients, but are nonetheless influential. Katz centrality is a measure of the diversity of people with advice and of power; hub centrality measures the degree to which agents send information to highly connected individuals; and cognitive demand identifies the degree of information each gent processes in performing its tasks. Notice that three of these five mechanisms, authority, Katz, and hub centralities, all deal with interactions and the sharing of knowledge. Essentially, they are related to social capital. Daly et al. (2014) also concluded that social capital was an important mechanism in producing test scores.

Table 5. Boosted tree for math BLUP mean center

Parameter	Value
Target	Math BLUP Mean Center
Validation Column	Validation
Number of Layers	21
Splits per Tree	2
Learning Rate	0.063
Number of training rows	87
Number of validation rows	29

Table 6. Overall statistics

Dataset	RSquare	RASE	N
Training	0.200	2.561841	87
Validation	-0.03	4.2853724	29

Table 7. Column contributions

Clustering Coefficient	11	377.495145		0.3229
Authority Centrality	9	281.455666		0.2408
X Katz Centrality	8	199.158348		0.1704
Hub Centrality	8	188.683717		0.1614
Cognitive Demand	5	96.6967593		0.0827

For comparison, we ran a stepwise regression and found that none of the predictors were significant. We compared our results with Jiang's (2017) analysis of comparable data. Jiang examined additional achievement scores for English language, Science, and Social Sciences using different analytical approaches. For Math achievement in Jiang's analysis, the mechanisms, Potential Boundary Spanner in the social network ($\beta = 0.20$) and Structural Holes Effective Network Size for the trust network (β.18) where statistically significant. Potential boundary spanner refers to agents who may ethe potential to serve as a bridge between two (or more) otherwise unconnected groups of people. Structural holes effective network size identifies each person's (called egos) direct contacts (called alters) and calculates the average number of untapped alters; that is, alters who could be unique resources of the ego. One would assume that structural holes, which measure each agent's personal network, would have a strong effect. It was relatively large in Jiang's data but was not influential for the math scores in our study.

Table 8. Definitions of variables in the analysis

Authority Centrality	Degree to which agents are connected to people who connect with many other agents. Participants high in this measure are "in-the-loop" with agents who are well-connected.
Hub Centrality	Degree each agent sends information to highly connected agents.
Cognitive Demand	Effort expended by each agent to do its tasks; agents high on this measure are considered emergent leaders.
Katz Centrality	Indicates who is connected to powerful agents. That is, it measures the relative influence of agents.
Clustering Coefficient	Average density of each agent's ego network (immediate links). Higher clustering supports local information diffusion and a decentralized infrastructure because agents are likely to share information and know what is happening in their work group.
Potential Boundary Spanner	Identifies agents that connect groups that are otherwise disconnected.
Structural Holes Effective Network Size	Degree to which agents are connected to highly connected individuals in their ego network

Discussion and Conclusion

In summary, our analysis of the effects of dynamic, changing mechanisms on student achievement did not yield results that rose to the level of a predictive model in machine learning. It did, however, identify mechanisms that may well have been significant in a more traditional regression analysis of theories and hypotheses. All these five mechanisms are potentially influential in this study, and they came from the advice network. We did run a follow-up boosted tree analysis that included he trust and social networks, but the final R_2 was with only a few hundreds of the advice-only analysis (0.20).

What can explain the failure to find a predictive machine learning model? First is the simplest explanation: Teacher's interaction does not affect student achievement to a significant degree. Positivistic analyses since the 1960s, when contextual variables such as SES were first added to prediction equations, have had difficulty finding effects for schooling on student achievement (Coleman, 1968; Coleman et al., 1966). Leithwood and Jantzi (2006) ran a set of regression analyses and found several school-related variables that affected student achievement, but, in their key analysis, the researchers entered the school variable first, before the contextual variables, so essentially his successful findings were posted before contextual variables were added. Flanigan et al. (1996) using path analysis found a modest effect for administrative funding on student achievement, but path analysis procedures were rather new at that time and their analysis did not include controls for error, among other things. Jabbar et al. (2022) performed a meta-analysis of charter schools on achievement and found only weak support for a relationship.

On the obverse side, a meta-analysis by (Lei et al., 2022) concluded that game-based teaching significantly improved science learning. Jiang (2017) examined English, science and social studies scores in addition to data from social and trust networks, finding some moderately strong mechanisms that affect achievement. Daly et al. (2014) examined in-degree, out-degree, and total centrality (SNA mechanisms) on student, English language achievement scores (summarized at the teacher level) and likewise found evidence that interactive data influences test scores. They used HLM methodologies, a positivistic methodology, for their analysis. This is a brief review of pertinent literature regarding achievement scores, but the point is, particularly given Jiang's and Daly's findings and our findings, all of whom calculated SNA data, that further research with interactive mechanisms may provide new insights into student achievement. Second, future research should address a possible weakness in our study: the number of respondents.118. This is complicated by the SNA requirement that one obtain high return rates from each bounded network; large numbers alone are not enough.

Implications for Practitioners and Theory

Readers are referred to the literature on complexity leadership theory (Lichtenstein et al., 2006; Schreiber et al., 2006; Uhl-Bien & Marion, 2009; Uhl-Bien et al., 2007) regarding the implications of this research. Complexity theory is the study of the effects of networked interaction on organizations; SNA measures interaction among people and is thus a useful research tool for is the study of how leaders can influence complex dynamics. Complexity leadership theorists propose that leaders maneuver SNA mechanisms that generate or suppress complex dynamics.

This paper represents a deliberate attempt to move organizational theory and research beyond its postpositivist roots toward a greater focus on macro dynamics—the interactions of people in bounded networks—rather than micro perspectives—the individual is an independent unit of analysis. It further encourages researchers to learn and adopt new analytical strategies—SNA and machine learning in this case. Qualitative procedures are widely used to address some of the issues we raise, but there are more "fields" to explore.

Limitations

We cannot claim that SNA-based research is always replicatable, because network mechanisms are sensitive to differing and changing conditions in an organization. The mechanisms that influence one organization likely differ from those that influence another. There are two responses applicable to this observation: First, As Cilliers (1998) argues, network agents interact, share information, and adapts to each other's worldviews, thus there could very well show a degree of constancy across various organizations. The second point is more important: researchers applying macro (e.g., SNA) methodologies should not typically hypothesize whether a given mechanism or mechanisms influence given outcomes; rather, they should ask whether outcomes are influenced by interactive processes generally, or by categories of mechanisms (those that measure social capital, for example). The research we propose is a hypothesis test in a broad, macro sense (e.g., interactive mechanisms influence an outcome). The findings can help explain what is going on in the organization.

Declarations

Acknowledgments: The authors wish to recognize the contribution made by William Bridges, professor of Math Science at Clemson University, for his significant help in developing the section of this paper on BLUPs.

Authors' contributions: All authors contributed to this work and reviewed and approved the final version for publication.

Competing interests: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethics approval and consent to participate: The necessary ethical approvals have been obtained for the study, and ethical guidelines have been strictly followed throughout.

Publisher's note: Culture, Education, and Future remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Orcid ID

Russ Marion https://orcid.org/0000-0001-9366-0291

Xiaorang Ma https://orcid.org/0000-0002-7551-2989

References

Alspaugh, J. W. (1996). The effects of geographic and management factors on the cost of pupil transportation. *Journal of Education Finance*, *22*(2), 180-194.

Altman, N., Carley, K., & Reminga, J. (2017). *ORA user guide 2017* (CMU-ISR-17-100). S. o. C. S. Institute for Software Research.

Bailey, D. E., Faraj, S., Hinds, P. J., Leonardi, P. M., & von Krogh, G. (2022). We are all theorists of technology now: A relational perspective on emerging technology and organizing. *Organization Science, 33*(1), 1–18.

Bartels, J. M., & Magun-Jackson, S. (2021). Achievement goals as predictors of motivation and academic persistence. *Social Psychology of Education, 24*(2), 491–507.

Bastos, J. A. (2022). Predicting credit scores with boosted decision trees. *Forecasting, 4*(4), 925–935.

Bedeian, A. G., & Hunt, J. G. (2006). Academic amnesia and vestigial assumptions of our forefathers. *The Leadership Quarterly, 17*(2), 190-205.

Boisot, M., & McKelvey, B. (2010). Integrating modernist and postmodernist perspectives on organizations: A complexity science bridge. *Academy of Management Review, 35*(3), 415–433.

Bond, R. M., Chykina, V., & Jones, J. J. (2017). Social network effects on academic achievement. *The Social Science Journal, 54*(4), 438–449.

Borgatti, S. P., Everett, M. G., & Johnson, J. (2013). *Analyzing social networks.* Sage.

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science, 323*(5916), 892–895.

Bossert, S. T., Dwyer, D. C., Rowan, B., & Lee, G. V. (1982). The instructional management role of the principal. *Educational Administration Quarterly, 18*(3), 34–64.

Bouchet, F., & Kizilcec, R. F. (2020). Self-concept as a predictor of student engagement in online courses. *Computers & Education, 152,* 103894.

Briley, B. C. (2016). *Exploring the effects of network dynamics on student test scores in a rural middle school* (Doctoral dissertation, Clemson University). Retrieved from https://open.clemson.edu/cgi/viewcontent.cgi?article=2748&context=all_dissertations

Bruun, J., & Brewe, E. (2013). Talking and learning physics: Predicting future grades from network measures and Force Concept Inventory pretest scores. *Physical Review Special Topics—Physics Education Research, 9*(2), 020109.

Butt, N. A., Mahmood, Z., Shakeel, K., Alfarhood, S., Safran, M., & Ashraf, I. (2023). Performance prediction of students in higher education using multi model ensemble approach. *IEEE Access, 11,* 136091–136108.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794).

Cilliers, P. (1998). *Complexity and postmodernism: Understanding complex systems.* Routledge.

Coleman, J. S. (1968). Equality of educational opportunity. *Equity and Excellence in Education, 6*(5), 19–28.

Coleman, J. S., Campbell, E., Hopson, C., McPartland, J., Mood, A., & Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity.* U.S. Government Printing Office.

Comprehensive introduction for beginners. *Visible Network Labs.* https://visiblenetworklabs.com/guides/social-network-analysis-101/#elementor-toc__heading-anchor-1

Crotty, M. (1998). *The foundations of social research.* Sage.

Cullen-Lester, K. L., & Yammarino, F. J. (2016). Collective and network approaches to leadership: Special issue introduction. *The Leadership Quarterly*, *27*(2), 173–180.

Daly, A. J., Moolenaar, N. M., Der-Martirosian, C., & Liou, Y. (2014). Accessing capital resources: Investigating the effects of teacher human and social capital on student achievement. *Teachers College Record*, *116*(7), 1–42.

Deniz, E. (2024). Evaluating the effectiveness of machine learning models in predicting student academic achievement. *ADBA Computer Science, 1*(1), 8–13.

Dess, G. G., & Shaw, J. D. (2001). Voluntary turnover, social capital, and organizational performance. *Academy of Management Review*, *26*(3), 446-456.

Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: Gradient boosting with categorical features support. *arXiv preprint arXiv:1810.11363.*

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109-132.

Edmonds, R. R. (1979). Effective schools for the urban poor. *Educational Leadership*, *37*(1), 15–24.

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*(4), 802–813.

Esposito, A. G., & Bauer, P. J. (2022). Determinants of elementary-school academic achievement: Component cognitive abilities and memory integration. *Child Development*, *93*(6), 1777-1792.

Flanigan, J. L., Marion, R. A., & Richardson, M. D. (1996). Causal and temporal analyses of increased funding on student achievement. *Journal of Research and Development in Education*, *30*(4), 222–247.

Frank, K. A., Zhao, Y., & Borman, K. (2004). Social capital and the diffusion of innovations within organizations: The case of computer technology in schools. *Sociology of Education*, *77*(2), 148-171.

Friedkin, N. E., & Slater, M. R. (1994). School leadership and performance: A social network approach. *Sociology of Education*, *67*(2), 139-157.

Friedman, M. (2001). *Dynamics of reason*. Csli Publications.

Gabidolla, M., Zharmagambetov, A., & Carreira-Perpiñán, M. Á. (2025). Boosted sparse oblique decision trees. Retrieved from http://faculty.ucmerced.edu/mcarreira-perpinan/papers/baylearn20b.pdf

Gremmen, M. C. (2018). Social network processes and academic functioning: The role of peers in students' school well-being, academic engagement, and academic achievement.

Gremmen, M. C., Dijkstra, J. K., Steglich, C., & Veenstra, R. (2017). First selection, then influence: Developmental differences in friendship dynamics regarding academic achievement. *Developmental Psychology*, *53*(7), 1356.

Hallinger, P. (2007). *Research on the practice of instructional and transformational leadership: Retrospect and prospect.* Retrieved from https://research.acer.edu.au/research_conference_2007/7

Hallinger, P., & Heck, R. H. (1998). Exploring the principal's contribution to school effectiveness: 1980–1995. *School Effectiveness and School Improvement*, *9*, 157–191.

Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

Hedström, P., & Swedberg, R. (1998). *Social mechanisms: An analytical approach to social theory.* Cambridge University Press.

Hendrickx, M. (2012). The role of the teacher in classroom peer relations [Doctoral dissertation].

Hughes, J. N., Luo, W., Kwok, O. M., & Loyd, L. K. (2008). Teacher-student support, effortful engagement, and achievement: A 3-year longitudinal study. *Journal of Educational Psychology, 100*(1), 549–569.

Hunt, J. G., & Dodge, G. E. (2000). Leadership déjà vu all over again. *The Leadership Quarterly*, *11*(4), 435-458.

Jabbar, H., Fong, C. J., Germain, E., Li, D., Sanchez, J., Sun, W.-L., & Devall, M. (2022). The competitive effects of school choice on student achievement: A systematic review. *Educational Policy, 36*(2), 247–281.

Jackson, J. C., Rand, D., Lewis, K., Norton, M. I., & Gray, K. (2017). Agent-based modeling: A guide for social psychologists. *Social Psychological and Personality Science, 8*(4), 387–395.

Jencks, C. (1972). *Inequality: A reassessment of the effect of family and schooling in America.* Basic Books.

Jiang, X. (2017). Understanding teacher effectiveness with complexity and network theories [Dissertation, Clemson University]. Clemson, SC.

Karaman, M. A., & Watson, J. C. (2020). Predictive role of growth mindset, coping, and self-efficacy on academic persistence. *Journal of College Student Retention: Research, Theory & Practice, 22*(4), 746–763.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems, 30*, 1–9.

Kezar, A. (2014). Higher education change and social networks: A review of research. *The Journal of Higher Education*, *85*(1), 91-125.

Klapp, T., Klapp, A., & Gustafsson, J. E. (2024). Relations between students' well-being and academic achievement: Evidence from Swedish compulsory school. *European Journal of Psychology of Education, 39*(1), 275–296.

Krackhardt, D. (1998). Simmelian ties: Super strong and sticky. In R. M. Kramer & M. A. Neale (Eds.), *Power and influence in organizations* (pp. 21–38). Sage.

Lei, H., Chiu, M. M., Wang, D., Wang, C., & Xie, T. (2022). Effects of game-based learning on students' achievement in science: A meta-analysis. *Journal of Educational Computing Research, 60*(6), 1373–1398.

Leithwood, K., & Jantzi, D. (2005). A review of transformational school leadership research 1996–2005. *Leadership and Policy in Schools, 4*(3), 177–199.

Leithwood, K., & Jantzi, D. (2006). Transformational school leadership for large-scale reform: Effects on students, teachers, and their classroom practices. *School Effectiveness and School Improvement, 17*(2), 201–227.

Li, L. (2013). The complexity of language teachers' beliefs and practice: One EFL teacher's theories. *The Language Learning Journal*, *41*(2), 175-191.

Lichtenstein, B. B., Uhl-Bien, M., Marion, R., Seers, A., Orton, J. D., & Schreiber, C. (2006). *Complexity leadership theory: An interactive perspective on leading in complex adaptive systems. Emergence: Complexity and Organization*, *8*(4), 2–12.

Liu, Y., Bellibaş, M. S., & Gümüş, S. (2020). The effect of instructional leadership and distributed leadership on teacher self-efficacy and job satisfaction: Mediating roles of supportive school culture and teacher collaboration. *Educational Management Administration & Leadership*, *49*(3), 430–453.

Louis, K. S., & Marks, H. M. (1998). Does professional community affect the classroom? Teachers' work and student experiences in restructuring schools. *American Journal of Education*, *106*(4), 532-575.

López-Sánchez, M., Arango Paternina, C. M., Petro Petro, J., Lema Gómez, L., Eusse López, C., Petro, J. L., & Perea Velásquez, F. (2023). Academic performance and social networks of adolescents in a Caribbean city in Colombia. *BMC Psychology*, *11*(1), 255.

Louis, K. S., Dretzke, B., & Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness and School İmprovement*, *21*(3), 315-336.

Lyotard, J.-F. (1984). *The postmodern condition: A report on knowledge.* Manchester University Press.

Ma, X., & Marion, R. (2024). How does leadership affect teacher collaboration? Evidence from teachers in US schools. *School Effectiveness and School Improvement*, *35*(2), 116–141.

Ma, X., & Marion, R. (2025). Linking distributed leadership to teacher job satisfaction in China: Mediating roles of teacher well-being and work motivation. *Educational Management Administration & Leadership.* Advance online publication. https://doi.org/10.1177/17411432241303299

Marion, R., Christiansen, J., Klar, H. W., Schreiber, C., & Erdener, M. (2016). Informal leadership, interaction, cliques and productive capacity in organizations: A collectivist analysis. *The Leadership Quarterly*, *27*(2), 242–260.

Marion, R., & Gonzales, L. D. (2013). *Leadership in education: Organizational theory for the practitioner*. Waveland Press.

Marks, H. M., & Printy, S. M. (2003). Principal leadership and school performance: An integration of transformational and instructional leadership. *Educational Administration Quarterly*, *39*(3), 370-397.

Mehra, A., Dixon, A. L., Brass, D. J., & Robertson, B. (2006). The social network ties of group leaders: Implications for group performance and leader reputation. *Organization Science*, *17*(1), 64-79.

Moolenaar, N. M., & Sleegers, P. J. C. (2010). Social networks, trust, and innovation: The role of relationships in supporting an innovative climate in Dutch schools. In A. J. Daly (Ed.), *Social network theory and educational change* (pp. 97–114). Harvard University Press.

Moolenaar, N. M., Sleegers, P. J. C., & Daly, A. J. (2012). Teaming up: Linking collaboration networks, collective efficacy, and student achievement. *Teaching and Teacher Education*, *28*(2), 251–262.

Moolenaar, N. M., Daly, A. J., & Sleegers, P. J. (2010). Occupying the principal position: Examining relationships between transformational leadership, social network position, and schools' innovative climate. *Educational Administration Quarterly*, *46*(5), 623-670.

Morinaj, J., de Moll, F., Hascher, T., Hadjar, A., Grecu, A., & Scharf, J. (2023). School alienation among adolescents in Switzerland and Luxembourg: The role of parent and peer supportive attitudes toward school and teacher autonomy support. *Youth & Society*, *55*(2), 187–212.

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics, 7.*

Nunes, C., Oliveira, T., Castelli, M., & Cruz Jesus, F. (2023). Determinants of academic achievement: How parents and teachers influence high school students' performance. *Heliyon, 9*(2).

Obstfeld, D. (2005). Social networks, the tertius iungens orientation, and involvement in innovation. *Administrative Science Quarterly*, *50*(1), 100-130.

Olivier, E., Archambault, I., De Clercq, M., & Galand, B. (2019). Student self-efficacy, classroom engagement, and academic achievement: Comparing three theoretical frameworks. *Journal of Youth and Adolescence, 48*, 326–340.

Pearce, C. L., Conger, J. A., & Locke, E. A. (2008). Shared leadership theory. *The Leadership Quarterly*, *19*(5), 622-628.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315–341.

Pil, F. K., & Leana, C. (2009). Applying organizational research to public school reform: The effects of teacher human and social capital on student performance. *Academy of Management Journal*, *52*(6), 1101-1124.

Ponomareva, N., Colthurst, T., Hendry, G., Haykal, S., & Radpour, S. (2017, December). Compact multi-class boosted trees. In *2017 IEEE International Conference on Big Data (Big Data)* (pp. 47–56). IEEE.

Popper, K. R. (1961). *The logic of scientific discovery.* Basic Books.

Popper, K. R., Freed, J., & Freed, L. (1961). The logic of scientific discovery. *Philosophy of Science, 28*(3), 319–324.

Prigogine, I. (1997). *The end of certainty.* The Free Press.

Ridgeway, G. (2024). Generalized boosted models: A guide to the gbm package. *Update, 1*(1), 1–15.

Ronfeldt, M., Farmer, S. O., McQueen, K., & Grissom, J. A. (2015). Teacher collaboration in instructional teams and student achievement. *American Educational Research Journal, 52*(3), 475–514.

Roorda, D. L., Koomen, H. M., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research, 81*(4), 493–529.

Sammons, P., Hillman, J., & Mortimore, P. (1995). *Key characteristics of effective schools: A review of school effectiveness research.* OFSTED.

Schreiber, C., & Carley, K. M. (2008). Network leadership. In *Complexity leadership* (pp. 291–331). Information Age Charlotte.

Schreiber, C., Marion, R., Uhl-Bien, M., & Carley, K. (2006). Multi-agent based simulation of a model of complexity leadership. Paper presented at the *International Conference on Complex Systems.* Boston, MA.

Scott, J. (2000). *Social network analysis* (2nd ed.). Sage Publications.

Siciliano, M. D. (2016). It's the quality not the quantity of ties that matters: Social networks and self-efficacy beliefs. *American Educational Research Journal*, *53*(2), 227-262.

Somech, A. (2010). Participative decision making in schools: A mediating-moderating analytical framework for understanding school and teacher outcomes. *Educational Administration Quarterly*, *46*(2), 174-209.

Somech, A., & Wenderow, M. (2006). The impact of participative and directive leadership on teachers' performance: The intervening effects of job structuring, decision domain, and leader-member exchange. *Educational Administration Quarterly*, *42*(5), 746-772.

Spillane, J. P. (2005). Distributed leadership. *The Educational Forum, 69*(2), 143–150.

Subramani, N., Veerappampalayam Easwaramoorthy, S., Mohan, P., Subramanian, M., & Sambath, V. (2023). *A gradient boosted decision tree-based influencer prediction in social network analysis*. *Big Data and Cognitive Computing, 7*(6), 1-18.

Supovitz, J., Sirinides, P., & May, H. (2010). How principals and peers influence teaching and learning. *Educational Administration Quarterly, 46*(1), 31–56.

Su-Russell, C., & Russell, L. T. (2021). Maternal autonomy support and children's social competencies, academic skills, and persistence: Social determinants and mediation. *Journal of Child and Family Studies*, *30*, 757-770.

Taboga, M. (2021). Boosted tree. In *Lectures on machine learning.* Retrieved from https://www.statlect.com/machine-learning/boosted-tree

Tan, E. (2014). Human capital theory: A holistic criticism. *Review of Educational Research*, *84*(3), 411-445.

Technical bulletin #2: Norms, scoring, scaling, and psychometrics. (2014). Retrieved from https://www.discoveractaspire.org/pdf/2014_ACT-AspireTechnicalBulletin2.pdf

Teddlie, C., & Stringfield, S. (1993). *Schools make a difference: Lessons learned from a 10-year study of school effects*. Teachers College Press.

Thapa, A., Cohen, J., Guffey, S., & Higgins D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research, 83*(3), 357–385.

Tortoriello, M., & Krackhardt, D. (2010). Activating cross-boundary knowledge: The role of Simmelian ties in the generation of innovations. *Academy of Management Journal*, *53*(1), 167-181.

Uhl-Bien, M., & Marion, R. (2009). Complexity leadership in bureaucratic forms of organizing: A meso model. *The Leadership Quarterly, 20*(4), 631–650.

Uhl-Bien, M., Marion, R., & McKelvey, W. (2007). Complexity leadership theory: Shifting leadership from the industrial age to the knowledge era. *The Leadership Quarterly, 18*(4), 298–318.

Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education*, *24*(1), 80-91.

Virgilio, I., Teddlie, C., & Oescher, J. (1991). Variance and context differences in teaching at differentially effective schools. *School Effectiveness and School Improvement, 2*, 152–168.

VisibleNetworklabs. (2023). Social network analysis 101: Ultimate guide. Retrieved from https://visiblenetworklabs.com/guides/social-network-analysis-101/#elementor-toc__heading-anchor-1 von Krogh, G., Roberson, Q., & Gruber, M. (2023). Recognizing and utilizing novel research opportunities with artificial intelligence. *Academy of Management Journal*, *66*(2), 367–373.

Wang, M. T., & Eccles, J. S. (2012). Social support matters: Longitudinal effects of social support on three dimensions of school engagement from middle to high school. *Child Development*, *83*(3), 877–895.

Wang, M. T., Kiuru, N., Degol, J. L., & Salmela-Aro, K. (2018). Friends, academic achievement, and school engagement during adolescence: A social network approach to peer influence and selection effects. *Learning and Instruction*, *58*, 148–160.

Wentzel, K. R., Russell, S., & Baker, S. (2016). Emotional support and expectations from parents, teachers, and peers predict adolescent competence at school. *Journal of Educational Psychology*, *108*(2), 242–250.

Will, T. E. (2016). Flock leadership: Understanding and influencing emergent collective behavior. *The Leadership Quarterly*, *27*(2), 261–279.

Williams, E. A., Zwolak, J. P., Dou, R., & Brewe, E. (2019). Linking engagement and performance: The social network analysis perspective. *Physical Review Physics Education Research*, *15*(2), 020150.

Yeung, S. S., King, R. B., Nalipay, M. J. N., & Cai, Y. (2022). Exploring the interplay between socioeconomic status and reading achievement: An expectancy-value perspective. *British Journal of Educational Psychology*, *92*(3), 1196–1214.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64–70.