

Chapter 1

Anticipating the Side Effects of Educational Reform Using System Dynamics Modeling

DAVID R. KEITH 

Melbourne Business School, University of Melbourne

AISHWARYA YADAMA

Sloan School of Management, Massachusetts Institute of Technology

ELLEN O'NEILL

SKIP Designed

SARAS CHUNG 

SKIP Designed

There have been countless efforts to improve academic achievement in public schools across the United States, especially in urban school districts. Few efforts, however, use rigorous analytic modeling tools to anticipate and prevent the potential side effects of change. This chapter proposes the application of system dynamics on educational reform efforts by offering a conceptual model, literature review, and research agenda to identify and mitigate the potential side effects of reform. The model is based on interviews and workshops with public school leaders, teachers, and parents in a mid-sized midwestern urban city. We identify a “capability trap” that may lead to policy resistance with educational reforms. Limitations to this approach are noted, and future areas of research are identified.

Public education has long been the target of interventions to improve outcomes for students, families, and educators. Interventions are multitudinous and varied and span topics from reducing educational achievement gaps between the United States and other countries (McDonnell, 2008; Rury et al., 2022), to banning instructional pedagogies such as critical race theory (Morgan, 2022), to mandating high-stakes testing (Elmore, 2004). Additionally, they are multilayered at local, state, and federal levels, adding external complexity to an educational system that is frequently under pressure to improve (Elmore, 2004).

Review of Research in Education

March 2024, Vol. 48, pp. 1–27

DOI: 10.3102/0091732X241260012



Chapter reuse guidelines: sagepub.com/journals-permissions

© 2024 The Author(s). journals.sagepub.com/home/rre

Policy interventions seek to improve the state of the system, but much like medical interventions, they can result in unforeseen consequences that can be negative or positive. In medicine, these “side effects” are widely accepted as consequences that can be anticipated because of the intervention and are required to be conveyed on pharmaceutical warnings and consent forms (Zhao, 2017). Meyer (1977) once described education as an institution of effects, affecting society through an allocation of socializing while conferring success on some and failure on others. Through allocation theory, he suggested the institutional effects of education as an intervention has both benefits and negative outcomes on some groups in society (Meyer, 1977). The anticipation and acceptance of unanticipated harmful effects of interventions in education is not common practice, leaving room for growth in a field that is increasingly expected to adopt rigorous practice and research (Bryk et al., 2015; Zhao, 2017).

In educational interventions, side effects are commonly depicted post hoc (after intervention) or with the goal of supporting or opposing an initiative. This is true in peer-reviewed processes, too: Advocates seeking to prevent the implementation of a policy may publish about potential side effects from a values-based assumption or expert conjecturing. Additionally, one may see postmortem analyses on educational interventions after implementation (e.g., Cawelti, 2006), but rarely are these side effects predicted using systematic and rigorous research methods.

There are, however, a growing number of studies that outline the side effects of various interventions in education. Yong Zhao and colleagues, among others, answered the recent call for side effects research in education in various published articles. By completing a deep review of literature on large-scale assessments in education, Emler et al. (2019) synthesized side effects from a postintervention perspective. When measuring the effect of school choice and vouchers on students, Zhao (2019) examined the effect beyond the average student to subsets of students, finding differential effects—positive for some and negative for others. In this study, he noted that “one student’s medicine may very well be another one’s poison” (Zhao, 2019, p. 66). These examples provide methods that examine post hoc intervention effects, yet few methods can be used to scientifically predict the effects of interventions before implementation. Most anticipatory studies of side effects from policy interventions are therefore based on values-based propositions and hypothetical assumptions.

No Child Left Behind (NCLB) provides numerous examples where values-based conjecturing and hypothetical assumptions are made without rigorous modeling or testing. The wide-sweeping bipartisan policy, which provides national standards for learning and test-based accountability, was highly contested both before, during, and after implementation. There were numerous voices warning of the potential negative effects of NCLB’s standards-based reforms; however, few rigorous analyses were conducted to demonstrate potential side effects and/or evidence of impending harm. For instance, in the *Harvard Educational Review*, Borkowski and Sneed (2006) noted the controversies surrounding NCLB by stating it was based on ideology and untested

federally mandated remedies. However, their own article did not propose scientifically based experiments, models, or research-based indicators of potential negative (or positive) outcomes. Rather, they provided legal perspectives on why they believed it would not work.

Others also noted their fear of NCLB's effect, but without a systematic method of formal analysis of the possible unintended consequences. Goldrick-Rab and Mazzeo (2005) wrote an article anticipating the impact of NCLB on college access based on a summary of research that outlined important factors contributing to student participation and success in college. They then compared components of NCLB to the factors that contribute to access and then studied other research examining the effects of NCLB for academic pressure and high-stakes testing to assess potential harm (Goldrick-Rab & Mazzeo, 2005). The method of examining the potential impact of a policy through rigorous review may be inadequate to predict outcomes because it relies on existing research studies and postulation of how the policy will affect parts of the system that may not be obvious.

Despite the lack of rigorous analysis on the potential side effects of NCLB, numerous impacts have been identified with the benefit of hindsight. Regarding intended outcomes, statewide data from the National Assessment of Educational Progress analyzed by Dee and Jacob (2011) demonstrated large and statistically significant increases in the math achievement of four graders (effect size = .23 by 2007), concentrated among White and Hispanic students and students eligible for subsidized lunches. But NCLB was overshadowed by the unintended consequences that resulted. For example, unanticipated effects such as teacher morale and neighborhood housing values were effects found after NCLB was implemented. One study noted that NCLB elevated teacher stress and decreased morale (Smith & Kovacs, 2011). Another study showed that schools receiving the designation of "failing," a key part of the policy, led to a decrease in home prices in nearby areas as families moved in search of better educational opportunities for their children (Bogin & Nguyen-Hoang, 2014). Linda Darling-Hammond (2007) noted that the unintended negative consequences from NCLB included a narrowed curriculum, focus on low-level skills generally reflected on high-stakes tests, inappropriate assessment of students with special needs and English language learners, and strong incentives to exclude students with low test scores from schools to achieve test score targets.

Although it is easier to see these side effects after implementation or to attribute them to bad luck, considering them somehow less valid than the intended consequence of the intervention, these side effects are no accident. Rather, they are the direct and endogenous consequences of feedbacks that exist in the structure of the system—feedbacks that can and should be anticipated *ex ante*. Anticipating these effects requires an understanding of causality and feedback effects over time. This practice is encapsulated in the practice of "closed loop thinking" or "feedback thinking," more popularly referred to as "systems thinking." Closed loop thinking happens when people see the world as a set of ongoing, interdependent processes rather than one-way relations between a group of factors and the phenomenon that these factors are causing (Richmond, 1993).

Exercising this type of thinking forces people to look at the circular cause-effect relations or loop of a system as responsible for generating behavior patterns exhibited by that system rather than external forces (Richmond, 1993).

Frequently, closed loop thinking that captures these feedbacks within a single integrative analysis does not occur when education policies are being developed, resulting in prescriptions that often underperform or even amplify existing problems. For example, zero-tolerance policies, an umbrella term for codes of conduct in schools that set ground rules for student behavior with mandatory consequences for committing an offense, were implemented widely in the late 1980s to curb violence and substance use in schools (Curran, 2017). Mandatory suspension or expulsion, thought to enhance school climate and ensure students who want to learn can do so with less disruption in the classroom, have negative effects on suspended or expelled students and their peers. In a study on school discipline reforms, suspensions not only decreased math and reading scores for suspended students, but their peers also experienced small but statistically significant decreases in math and English language arts scores (Lacoe & Steinberg, 2019). Additionally, the counterfactual reality of what happens without intervention is hardly explored with scientific rigor. In randomized control trials, we control for a nontreatment effect and report on these conditions over time with multiple levels of follow-up. Although this “counterfactual” is difficult to identify in policymaking because the future state of the world is not yet known, it is equally important to analyze as the policy itself, providing the baseline against which the consequences of the intervention should be compared.

Anticipating and mitigating the side effects of well-intentioned interventions requires the use of methods that can account for the complex structures in which these efforts are made. Bryk et al. (2015) argued that improving educational systems for the long haul requires us to fully understand the problem before jumping to solutions. The authors then indicated that seeing the system that produces the current outcomes is necessary to adequately address the problems the system produces and mitigate any unexpected side effects (Bryk et al., 2015). These concepts, based on principles of systems thinking and its subset, system dynamics (SD), takes into consideration the interdependencies of a system and the effects of intervening on one part of the system to another (Richmond, 1993).

System dynamics is a method for conceptualizing and analyzing complex problems that are characterized by multiple interconnected and often nonlinear feedbacks, like those previously described. Using qualitative maps and formal computational models, SD accounts for the complex structure of systems and can be used to analyze the future effect of interventions over time (Richardson, 2011; Sterman, 2000). In contrast with linear models that are “open loop” (i.e. lacking feedback), SD is an approach that is characterized by (a) a broad model boundary that seeks to capture endogenously all the mechanisms relevant to the problem at hand, including (but not limited to) stakeholders of all types, the built environment, markets, and policies and regulations; (b) realistic portrayal of the decision-rules being employed by stakeholders in the system; and (c) explicit recognition of time

and the existence of time delays between cause and effect. This method of analyzing a system and its future behavior can be useful for understanding the root causes of problematic behaviors and designing appropriate interventions.

SD can be used to map the interconnections and processes that represent the way a system is structurally biased to create varying outcomes. In education, there is much discussion on how the system is structurally set up to produce inequitable outcomes for people based on race, gender, and other demographic factors. John Powell (2008) described structural racism as the interaction of multiple institutions in an ongoing process that produces racialized outcomes. He noted that a systems approach can help illuminate the ways in which individual and institutional behaviors interact across domains and over time to produce these side effects with clear racialized effects. By explicitly mapping how the system becomes racially, economically, and politically divided through enrollment patterns, investments, and dollars available per pupil, SD can demonstrate the structure of how racism and economic inequality work within the education system.

In this chapter, we review the use of SD for the integrative analysis of side effects from education policy interventions. First, we introduce the fundamentals of the SD methods for the modeling and analysis of complex systems, such as stocks, flows, balancing, and reinforcing feedback loops. We explain how these building blocks can be used to develop both qualitative models (“causal loop diagrams” [CLDs]) and quantitative (simulation) models. We then demonstrate the application of these tools by sharing the development of a model to address educational underperformance, teacher burnout, parent dissatisfaction, and declining enrollment in a U.S. public school system. Through this example, we discuss how to design and test interventions to address these issues. We end by discussing the limitations of this method and recommendations for pursuing educational reforms based on methods of rigorous SD to better understand the effects, both positive and negative, of reform efforts in complex educational systems.

MODELING THE BEHAVIOR OF COMPLEX SYSTEMS WITH SYSTEM DYNAMICS

Policy analysis is frequently conceptualized as a linear process that starts with problem definition and determination of evaluation criteria, then moves through identification and evaluation of possible policy interventions, and ends with the selection and implementation of the preferred policy (e.g., Patton et al., 2016). Yet this analytical framework is open loop in that it frequently fails to capture the multiple and often nonlinear feedbacks that exist in real-world systems and often results in policy prescriptions that have negative side effects. These effects can diminish the intended benefit or even exacerbate the problem in the worst case. The state of the world today is an endogenous consequence of the many decisions made by stakeholders in days past, who responded to the state of the world at the time. Similarly, the decisions being made today will shape the state of the world into the future in many

and varied ways. The more our analysis during policymaking reflects this endogenous reality, the more effective our policy prescriptions will be.

John Dewey and Lev Vygotsky, popular educational theorists, pointed to bidirectional feedback loops and complex relationships as key constructs in educational development (e.g., Dewey, 1916; Vygotsky, 1978). Advanced statistical tests have been applied to understand the bidirectional nature of these relationships, mostly with some form of structural equation modeling (Teo et al., 2013). For instance, multilevel structural equation modeling was used by Yarnell and Bohrnstedt (2018) to explore the association between student-teacher matches and Black student achievement and the relationship between social support, classroom management, and school adjustment (Aldrup et al., 2018). Latent change analysis, a type of structural equation modeling, was applied to understand the reciprocal effects of school capacity on student learning (Hallinger & Heck, 2010). These types of statistical applications can establish general associations between parameters and patterns over time and ways to represent nonrecursive relationships. However, these methods are inadequate to model issues of feedback loop dominance (Hovmand & Chalise, 2015). Additionally, structural equation modeling and similar analyses are not able to capture the complexity of multiple feedback loops operating over time in a closed system. Methods that account for these challenges can help better predict the outcome of efforts to intervene in complex systems such as education.

System dynamics is a set of modeling tools, qualitative and quantitative, that seeks to provide an improved understanding of the behavior of complex systems by formalizing the feedback structure of the system and builds on the principle that structure generates behavior. At the heart of SD models are causal loop diagrams—visual maps that articulate an explicit causal theory of how the variables in the system relate to each other. Each causal arrow in a CLD is assigned a polarity. A positive polarity indicates that the cause-and-effect variables move in the same direction. For example, an increase (or decrease) in teacher salaries leads to an increase (or decrease) in teacher retention, all else being equal. A negative polarity indicates that cause and effect move in opposing directions. For example, an increase in the number of teachers in a school leads to a decrease in student-to-teacher ratio, all else being equal. As the causal relationships present in the system are mapped, feedback loops emerge that give rise to dynamic and nonintuitive behavior. Feedback loops in which an initial change in a variable is further amplified are labeled as “reinforcing” loops. For example, it is a reinforcing feedback if an increase in student attainment leads to greater student satisfaction and effort, resulting in yet higher student attainment. When the effect of the feedback is to resist the initial change, the term “balancing” loop is used. For example, it is a balancing feedback loop if an increase in student attainment at a school attracts more students that results in an increase in enrollments and class sizes, leading to a decrease in teacher attention per student and a subsequent (relative) decrease in student attainment. Eliciting mental models in this way results in an explicit causal theory of how the system operates, with the CLD serving as an artifact around which the analysis of consequences, intended and unintended, can occur.

Initial insights are gained simply from mapping the structure of the system. Formalizing the model as a series of coupled differential equations and adding elements including accumulations (stocks) and time delays allows for simulation of the model's behavior over time, the testing of the model against historic data, and the explanation of alternative future scenarios (Sterman, 2000).

We do not attempt to provide a full description of how to build an SD model here, which is well beyond the scope of this chapter. Excellent learning resources are available, such as the seminal *Business Dynamics* textbook (Sterman, 2000). However, we do make some observations here about why SD is uniquely well suited to understanding the potential for side effects of education policy interventions:

- SD is a top-down modeling approach that emphasizes a broad model boundary, with all feedbacks relevant to the problem being represented endogenously within the boundary of the model. No single variable represents the “main effect” or dependent variable. And there are therefore no side effects—only consequences, both intended and unintended, as the interactions between feedback loops play out.
- Central to SD is a realistic portrayal of human decision-making. If the behavior of the model is to reflect reality, it is necessary that the decision rules encoded in the model faithfully represent those used by decision makers (e.g., how funding is allocated to schools and how the amount of funding changes as the number of students enrolled increases/decreases). It is critical that these decision rules can only make use of information available to decision makers in the real world, the so-called Baker criterion (Sterman, 2000). The potential for unintended consequences emerges naturally from the interaction of these boundedly rational decision makers, who may make the best decision possible given the information at hand but who do not and cannot appreciate the flow-on consequences of their decisions.
- The existence of significant time delays between cause and effect can be particularly problematic for traditional policy design. For example, changes in curriculum or pedagogy may take years to roll out and even longer to influence students' educational attainment, slowing the rate at which learnings about policy effectiveness can be gained and potentially leading to incorrect attribution about effectiveness. Although SD analysis cannot simplify these dynamics, it does require analysts to explicitly recognize the existence and extent of time delays, providing the tools needed to understand how consequences play out over different time scales.

Although SD has a long history of application in for-profit management settings, it is a methodology that is equally relevant to the analysis of nonprofit and public policy problems. For example, Keith et al. (2024) analyzed how pressure to deliver results and keep overhead low affects the decision-making of nonprofit managers, Landry and Sterman (2017) explored capability trap dynamics in the foster care system, and

Lim et al. (2022) modeled the opioid crisis in the United States to explore policy impacts and solutions. SD is a descriptive and human-centered approach that embraces the complexity and uniqueness of real-world systems, as distinct from more prescriptive approaches to the “scientific management” of education systems considered in decades past that frequently did not deliver projected improvements (Gough, 2010).

ENGAGING STAKEHOLDERS TO REVEAL THE STRUCTURE OF A SYSTEM

Anticipating the side effects of policy interventions in the education system requires an understanding of how current systems are structured such that they give rise to problematic behaviors over time. Causal mapping is one way to depict these structures for future formalization and testing. Once models are created, then we can test interventions. However, the accuracy of a system model depends on how and who was engaged in mapping it. To fully and most accurately depict the structure of a system, we must gain multiple perspectives from stakeholders engaged in the system and use a model as a boundary object to seek consensus (Hovmand, 2014). Although literature reviews and expert engagement can help with identifying causal links and ordering, the everyday experience of people living in a system can provide rich observational data that are often missing from extant published literature and analysis. Expertise, defined broadly, should not only include subject matter experts but also the perspectives of people who live through a system every day.

Lived experience experts, or LEEs, as referenced in the medical field, are people who develop a very special expertise by living with a challenge every day (Vázquez et al., 2023). They know, better than anyone else, how it affects their lives, what they must do to get through the system, how interventions affect them, and how symptoms or side effects impact their everyday lives and relationships (Vázquez et al., 2023). From teachers, cafeteria workers, bus drivers, students, and their families, each person in a system has a unique vantage point of identifying parts of the system that are integral to its functioning. These observational data provide a rich database of information that expands on what we know from formal parameters and data collected in the world. Forrester (1993) acknowledged that this mental database is rich with insight but is often missed and hard to elicit. To address the challenge of engaging with stakeholder knowledge and reasoning, which may also be bounded by scientific rationality, methods exist to elicit this knowledge in a way that is also meaningful and empowering to LEEs.

For instance, a participatory method of mapping and modeling with community members using SD modeling to understand and change systems from an endogenous perspective is called community based system dynamics (CBSD; Hovmand, 2014). CBSD is a facilitated process that elicits the mental models of people in a system that ultimately affect the way a system behaves. Using scripts or facilitated activities through group model-building sessions, CBSD engages people with experience in

the system, often those most marginalized by the system, to develop informal CLDs and formal stock and flow SD models. These maps make mental models explicit and test hypotheses about the logical implication of assumptions on system behavior with computer simulation (Hovmand, 2014).

The ethos and person-centered values of CBSD are admirable. On a practical level, however, it may be challenging for participants who are most marginalized by a system to engage over time. CBSD, as it is currently described, requires that people attend and participate in group model-building workshops. These sessions tend to be time-intensive (they range from 3 hours to a week-long session at multiple points in a project) and are perhaps impractical to request from people who have rigid work schedules, limited literacy skills, and constraints with childcare.

Expanding CBSD with the inclusion of data from published literature, quantitative and qualitative research elicits more robust SD models that can be tested to replicate historical patterns of behavior over time. Taking the steps to engage individuals in this way leads to a better understanding of the side effects of well-intentioned interventions. The next section describes a case study in which we use various types of data—including published data, subject matter experts, qualitative interviews, and CBSD group model building—to map the decline of student enrollment in public schools in St. Louis City over time.

ANALYZING DECLINING PUBLIC SCHOOL ENROLLMENT IN ST. LOUIS

At peak enrollment, the public school system in St. Louis City boasted 115,543 students in 1967. Today, it has just under 20,000. To understand this decline, we begin over a century ago with a revolution in St. Louis's public education system. St. Louis experienced rapid growth in the late 19th century, becoming the fourth largest city in the United States, following New York, Philadelphia, and Chicago. Accelerating industry led to population growth that in turn strained the city's public infrastructure system, which had been largely neglected during the Civil War years. Students at that time learned in unheated, dimly lit buildings. Playgrounds and sanitary conditions were rare, and in 1897, city officials contracted with renowned architect William Ittner to design a fleet of new public schools to support the city's growing population. Ittner schools were constructed to be conducive to learning, with spacious, sunlit common areas and classrooms. Forty-eight of 50 Ittner schools stand today, a testament to the city's past investment in its students.

Although the majority of Ittner's school buildings stand today, only remnants of the original vision remain as the city struggles to support its core infrastructure, including schools. The following century brought population decline: Between 1950 and 2010, the city's population declined by 63%. Destructive patterns of urban renewal left many neighborhoods empty as families fled to surrounding suburbs, leaving few students to sustain neighborhood schools (Taylor, 2013). The outflow of students higher on the socioeconomic ladder left a concentration of students and

families with more complex needs. The longest running desegregation program in the United States also contributed to the dwindling number of students in St. Louis Public Schools. By its end in 2019, St. Louis's desegregation program bussed 70,000 inner-city students to predominantly White schools in the surrounding suburbs (Associated Press, 2016). The deep economic well and population base that once supported construction of 50 new, state-of-the-art schools at the end of the 19th century was nonexistent by the end of the 20th.

Reform was attempted in 1998 when Missouri passed legislation enabling charter schools to open in two geographic locations: St. Louis and Kansas City. In 2000, the first charter school opened in St. Louis, quickly followed by 16 additional, independent charter networks between 2008 and 2023. A choice environment in St. Louis did not reverse declining enrollment trends in the public school system and drove some families to move to the surrounding suburbs. Under pressure, St. Louis Public Schools reorganized into a corporate-like structure epitomized by hiring a former Fortune 500 CEO as the superintendent in 2003. Fiscal efficiency was the name of the game during this period, and severe funding cuts were applied across special education services, curriculum development, teacher development programs, school counselors and social workers, and building maintenance (Dobbs, 2004).

The options available to St. Louis students today hardly compare to the relatively state-of-the-art buildings, teachers, and resources their 19th-century predecessors experienced. Both traditional districts and charter schools have bright spots of academic excellence, but those who enroll enter districts on the tail end of deep cuts to critical services. In a resource-scarce environment, tasked with educating students experiencing higher and higher rates of poverty, educators on both sides of the public school sector struggle to achieve acceptable academic outcomes with their students. Faced with another round of building closures due to declining enrollment, leadership from St. Louis Public Schools began conversations with charter school leaders interested in informal collaboration in 2019. Assembled educators began with the premise that no school district could single-handedly improve the system of public education in the city. Some level of collaboration and problem solving would be needed to address deeply entrenched issues.

St. Louis school leaders joined their peers in Washington, DC; Denver; Chicago; Cleveland; and Minneapolis (among others) that have formal district-charter cooperation. District-charter compacts, as they are known, seek to create a shared vision of education in a city, focus on long-term outcomes, adopt win-win activities, and seek common metrics used to evaluate success (Lake et al., 2017). Few incentives for school decision makers to work together currently exist. Structural barriers to collaboration, such as separate decision-making bodies and zero-sum state funding structures, increase perceived competition between schools. The result of this structure is a hard-to-navigate system of public schools in St. Louis City. Many families choose to opt out altogether and move to a surrounding suburb with one public school option—a significant reduction in choice and effort to enroll their student each successive school year. Spurred by accelerating enrollment decline plus the

increasing complexity of social and economic challenges facing students—partly due to the 2020 pandemic—leaders from St. Louis Public Schools and city charter schools formally embarked on an 18-month research and design process in 2022. School leaders sought to create a shared vision and plan for system change that sustainably transforms the way St. Louis City’s system of schools operates.

Problem Formulation

Formulation of a clear problem statement is one of the foundational activities for leading change efforts. Considering the discipline of problem formulation, researchers wrote that a clear problem statement can “unlock the energy and innovation that lies within those who do the core work of the organization” (Repenning et al., 2017). This is good practice for any organization: Without pausing to clearly define the problem, human nature is prone to leap straight from situation to solution. This is not inherently problematic when it is experts working under extreme time pressure doing the leaping—first responders or trauma surgeons, for example. Cognitive science warns us that solutions generated this way rely on patterns identified from experience. An inherent bias toward the status quo is introduced, and the likelihood of innovative solutions decreases. Repenning et al. (2017) commented on this phenomenon: “It should come as little surprise that breakthrough ideas and technologies sometimes come from relative newcomers who weren’t experienced enough to ‘know better.’”

A complement to fast thinking, described previously, is a more deliberate approach. The discipline of problem formulation entails developing a logical argument that links the observed data to root causes and eventually, to a solution (Repenning et al., 2017). A good problem formulation has five ingredients: It references something the organization cares about and connects that element to a clear and specific goal; it contains clear articulation of the gap between the current state and the goal; the key variables—the target, the current state, and the gap—are quantifiable; it is as neutral as possible concerning possible diagnosis or solutions; and it is sufficiently small in scope that you can tackle it quickly. (Repenning et al., 2017). School leaders and researchers defined their problem as such:

The system of public schools in St. Louis does not meet the expectations of City students and their families. A vicious cycle of enrollment decline exists as families seek alternative school options, leading to fewer dollars for schools to meet the needs of students, increasing teacher turnover due to burnout, and lowering community investment and overall confidence in the public education system.

Data Collection

Once our problem was defined, we recruited parents/caregivers, teachers, and school administrators with experience in the public education system in St. Louis City to participate in one-to-one interviews ($N = 96$ participants). Parents/caregivers were recruited in neighborhoods characterized by high concentrations of poverty. Due to the difficult nature of recruiting parents or guardians who are least likely to

engage in formal institutions, participants were recruited and interviewed using purposive and word-of-mouth sampling. We sought to include participants from groups historically underrepresented in educational decision-making, including Black parents and those with immigrant and refugee experiences. Additionally, we targeted parents living in majority non-White communities in north St. Louis City and sections of south St. Louis City (e.g., the South Grand area, Benton Park, Fox Park, and Dutch Town). Parent/caretaking participants who met the following characteristics were excluded from the study: parent/caregivers whose children had no public school experience, parent/caregivers whose children had no public school experience in the past 5 years, and parent/caregivers under 18 years of age. At the time of the interview, each participant filled out a demographic questionnaire and were interviewed face-to-face with a semistructured interview guide. All interviewees were compensated with a \$25 gift card for groceries or gas. Audio files of the interview were made and securely stored on an encrypted drive and then transcribed.

We conducted 78 semistructured interviews with parents and caregivers of a student enrolled in any St. Louis City public school. Eighty-one percent of parents or caregivers identified as Black or African American. Seventy percent had a student either currently or recently enrolled in the city's traditional district school, St. Louis Public Schools (SLPS). The remaining 30% of students were enrolled in one of the city's independent charter schools. Fifty-two percent of parents and caregivers interviewed were eligible for Supplemental Nutrition Assistance Program and/or Temporary Assistance for Needy Families, and a high school degree or GED was the highest level of education for 46% of parents or caregivers interviewed. Interviews started with basic questions: What school does your student attend? What influenced your decision to enroll? What do you like about your student's school? Interviewers then asked questions designed to elicit more detail: What is your student's teacher like? Tell us about your student's academic growth. Does your kid get any special support at school? What's that like? What makes a good school? What would make you leave a school? If you have left a school, tell us about your experience. Why do you think people leave their current school or city schools? Interviewers also asked questions about the interviewees' general perceptions of St. Louis City schools and their experience navigating traditional district, charter, private, and parochial options.

We then conducted 18 semistructured interviews with teachers, principals, and school administrators currently employed or recently departed from a public school in St. Louis City. Professionals with no public school experience, no public school experience in the past 5 years, or people under 18 years of age were excluded. We used similar recruitment methods, including word-of-mouth and purposive sampling. Interviews began after a demographic questionnaire was completed, and a semistructured interview guide was used. Interviews took place on Zoom and were recorded, securely stored on an encrypted drive, and transcribed.

Interviewees were compensated with a \$25 Amazon gift card. Sixty-one percent of the sample identified as White, 33% as Black or African American, and the remaining 5% as “other”—either Hispanic or Latinx, Asian Pacific Islander, or Indigenous American. Thirty-eight percent of participants were employed by SLPS, and 62% were employed by a charter school. Questions began with basics, such as the following: How long have you been a teacher/principal/educator? Why did you become an educator? They were then asked about their day-to-day duties: Do you feel that your school-day duties are realistic? If you put in more effort than what is expected from your job description (e.g., staying late, offering out-of-scope support), how does it affect your students? What kind of support do you provide and/or receive? Educators were also asked about their relationships at school: How do you approach working with the parents/caregivers of your students? In your experience, what has worked well in building trust with parents? What are the relationships between school staff like at your school? Interviews also included questions about the educator’s perception of workload, what would make it sustainable, support they wish staff had, and reflections on their training and preparedness for their role. Interviews concluded with the question, “If you could tell school leadership anything, what would you say?” All responses were de-identified.

Themes were found using grounded theory (Charmaz, 2006). Grounded theory offers systematic yet flexible guidelines for making sense of qualitative interviews through open and axial coding and analysis. The themes gathered from the coding process were used to outline a conceptual diagram of the underlying structure driving the decline in educational attainment and enrollments in SLPS. The themes included: a) challenges navigating school options; b) parental trust in schools precedes changes in enrollment; c) parent trust in teachers, however, may affect decisions to stay even if trust in the school decreases; and d) parents consider high quality teaching to include more factors than a student’s academic progress. Parents reported that they look for teachers who work with them, communicate regularly, take time to work with the students one-on-one, and offer individualized attention.

Model

In this section, we demonstrate how our fieldwork informed the development of a CLD that depicts the system that leads to declining enrollment of students and achievement in the public school system of St. Louis City. Using SD qualitative mapping, we demonstrate the feedback loops, or circular causal relationships between variables where increases or decreases in one variable have self-reinforcing or self-correcting effects on the other and then itself over time, to understand nonlinear system behaviors (Richardson, 2011; Sterman, 2000). These nonlinear behaviors create side effects of interventions and help us determine what we could expect to happen over time.

The Public School Capability Trap

At the core of our model is school performance, which is measured by educational attainment. Building on the organizations literature, which has long recognized that organizational capabilities (e.g., processes, routines, relationships, and institutional knowledge) are a key driver of organizational performance (e.g., Gibbons & Henderson 2012; Rahmandad et al., 2018; Repenning & Sterman 2002), we conceptualize school performance as a function of its educating capabilities and its effort spent on addressing immediate student needs (Figure 1). That is, performance can be increased by either investing in the improvement of educating capabilities or by putting additional effort into day-to-day work with students. Educating capabilities is represented as a stock (accumulation) variable, shown by the rectangle around the variable name, emphasizing that capabilities must accumulate over time. The stock of educating capabilities increases with the inflow capability building and declines with capability erosion (e.g., due to forgetting, staff turnover, and redundancy).

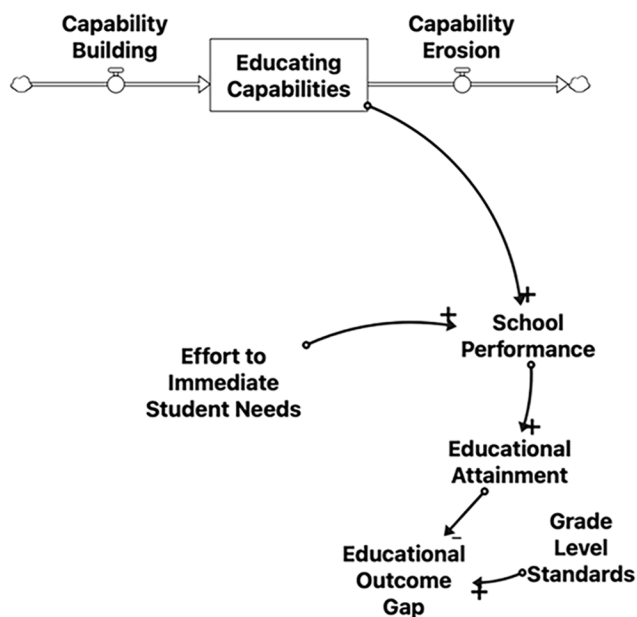
The educational outcome gap that is the key driver of school effort is conceptualized as the difference between educational attainment of the school and the grade-level standards that they are trying to meet. The St. Louis public schools are facing significant attainment gaps. This came up repeatedly in our interviews and constantly is on the minds of educators. For example:

I know that as the years have progressed, I am handed students less literate. So they're not ready for third grade. They're not even ready for second grade. . . . It's like, you don't want to blame the teachers below you. But you're also like, "What are you doing?" The fourth-grade teachers are feeling like that to me when they get my kids, and then it's just this, "Well, you should have seen them when I got them."
(Teacher 37)

This is the narrative we heard in most of the schools within SLPS—teachers and administrators experience the weight of educational attainment gaps, often in the face of increasingly inadequate resources. As the gap increases, there is greater pressure on teachers and administrators to respond by working ever harder for students on a day-to-day basis. The teachers in SLPS can be found staying after school with students and working overtime hours to manage the many small issues that emerge throughout the school day. High levels of teacher effort can help close attainment gaps in the short term and is appreciated by parents; we capture this response in the “working harder” balancing loop B1 (Figure 2). As intended, spending more time and energy doing work does help with educational attainment and helps to close the performance gap in the short term.

Another way to respond to the attainment gap is to invest more effort into developing school capabilities so that each hour of effort committed by teachers is more effective at delivering improved school performance. In our interviews, it was clear that teachers and school leaders are well aware that time spent investing in improvement will yield long-term returns:

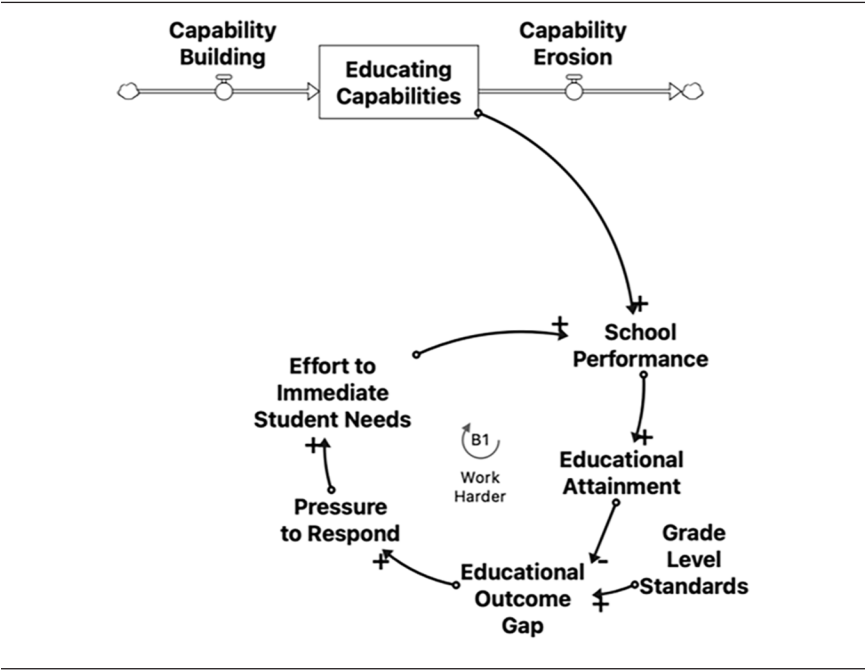
FIGURE 1
The Educational Outcome Gap



I think for a lot of us, and this goes for teachers and admin, you know, anybody on the staff, there are immediate issues that crop up, that will take us away from, you know, doing that work. It's like, "Oh, we have a disciplinary problem to deal with here," or there's a parent or, you know, something came up. And so that kind of chips away at really feeling secure. We need to deal with those immediate needs, rather than these higher function needs. (Teacher 32)

Investing effort to build capabilities is also a balancing process because investment in programs such as teacher training, mentoring, and curriculum development are all intended to increase the stock of educating capabilities the school possesses, increasing school performance and closing the attainment gap. We capture this mechanism as the "working smarter" feedback loop B2 in Figure 3. However, the previous quote reveals two further insights about the consequences of how teachers choose to allocate their available time. First, capabilities take time to accumulate ("doing that work"), with a temporal difference between responding to immediate needs (e.g., behavioral issues) that has immediate but often ephemeral rewards versus making a time investment that will pay off in a sustained way in the longer term but deliver little in the short term. Second, a resource trade-off exists because with finite human capital available, time spent responding to immediate student needs (i.e., working harder) necessarily takes

FIGURE 2
Responding to Student Needs by Working Harder



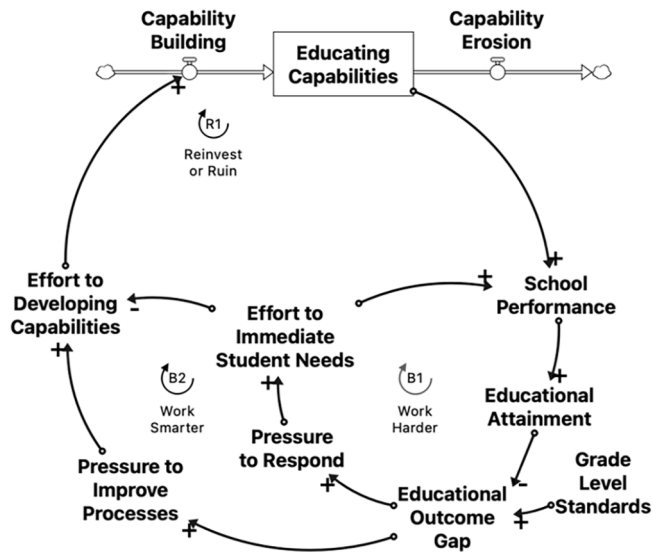
teachers away from investing in capabilities (i.e., working smarter), just as choosing to invest in capabilities would take teachers away from responding to immediate student needs, a choice that many teachers would understandably feel is impossible to make. We capture this trade-off between working harder and working smarter with the “reinvest or ruin” reinforcing loop R1 in Figure 3.

The reinvest or ruin reinforcing loop is particularly pernicious, capturing the unintended consequences of a reliance on working harder over working smarter. It is entirely understandable that educators have made working harder the norm, given the immediate and salient need that St. Louis students have, when the alternative, working smarter, delivers uncertain benefits out into the future. The teachers at SLPS understand this all too well, reporting that investments in their own development as educators were few and far between:

Okay, well, with social emotional learning [SEL], like, yeah, let's get an SEL curriculum, this is going to fix everything. But then no one's the expert on it. And so I'm like, okay, [I] really see that this is a great thing, I'm gonna try it out. But also [I'm] responsible for doing all this curriculum. (Teacher 37)

The decision to respond to an educational outcome gap is not surprising, with similar responses having been observed in other industries, such as manufacturing and

FIGURE 3
The Trade-Off Between Working Harder and Working Smarter



quality management (Repenning & Sterman, 2001). But the critical unintended consequence of spending more time working harder is that less time is spent working smarter. If the rate of capability building falls below the rate of capability erosion, the level of educating capability will decline over time, leading to a reduction in educational attainment, an increase in the educational outcome gap, and yet more pressure to work harder. Repenning and Sterman (2001) called this the “capability trap”—simultaneously doing more and getting less because the organization lacks capabilities to effectively translate effort into results and no free capacity to reinvest in capability building.

Additional Reinforcing Loops in Public Education

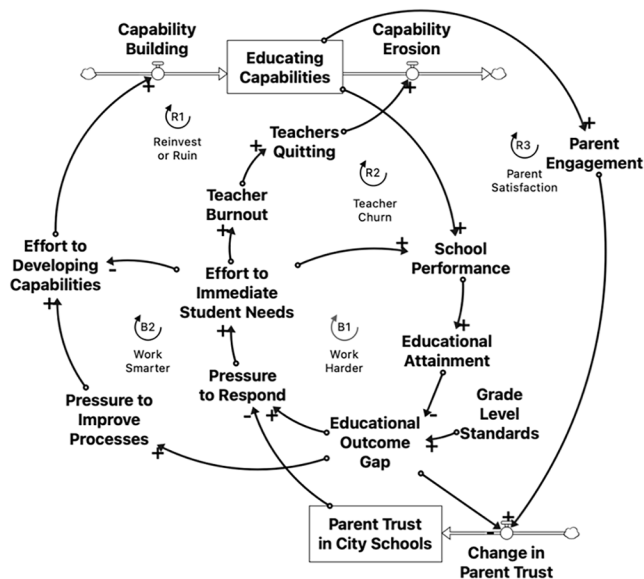
Although the aforementioned capability dynamics echo what have been observed in other industries, our fieldwork with SLPS revealed the existence of several other reinforcing feedbacks that amplify these capability trap dynamics.

First, the notable commitment of teachers to work harder without enrichment of resources to support their work eventually causes burnout. It became clear in our fieldwork that the teachers did not feel that they had the resources they needed to perform in their roles. The teachers at SLPS observed burnout and quitting across the school district, as people sought more attractive jobs or left the profession entirely:

Quitting presents several challenges to the school. First, it quickens the decay of capabilities because the school loses critical institutional knowledge and experience. Furthermore, it has a negative impact on educational attainment of students due to a less cohesive education staff and higher student-teacher ratios. Lower educational attainment leads to higher demands on the remaining teachers, which reinforces burnout and little to no investment in capabilities. We conceptualize this reinforcing mechanism with the “teacher churn” loop (Figure 4).

Second, another critical reinforcing mechanism is the engagement of parents as educational attainment and parent engagement evolves. A core part of educating capabilities is a school's ability, via teachers and school staff and administrators, to communicate with and engage parents in the education process. The higher the capabilities of a school, the better it can maintain ties with parents and facilitate engagement. Parent pickups and drop-offs where administrators greet the parents or social media and translation services for English as a second language families create connections between the school and parents and increase engagement, which is a key driver of parents' trust in their schools. Teachers are critical to engagement because they are often the representatives of the school who have the most interaction,

FIGURE 5
The Effect of Educational Outcomes and Parent Engagement on Parent Trust



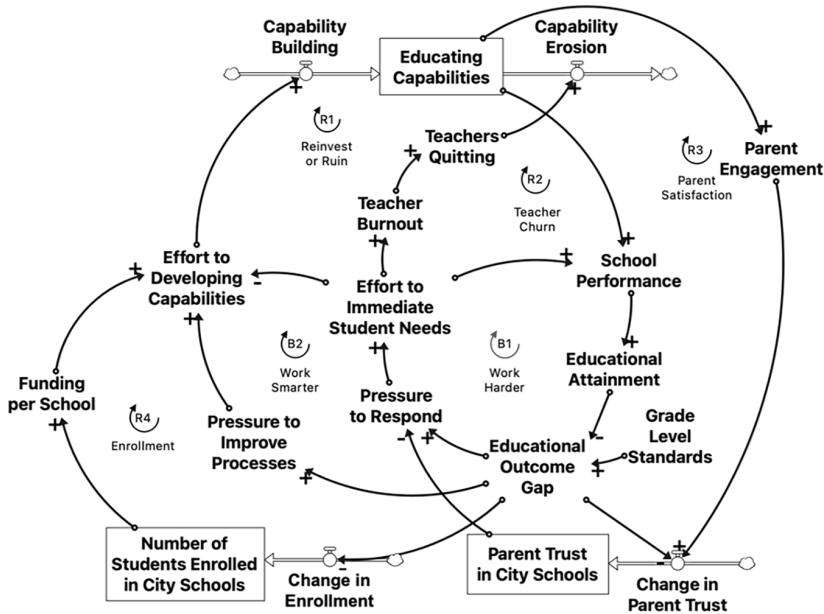
formally through planned communication and informally via student perceptions and stories, with families. When parents have low levels of trust in city schools, this frustration often gets channeled toward teachers and increases teachers' pressure to respond. Teachers then spend more of their time trying to meet the immediate needs of students and satisfy parents in the short term at the expense of long-term improvement, thus completing the "parent satisfaction" reinforcing loop (Figure 5).

Third, educational attainment drives parents' decisions to enroll or unenroll their children. The "enrollments" reinforcing loop (Figure 6) illustrates that when the school is not performing, those who can leave the system do leave the system. We spoke to one parent who, when asked why she moved her children out of the city, said:

Because I wanted them to have a better education. You know people say, "You're a traitor." I said "no." Because if you go out there and look at the city schools, the city schools are not doing what the county schools are. You don't hear from a city school. (Parent 53)

Families leaving the district contributes to underfunding and fewer resources for schools because a significant fraction of school funding is variable with student enrollments while fixed costs and complexity of student needs remain high. This underfunding leads to even less effort dedicated toward capability building, thus reinforcing the low-performance equilibrium.

Figure 6
The Effect of Educational Outcomes on Student Enrollment in City Schools

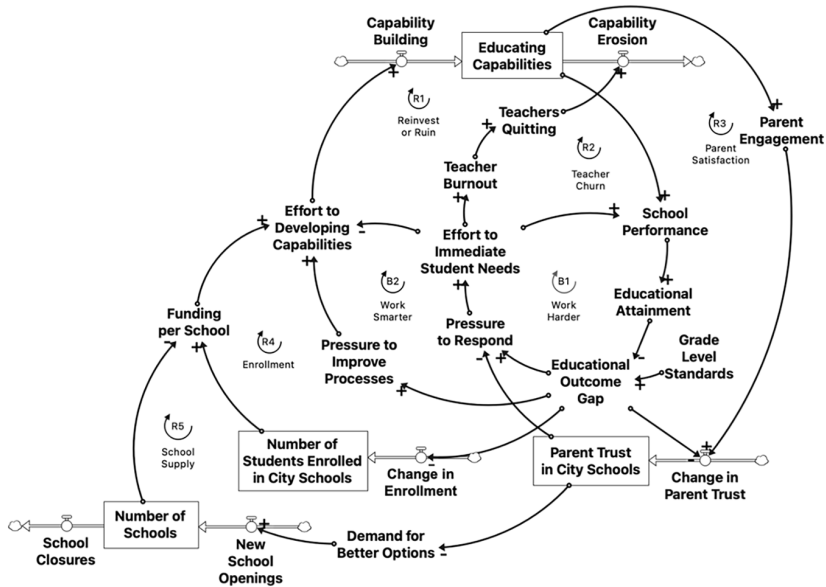


Fourth, over time, as parent trust erodes, the perceived need for new schools grows, as independent actors and their sponsoring organizations open new schools to address performance gaps. We see empirically that the number of schools in the public school system is increasing over time. Adding schools to the system reduces the number of children per school, which leads to less funding available per school, and adding more fixed costs to the system reduces the available funding, thus reducing the resources available to build capabilities. This mechanism is described through the “school supply” reinforcing loop (Figure 7).

Policy Analysis and Learnings About the Potential Unintended Consequences

Using SD to make mental models explicit, working in partnership with people in different vantage points of the system and pairing it with published literature, data, and rigorous analysis, can help us understand how the structure of the system works. This type of practice can aid in the ever-evolving process of reform for districts and educational systems. The district level, in particular, has had little examination as a subject of inquiry (Gamson & Hodge, 2016), let alone as a subject by which change or interventions can be systematically examined for side effects and long-term collateral changes.

FIGURE 7



In our study of the public school system in St. Louis, where the historical effects of racial discrimination, segregation, and poverty have influenced the resources available for growth, the structure of the system challenges leaders' efforts to improve conditions for students. In this environment, enrollment decline in all schools is a function of diminishing parent trust leading to the hope for more alternatives, the development of new schools and charter schools in particular, and thereby the spread of already thin financial resources to the system as a whole. In the absence of efforts to combine or close existing schools with dwindling enrollment, student movement out of schools constrains the amount of funds available to adequately meet the needs of students and educators in the system, diminishing the time and resources available to build educational system capabilities, leading to more "fires" that need fighting. This vicious cycle, resulting from the well-intentioned decisions of parents, teachers, and funders, will continue to diminish public school enrollment in the city without major interventions. It is important to recognize, however, that these same reinforcing loops can operate in a virtuous manner if conditions in the school system change. Exactly the same reinforcing loops that have driven this system to vicious decline could operate in a virtuous manner: If enrollments increase, so does school funding, allowing for more investment in capabilities, higher school performance, greater

educational attainment, and less short-term pressure, allowing for yet more investment in educating capabilities. The challenge, however, is how to reverse the sustained decline in enrollments occurring currently.

Policy implications of this work are numerous. First, focusing on the problem definition phase of this work helped us understand the various ways in which side effects are interconnected to other challenges in the system. For example, focusing on declining enrollment helped school leaders across charter and district schools to ground themselves in a shared operational concern. With a pool of fewer children to recruit from, more aggressive tactics to attract and retain students have been used to market to a dwindling number of families. The side effect of efforts to enroll children was leading to increased funding spent on marketing, with less dollars available to invest in educational operations, leading to less resources to address educational achievement and thereby more parents pulling their students out of schools. Similar issues have been observed in New York City, where one impact of school choice was increasing competition for student applicants, leading to increased per-pupil expenditures on noninstructional functions and thereby reducing resources for instruction and educational improvements (Rothbart, 2020).

The use of SD with the community members most affected by the educational system provides a map that has allowed us to extensively hypothesize about possible side effects of interventions. Causal loop diagramming forces us to articulate an explicit causal theory articulating how different parts of the system interact, helping us to think through all the parts of the system that will be affected—consequences that might not have otherwise been considered. Doing so allows us to understand *ex ante* the possible consequences of policy interventions and to mitigate potential side effects before they occur. Although there is a natural inclination to want to predict benefits and costs during the policymaking process, all predictions about the future (far into the future in particular) will be wrong because the state of the world will inevitably change, and the model used to make the prediction is necessarily a simplification of the real world. We believe in focusing instead on developing improved mental models that will allow stakeholders to make better and robust policy decisions in whatever circumstances they find themselves in. Beyond the scope of this chapter, the further specification of the CLD as a formal simulation model and calibration of the model to data allow for the consequences of various policy interventions to be quantified. Doing so allows decision makers to experiment freely in a virtual world, testing out policies that may be costly, unethical, or politically infeasible (Sterman 2014), developing improved policy prescriptions that have fewer side effects.

LIMITATIONS OF SYSTEM DYNAMICS

Like all analytical methods, SD is not a panacea, and it is just as important to recognize the types of problems for which SD is not well suited just as much as the types of problems for which SD is well suited. SD analysis is most particularly insightful for problems that are characterized by “dynamic complexity”—the existence of

evolving and often counterintuitive patterns of behavior over time, including (but not limited to) S-shaped growth, oscillations, and overshoot-and-collapse. These dynamics result from the interplay of feedback loops, nonlinearities, accumulations, and time delays, system-level mechanisms that are not well captured in other analytical methods, such as optimization and regression (Yasarcan, 2023). SD is less well suited to informing problems characterized by “detail complexity”—how to schedule a school timetable that ensures that all students can attend all classes given the available teachers and classrooms, for example. The level of model (dis)aggregation chosen also matters. Rahmandad and Sterman (2008) compare differential equation (DE) models classically used in SD with agent-based models, a related bottom-up simulation methodology that explicitly represents individual agents with their own preferences and decision rules, finding that aggregate DE models are computationally efficient and flexible but are relatively less able to capture high degrees of agent heterogeneity and network structure (physical and social) in how agents interact.

CONCLUSION

Education reforms and research on interventions should include an explicit discussion of anticipated side effects of intervention (Bryk, 2015; Zhao, 2017); this chapter suggests system dynamics as a method to map the interconnected feedback structures to aid one’s understanding of side effects as a result of intervention. Currently, methods to report side effects are either based on speculation prior to implementation or often poorly reported in post hoc analysis. Using public school enrollment decline as an example, this chapter demonstrates how causal loop diagrams and feedback models can add to existing education research practices by accounting for complex, interrelated feedback relationships and its effect on other areas of the system over time. Limitations of this modeling approach include inherent issues with model confidence and accuracy of causal links. Additionally, challenges with systems reorganizing over time can limit one’s ability to accurately predict side effects long into the time horizon.

Future efforts to test educational interventions using system dynamics should (a) clearly define the problem(s) one is seeking to change over time, (b) prioritize mapping efforts that are informed by multiple stakeholders experiencing the system to most accurately depict the system for testing, and (c) include mathematical specification of causal feedback relationships with existing or gathered data to build confidence in one’s model’s accuracy to predict future dynamics. When applied to pernicious challenges in educational improvement, system dynamics can move the field forward to more rigorous testing prior to implementation, which can potentially reduce harm to students, educators, and communities. Additionally, this method of modeling for possible side effects helps the field gain a better understanding of the trade-offs of intervention. With few examples of this type of educational intervention testing in the field, there are numerous areas of application, such as teacher retention, workforce development, preschool scale-up, or educational persistence.

ACKNOWLEDGMENTS

We are grateful to the St. Louis School Leaders Collaborative for their insights developing the system dynamics model in this chapter, specifically, Candice Carter-Oliver, Kelly Garrett, Kelvin Adams, Christie Huck, Sarah Christman, Isaac Pollack, Sarah Ranney, Katrice Noble, and Meghan Hill. We also thank Janet Velasquez, who coordinated interviews and workshops with parents and teachers.

ORCID IDS

David R. Keith  <https://orcid.org/0000-0003-4938-0617>

Saras Chung  <https://orcid.org/0000-0003-4432-1407>

REFERENCES

- Aldrup, K., Klusmann, U., Lüdtke, O., Göllner, R., & Trautwein, U. (2018). Social support and classroom management are related to secondary students' general school adjustment: A multilevel structural equation model using student and teacher ratings. *Journal of Educational Psychology, 110*(8), 1066–1083. <https://doi.org/10.1037/edu0000256>.supp
- Associated Press. (2016, November 29). St. Louis area to phase out desegregation program. *Education Week*. <https://www.edweek.org/leadership/st-louis-area-to-phase-out-desegregation-program/2016/11>
- Bogin, A., & Nguyen-Hoang, P. (2014). Property left behind: An unintended consequence of a No Child Left Behind “failing” school designation. *Journal of Regional Science, 64*(5), 788–805.
- Borokski, J. W., & Sneed, M. (2006). Will NCLB improve or harm education? *Harvard Educational Review, 76*(4), 503–525.
- Bryk, A. S. (2015). 2014 AERA Distinguished Lecture: Accelerating how we learn to improve. *Educational Researcher, 44*(9), 467–477. <https://doi.org/10.3102/0013189X15621543>
- Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Harvard Education Press.
- Cawelti, G. (2006). The side effects of NCLB. *Educational Leadership, 64*(3), 64–68.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. SAGE Publications.
- Curran, C. (2017). The law, policy, and portrayal of zero tolerance school discipline: Examining prevalence and characteristics across levels of governance and school districts. *Educational Policy, 33*(2), 319–349. <https://doi.org/10.1177/0895904817691840>
- Darling-Hammond, L. (2007). Race, inequality and educational accountability: The irony of ‘No Child Left Behind.’ *Race Ethnicity and Education, 10*(3), 245–260.
- Dee, T., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management, 30*(3), 418–436.
- Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. MacMillan.
- Dobbs, M. (2004, December 5). Corporate model proves an imperfect fit for school system. *Washington Post*. https://www.washingtonpost.com/wp-dyn/articles/A35712-2004Dec4.html?itid=lk_inline_manual_107
- Elmore, R. F. (2004). *School reform from the inside out: Policy, practice, and performance*. Harvard Education Press.
- Emler, T. E., Zhao, Y., Deng, J., Yin, D., & Wang, Y. (2019). Side effects of large-scale assessments in education. *ECNU Review of Education, 2*(3), 279–296.

- Forrester, J. W. (1993). System dynamics and the lessons of 35 years. In K. B. De Greene (Ed.), *A systems-based approach to policymaking* (pp. 199–240). Springer. <https://doi.org/10.1007/978-1-4615-3226-2>
- Gamson, D. A., & Hodge, E. M. (2016). Education research and the shifting landscape of the American school district, 1816 to 2016. In P. A. Alexander, F. J. Levine, & W. F. Tate, IV (Eds.), *Review of research in education* (Vol. 40, pp. 216–249). Sage. <https://doi.org/10.3102/0091732X16670323>
- Gibbons, R., & Henderson, R. (2012). Relational contracts and organizational capabilities. *Organization Science*, 23(5), 1350–1364.
- Goldrick-Rab, S., & Mazzeo, C. (2005). What No Child Left Behind means for college access. In *Review of research in education* (Vol. 29, pp. 107–129). Sage. <https://doi.org/10.3102/0091732X029001107>
- Gough, N. (2010). Can we escape the program? Inventing possible-impossible futures in/for Australian educational research. *The Australian Educational Researcher*, 37(4), 9–42.
- Hallinger, P., & Heck, R. H. (2010). Collaborative leadership and school improvement: Understanding the impact on school capacity and student learning. *School Leadership and Management*, 30(2), 95–110. <https://doi.org/10.1080/13632431003663214>
- Hovmand, P. S. (2014). *Community based system dynamics*. Springer Science + Business Media. <https://doi.org/10.1007/978-1-4614-8763-0>
- Hovmand, P. S., & Chalise, N. (2015). Simultaneous linear estimation using structural equation modeling. In H. Rahmandad, R. Oliva, & N. D. Osgood (Eds.), *Analytical methods for dynamic modelers* (pp. 71–94). The MIT Press.
- Keith, D. R., Taylor, L., Paine, J., Weisbach, R., & Dowidowicz, A. (2024). When funders aren't customers: Reputation management and capability underinvestment in multi-audience organizations. *Organization Science*, 35(2), 387–404. <https://doi.org/10.1287/orsc.2022.1579>
- Lacoe, J., & Steinberg, M. P. (2019). Do suspensions affect student outcomes? *Educational Evaluation and Policy Analysis*, 41(1), 34–62.
- Lake, R., Yatsko, S., Gill, S., & Opalka, A. (2017). *Bridging the district-charter divide to help more students succeed*. Center on Reinventing Public Education.
- Landry, E., & Sterman, J. (2017). The capability trap: Prevalence in human systems. In *35th International Conference of the System Dynamics Society* (pp. 963–1010). System Dynamics Society. <https://proceedings.systemdynamics.org/2017/proceed/papers/P1325.pdf>
- Lim, T. Y., Stringfellow, E. J., Stafford, C. A., DiGennaro, C., Homer, J. B., Wakeland, W., Eggers, S. L., Kazemi, R., Glos, L., Ewing, E. G., Bannister, C. B., Humphreys, K., Throckmorton, D. C., & Jalali, M. S. (2022). Modeling the evolution of the US opioid crisis for national policy development. *Proceedings of the National Academy of Sciences*, 119(23), Article e2115714119. <https://doi.org/10.1073/pnas.2115714119>
- McDonnell, L. M. (2008). The politics of educational accountability: Can the clock be turned back? In K. E. Ryan & L. A. Shepard (Eds.), *The future of test-based educational accountability* (pp. 47–67). Routledge.
- Meyer, J. W. (1977). The effects of education as an institution. *American Journal of Sociology*, 83(1), 55–77. <https://doi.org/10.1086/226506>
- Morgan, H. (2022). Resisting the movement to ban critical race theory from schools. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 95(1), 35–41. <https://doi.org/10.1080/00098655.2021.2025023>
- Patton, C. V., Sawicki, D. S., & Clark, J. J. (2016). *Basic methods of policy analysis and planning*. Routledge.
- Powell, J. A. (2008). Structural racism: Building upon the insights of John Calmore. *North Carolina Law Review*, 86(3), Article 791. <https://scholarship.law.unc.edu/nclr/vol86/iss3/8>

- Rahmandad, H., Henderson, R., & Repenning, N. P. (2018). Making the numbers? “Short termism” and the puzzle of only occasional disaster. *Management Science*, 64(3), 1328–1347.
- Rahmandad, H., & Sterman, J. (2008). Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science*, 54(5), 998–1014.
- Repenning, N. P., Kieffer, D., & Astor, T. (2017). The most underrated skill in management. *MIT Sloan Management Review*, 58(3). <https://sloanreview.mit.edu/article/the-most-underrated-skill-in-management/>
- Repenning, N. P., & Sterman, J. D. (2001). Nobody ever gets credit for fixing problems that never happened: Creating and sustaining process improvement. *California Management Review*, 43(4), 64–88.
- Repenning, N. P., & Sterman, J. D. (2002). Capability traps and self-confirming attribution errors in the dynamics of process improvement. *Administrative Science Quarterly*, 47(2), 265–295.
- Richardson, G. P. (2011). Reflections on the foundations of system dynamics. *System Dynamics Review*, 27(3), 219–243.
- Richmond, B. (1993). Systems thinking: Critical thinking skills for the 1990s and beyond. *System Dynamics Review*, 9(2), 113–133. <https://doi.org/10.1002/sdr.4260090203>
- Rothbart, M. W. (2020). The impact of school choice on public school budgets: Evidence from open enrollment in New York City. *Public Budgeting & Finance*, 40(4), 3–37. <https://doi.org/10.1111/pbaf.12267>
- Rury, J. L., Belew, R., & Hurst, J. (2022). The origins of American test-based educational accountability and controversies about its impact, 1970–1983. *Teachers College Record*, 124(1), 143–163. <https://doi.org/10.1177/01614681221086094>
- Smith, J. M., & Kovacs, P. E. (2011). The impact of standards-based reform on teachers: The case of ‘No Child Left Behind.’ *Teachers and Teaching*, 17(2), 201–225. <https://doi.org/10.1080/13540602.2011.539802>
- Sterman, J. (2014). Interactive web-based simulations for strategy and sustainability: The MIT Sloan LearningEdge management flight simulators, Part I. *System Dynamics Review*, 30(1–2), 89–121.
- Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Irwin McGraw-Hill, Boston.
- Taylor, A. N. (2013). Segregation, education, and blurring the lines of division in St. Louis. *St. Louis University Public Law Review*, 33(1), Article 11. <https://scholarship.law.slu.edu/plr/vol33/iss1/11>
- Teo, T., Tsai, L. T., & Yang, C. C. (2013). Applying structural equation modeling (SEM) in educational research: An introduction. In M. S. Khine (Ed.), *Application of structural equation modeling in educational research and practice* (pp. 3–22). Sense Publishers.
- Vázquez, E., Kim, M., & Santaella, M. E. (2023). Lived experience experts: A name created by us for us. *Expert Review of Hematology*, 16(Suppl. 1), 7–11. <https://doi.org/10.1080/17474086.2023.2178410>
- Vygotsky, L. S. (1978). *Mind in society* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.
- Yasarcan, H. (2023). The four main elements of dynamic complexity. *System Dynamics Review*, 39(2), 171–179.
- Yarnell, L. M., & Bohrnstedt, G. W. (2018). Student-teacher racial match and its association with Black student achievement: An exploration using multilevel structural equation modeling. *American Educational Research Journal*, 55(2), 287–324. <https://doi.org/10.3102/0002831217734804>

- Zhao, Y. (2017). What works may hurt: Side effects in education. *Journal of Educational Change*, 18(1), 1–19.
- Zhao, Y. (2019). Side effects in education: Winners and losers in school voucher programs. *Phi Delta Kappan*, 100(5), 63–66. <https://doi.org/10.1177/0031721719827553>

AUTHORS

DAVID R. KEITH, PhD, is an associate professor of strategy at Melbourne Business School. His research and teaching focuses on the use of system dynamics modeling for effective strategy and policy decision-making in complex operating environments. He has published articles on a diverse range of topics, including technology strategy, supply chains, platform markets, organizational capabilities, and public policy.

AISHWARYA YADAMA is a PhD candidate in the System Dynamics Group at the MIT Sloan School of Management. Her research focuses on economic mobility of women in low-wage work and diversity practice adoption.

ELLEN O'NEILL is the director of strategy and operations at SKIP Designed. She dedicates her work to catalyzing systemic transformation for children, families, and communities historically marginalized by past change efforts.

SARAS CHUNG, PhD, MSW, is a faculty affiliate at Washington University in St. Louis and the executive director of SKIP Designed, LLC. Her research focuses on adolescent socio-emotional support in schools. She serves as the vice president of professional practice for the International System Dynamics Society and is a member of the Complexity Theories in Education SIG of the American Educational Research Association.