Here is why I have become hooked on determining what “success” and “failure” mean in American public schools.

In the mid-1960s, I taught in and later directed a federally funded teacher-training program located in several Washington, DC, public schools.¹

The Cardozo Project in Urban Teaching (CPUT), as it was then called, prepared in one year returned Peace Corps volunteers to teach in urban elementary and secondary schools. The “interns” taught for half-days under the supervision of master teachers in elementary school grades and English, science, math, and history in junior and senior high schools. They took on-site university-sponsored seminars after school, and in evenings developed curriculum materials for their lessons and worked in the community. At the end of the year the CPUT “interns” were certified to teach in the District of Columbia and were on the way to earning a master’s degree in their field through local universities.²

An independent evaluation confirmed that sixty-one interns had completed training between 1963 and 1967. Of the fifty-six who had finished (two had died and three left the program for various reasons), forty-two (or 75 percent) were teaching in urban schools, other federally funded programs, or overseas—one goal of the program. The on-site training, the supervision by DC master teachers, and after-school seminars involving both interns and regular faculty from the participating schools were put into practice and seemed to be a fruitful mix for channeling rookie teachers into the system. The evaluation and praise for the program led
the DC school board to fund and rechristen the program as the Urban Teacher Corps in 1967. Getting a school board to use its limited monies to continue a federally funded pilot program meant that school officials saw its worth in attracting a different pool of teaching candidates.³

Consider further that the CPUT model of recruiting and training new teachers became the poster-child for funding a federal initiative to prepare teachers nationally for high-poverty urban and rural schools. The National Teacher Corps legislation (1966) adopted the model used by CPUT for training teachers on-site, but rather than fund districts, federal officials funneled monies to universities that took responsibility for the training program and awarding degrees.⁴

Surely, the pilot program had achieved its goals: three of four interns became full-time teachers after completing the program. And the program was adopted by an urban district using locally budgeted funds. Accomplishing both goals suggests program effectiveness, a sign of clear success. That the pilot program became the model for a National Teacher Corps further cements the sweet smell of success.

As a twenty-nine-year-old teacher/director of the project, I did inhale that sweetness. After leaving the project to return to teaching high school history, I believed I had been part of a “successful” reform. And I felt good about my part in the innovation.

There is a “however” to this seeming “success” story that needs to be noted. After the Urban Teacher Corps became part of the DC schools, the board of education fired its superintendent and in 1970 appointed Hugh Scott as its first black superintendent. During Scott’s brief administration (he resigned in 1973), he dismantled the Urban Teacher Corps. Together the federally funded pilot program and locally funded UTC existed for just under a decade.

Similarly, with the election of Ronald Reagan in 1980, the National Teacher Corps disappeared as federal monies for education went to the states in block grants rather than to specific programs. The NTC had lasted just over a decade.

The disappearance of both teacher-training programs within ten years suggests “failure” even though they could be fairly characterized as “successes” in achieving their primary goals.
Program “success” or “failure”? If it is a binary either/or conclusion, then determining whether CPUT should go down in the books as a “successful” reform in a district recruiting and training new teachers and then have it disappear raises questions about how one defines program “success”—Being adopted? Fully implemented? Achieving goals? Longevity? Or are there gray areas in defining “success” that seldom get attention?

I want to add two pieces to finish the story.

First, an assumption (not an explicit goal) was that the pilot project with ten returned Peace Corps volunteers, seven of whom were white and three were African American, would somehow influence staffing Cardozo Project schools whose faculties were then 80–90 percent black. That is, a majority white teacher-training project would produce, over time, an integrated faculty at each of the schools.

Changes in faculty demographics did occur in subsequent decades but can hardly be attributed to CPUT or the subsequent Urban Teacher Corps. In 2008, for example, Cardozo High School was closed for persistently poor academic performance. It reopened with different teachers hired to replace existing ones. As a result, percentages of white teachers increased to over one-third in 2009. While in 2017 the district has 32 percent white administrators and teachers, the original schools in CPUT continue to have predominantly black staffs.

Second, before leaving the Cardozo Project in Urban Teaching in 1967, I asked a professor at the University of Maryland to find out if the students in classes of elementary and secondary school interns achieved less, about the same, or more than students who had nonintern teachers. While raising student achievement was not an explicit goal for the Cardozo Project in Urban Teaching or the Urban Teacher Corps, there was a staff assumption (which I shared) that well-trained teachers creating engaging lessons and working in the community would eventually lead to better teaching and better teaching would lead to higher student achievement.

The professor designed a study where students in classes taught by interns were matched with students taught by regular DC teachers. With no districtwide standardized test available then, the professor used the Iowa Test of Basic Skills as the outcome measure of teaching in reading,
math, and other skills. About a year later, the professor called me up—I was then teaching history at a DC high school—and said he had the results. We met for coffee and he showed me what he had found.

In both elementary and secondary classrooms of interns and regular teachers, students in regular classrooms did marginally better than those in “intern” classrooms. While the percentile scores in both sets of classes were fairly low compared to the national average, I was still shocked. I had believed that the teacher-training program I had taught in and eventually directed was so strong that even in one year with “interns,” DC students would do better academically than in nonintern classrooms. I was wrong.

Although standardized testing was becoming common—the year is 1968—as a consequence of the Coleman Report (1966) and the federal Elementary and Secondary Education Act (1965) that required outcome measures to hold districts accountable for the federal dollars they received, I knew little then about the design and methodology the professor had used to evaluate student achievement.

Of course, now I realize that there were flaws in the evaluation design—it was not a random sample of students or interns; the test questions covered content and skills that students had not yet been taught in the DC curriculum; only one year was covered—still I was shaken by the results.

So I come to the end of my story and the puzzle of defining program “success,” one that a half-century later still bothers me. Here is an example of a pilot that initially appeared as a “success” in achieving its primary goals. The Cardozo Project in Urban Teaching baptized later by the DC Board of Education as the Urban Teacher Corps serving the entire district had the fragrance of success. Adding to the scent was the founding of the National Teacher Corps. Yet within a decade both teacher-training programs had disappeared.

And as an afterthought, in examining the underlying assumption that a majority white teacher-training project might have effects on prevailing faculty demographics, I found that African American teachers with a gradual smattering of Hispanics in the early 2000s continued to dominate the racial profile of faculties in these schools and across the district. Faculty demographics had changed minimally over time.
Finally, I discovered that students achieved less well in classes taught by interns than did students in regular classrooms. Even though raising student achievement was not one of the program’s goals, the results turned my assumptions inside out. Reform outcomes, I discovered, are seldom tidy.

Why this story? Puzzling then as it was to me, I still have not figured out a satisfying answer to the direct question of whether the Cardozo Project (and its progeny—the Urban Teacher Corps and the National Teacher Corps) was a “success.” Surely, it was not a “failure.” What lies between “success” and “failure” in a reform or when applied to a particular school or program?

Any answer to the question has to deal with what the concepts mean when applied to reform, organizations or individuals, who decides what they mean, the criteria used in making judgments, and whether these concepts have shifted in their meanings over time. Although determining “success” and “failure” remains puzzling to me, I have discovered that policy makers, administrators, practitioners, and political leaders also stumble over judging a particular innovation, program, school, classroom teacher, and individual students. And unraveling this enduring puzzle is one to which this book is devoted.

Looking Backward at Determining “Success” and “Failure” in Schooling

My experience in one reform-driven project does not mean that there have not been clear “successes” in improving schools. I describe briefly one educational institution that has weathered decades of political conflict and demonstrates resilience in its durability even today. I use the criteria of institutionalization of a reform and its longevity as measures of “success.” Political action, that is, gathering support from teachers, parents, voters, and taxpayers to adopt and sustain an organization is essential in determining “success” and “failure.” If a structure, program, or school process began as an innovation, was fully implemented, and then incorporated into the organization long enough to become taken for granted as is a drain in the kitchen sink or a sofa in the living room, the reform has been a “success.” Where these notions of “success” come from I take up in chapter 2.
Consider the mid-nineteenth century age-graded school imported from Prussia as an innovative reform to the then-dominant public school organization: the one-room schoolhouse. Horace Mann, Henry Barnard, and others were evangelists for age-graded Common Schools in New England and elsewhere. These reformers built political coalitions in various states that persuaded legislatures and town officials to fund these Common Schools. They succeeded in establishing such age-graded schools across New England, the mid-Atlantic states, and the Midwest prior to the Civil War.8

Since the late nineteenth century, the age-graded school (e.g., K–5, K–8, 6–8, 9–12) has become the mainstay of school organization in the twenty-first century. Today, most taxpayers, voters, and readers of this book have gone to kindergarten at age five, studied Egyptian mummies in the sixth grade, taken algebra in the eighth or ninth grade, and then left twelfth grade with a diploma.

If any school reform—in the sense of making fundamental changes in organization, curriculum, and instruction—can be considered an institutional “success,” it is the age-graded elementary and secondary school. In providing access to all children and youth, longevity as a reform, and global pervasiveness, the age-graded school is stellar.

But there is far more to the age-graded school. Its structures and rules such as a teacher for each classroom, daily schedules, homework, tests, report cards, and annual promotion are intrinsic to the organization. Students move sequentially through the school curriculum each month. If they pass tests and make “normal” progress, they advance to the next grade or can move from algebra to geometry. The age-graded school defines what is “normal” academic progress and behavior. These structures become what others and I call the “grammar of schooling.”

These structures and rules also instruct those in the organization about what adults and children are to do and how to behave inside classrooms, hallways, lunchrooms, and auditorium assemblies. In short, the “grammar of schooling” is the genetic code of the age-graded school.9

The age-graded school and its “grammar of schooling,” then, have been a robust “success.” Think about its longevity—the first age-graded structure of eight classrooms appeared in the Quincy Grammar School
in Boston, Massachusetts, in 1848. Within a half-century, it had begun to replace one-room schoolhouses in urban and rural schools.10

Or consider access. Between 1850 and 1913, over thirty million Europeans crossed the Atlantic and settled in the United States. The age-graded school has enrolled millions of students over the past century and a half, assimilated immigrants into Americans, sorting out achievers from nonachievers, and now graduates over 80 percent of those entering high school.11

Or ubiquity. The age-graded school exists in Europe, Asia, Africa, Latin America, and North America covering rural, urban, and suburban districts. What other school reform has been this “successful”?12

Habit and tradition play a part in the longevity of the age-graded school. The lack of recognizable alternatives that have been around sufficiently long to compete with the prevailing model is another. Sure, occasional reformers created nongraded schools and similar singletons, but they were outliers that disappeared after a few years.13

What is too often ignored in explaining the durability of the age-graded organization, however, are the widely shared social beliefs among parents and educators about what a “real” school is. After all, nearly all US adults—save for the tiny number who are home schooled—have attended both public and private age-graded schools. Addition, subtraction, and multiplication are taught in primary grades; the nation’s history is taught in the fifth, eighth, and eleventh grades. This is what a school is and does. American as apple pie and the Thanksgiving holiday.

For example, when charter school applicants propose a brand-new innovative school, the chances of receiving official approval and parental acceptance increase if it is a familiar age-graded one, not one where most teachers team teach and groups of multi-age children learn together.14

External pressures also constrict reformers’ maneuverability in trying other organizational forms. State-mandated grade-by-grade curriculum standards—college entrance requirements calling for which academic subjects have to be taken and passed are located in the ninth to twelfth grades, and the federal Every Student Succeeds Act determining what grades elementary and secondary school will be tested—are all married to this taken-for-granted institution.
The unintended (and ironic) consequence of frequent and earnest calls for radical change in instruction through nontraditional teachers and administrators, charter schools, nifty reading and math programs, digital devices for kindergartners, and other reforms assumes that such innovations will occur within the traditional school organization, thus inadvertently preserving the age-graded school and freezing classroom patterns that so many reformers and entrepreneurs want to alter.

Beyond the age-graded school, there have been other fundamental and incremental changes that have, intentionally or not, sustained the structure and culture of this organization. The striking emergence of the junior high school and then the comprehensive high school in the 1920s, offering students an array of academic and vocational courses leading to graduation, remains a mainstay in US secondary schools in 2020. These new age-graded organizations appealed to families wanting their sons and daughters to have a high school diploma, at that time a prized piece of parchment.

Cementing that high school structure in grades nine through twelve has been the Carnegie unit—student contact of 120 hours in a class over a school year of at least twenty-four weeks—installed as another innovation in the early twentieth century. It has been used as a basis for students graduating high school and continues into the twenty-first century.¹⁵

These triumphs of reform came about because every policy aimed at solving a district or school problem contains three essential elements that demand attention prior to making judgments about overall “success” and “failure”: the process of adopting a policy, the programs that enact the policy, and the politics necessary to put the policy into practice. Thus, there can be a lasting institutional “success” in all three realms—see above examples—but more often than not, there are temporary and partial “successes” leading to mixed judgments; for example, the policy was a “success” in getting adopted but “failed” in securing sufficient political support from teachers for classroom implementation (e.g., school vouchers, open space schools). Noting these three strands of organizational policy recognizes that “success” and “failure” are not binary judgments. There may be early victories and later defeats in one
realm and not another. Puzzling contradictions and ambiguities that accompany so many judgments about policies as “winners” or “losers” become clearer and richer in acknowledging these political dimensions of reform-driven policies.

Recognizing these often-overlooked facets of policy helped me make sense of my experiences in a federally funded program to train teachers on school sites a half-century ago. Such analyses of policy also help unravel ambiguities and contradictions in assessing “success” and “failure” in noneducational domains such as starting and sustaining business and providing health care.

This brings me to the questions that drive this study of organizational “success” and “failure” in school reform and judging the quality of public schools in the nation.

1. How have “success” and “failure” been defined and applied to schools past and present?
2. From where do these ideas of “success” and “failure” come?
3. How were these ideas transmitted to Americans then and now?
4. Who decides (and how) whether schools “succeed” and “fail”?
5. What does institutional “success” and “failure” look like in two contemporary schools?
6. So what?

My journey from the Cardozo Project in Urban Teaching in the mid-1960s to now has been both meandering and long. Whether that project, one that led to a career as an administrator and then professor, was a “success” or “failure” is a question that has dogged me for over a half-century. As a teacher and superintendent, I thought I knew the answer, but as a researcher looking back to what I did and what I have learned from investigating reforms and similar projects, I now understand how notions of “success” and “failure” I took for granted, I had never questioned. Reflecting on my experiences and research over decades, I remain uncertain and ambivalent about these common terms to describe schools and life itself. Thus, this book.

The following chapters answer the above questions. The answers I constructed give me a hand-hold—but not a full grip—on figuring out
what these common terms mean in judging school reform, applying them to school innovations, programs, individual teachers, and students. And, I have discovered that reconsidering key events in my life has brought me full circle to where I began in Washington, DC, where I became hooked on determining what “success” and “failure” mean in American public schools.