Student Progress to Graduation in New York City High Schools

PART II

Student Achievement as *Stock* and *Flow*: Reimagining Early Warning Systems for At-Risk Students

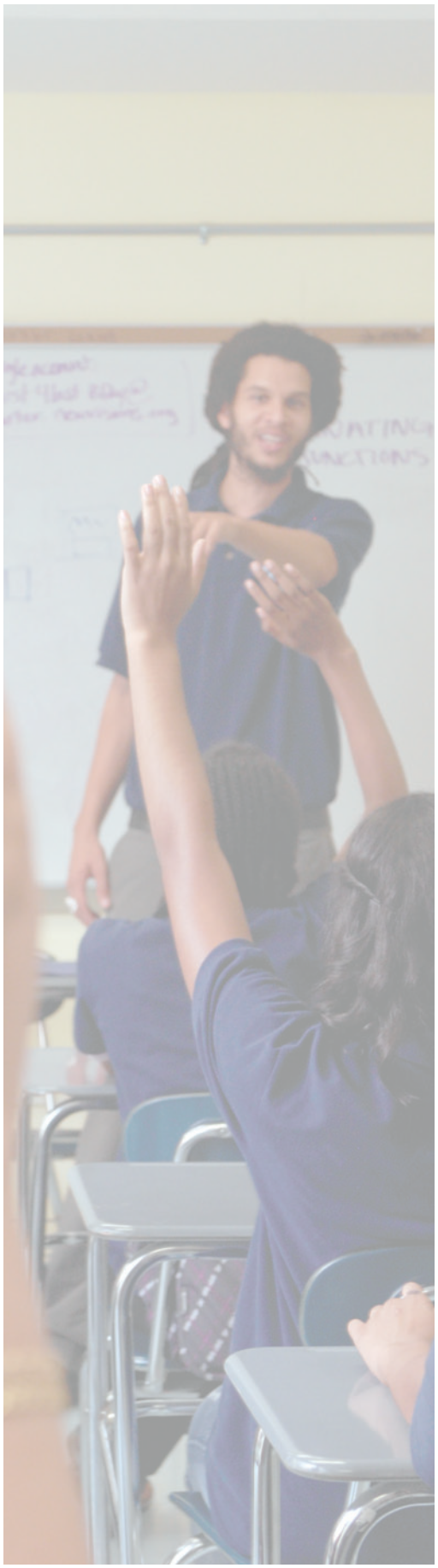
Authors:
Susan Fairchild
Gerard Carrino
Brad Gunton
Chris Soderquist
Andrew Hsiao
Beverly Donohue
Timothy Farrell

Data Visualizations:
Sarah Slobin and Andrew Garcia Phillips

Acknowledgments:
Amanda Smith and Russell West

NEW VISIONS
FOR PUBLIC SCHOOLS

Table of Contents



This report is the second in a series documenting the design, analysis, revision, and implementation of a new metric for measuring student progress to graduation and college readiness in New York City high schools.

Recommended citation: Fairchild, S., Carrino, G., Gunton, B., Soderquist, C., Hsiao, A., Donohue, B., Farrell, T. (2012). *Student Progress to Graduation in New York City High Schools. Part II: Student Achievement as Stock and Flow: Reimagining Early Warning Systems for At-Risk Students*. New York, NY: New Visions for Public Schools.

"No statement is complete or comprehensible in itself. … Its speaker must stand by it: must believe it, be accountable for it, be willing to act on it."

—Wendell Berry, "Standing by Words"

Poet and activist Wendell Berry may seem an unusual starting point for a report on identifying at-risk students in New York City public schools. But Berry's concern for the "accountability of language" strikes us as eminently appropriate when discussing a policy environment that itself stresses accountability of principals, teachers, and schools. Central to Berry's argument in his essay "Standing by Words" is that the language of the technician or specialist plays a degenerative role in our culture. When we resort to sterile, specialist jargon, we mask real crises, passing them off as mere technical events or problems. Words can either demonstrate ownership and a connection to unfolding events or they can isolate and separate us from those events. He argues that politic language obscures and disconnects us from crises and those affected by them.

In stark contrast to the sterile, specialist language against which Berry warns, the language of "A Nation at Risk: The Imperative for Education Reform," released in 1983 by the Commission on Excellence in Education, is immediate and alarming in tone. The report highlighted the increasing rates of illiteracy, declining quality of teacher preparation, and increasing number of high school graduates inadequately prepared for the real world. "A Nation at Risk" is remarkable not only because it surfaced grim, new trends in the nation's educational health, but also because it sounded an alarm. The report's epic language and imagery speaks to core values: the "educational foundations of our society are presently being eroded by a rising tide of mediocrity"; rising levels of mediocrity "threaten our very future as a Nation and a people" — specifically, our "prosperity, security, and civility." The commission goes so far as to liken educational mediocrity to "unilateral educational disarmament," positing that if "an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war."[1]

The Commission on Excellence in Education used bold, passionate language designed to invoke urgency, danger, and ownership (e.g., "our society," "our future as a Nation and a people") on a national scale. Key concepts such as "early warning," "at risk," and "indicators of risk" emerged from this report that, 30 years later, continue to frame the national discourse on education reform and influence the types of data we examine from our student information systems. Since the publication of "A Nation at Risk," important and ongoing efforts to identify early indicators of risk have emerged;[2] transitions that are key to success along the educational pipeline have been

1 The Commission on Excellence in Education, 1983, p. 9.
2 E.g., Allensworth & Easton, 2005; Roderick, 1994; Roderick, Nagaoka, & Coca, 2009.

highlighted;[3] accountability structures have been designed that reward (or punish) schools for their at-risk students' academic achievement;[4] and on-track metrics and data systems have proliferated and have become more sophisticated.[5]

Along with many of our colleagues across the nation who have internalized this message, New Visions for Public Schools has leveraged student-level data to help schools identify at-risk students, designed metrics to capture student progress toward graduation, developed data tools and reports that visualize student progress at different levels of aggregation for different audiences, and implemented real-time data systems for educators.

Central to New Visions' early warning system is the point-in-time index — the four-color Progress to Graduation Metric — made possible through the greater availability of real-time student data such as graduation rates, attendance rates, credit accumulation, Regents' passage rates, student assessments, and other key indicators of student performance. Student performance benchmarks like New Visions' Progress to Graduation Metric serve as one of a school's tools for identifying at-risk students and help to inform early-warning interventions. These same student performance benchmarks, when aggregated at the school level, become one of a district's tools for identifying at-risk schools and informing interventions.

Early warning systems, like New Visions' Progress to Graduation Metric, while providing a solid basis for characterizing a student, a classroom, a department, a school, or a district at a moment in time, are also limited. The primary intent of this paper is to present a new framework that will guide the next phase of New Visions' early warning data work. These are the goals of this paper:

- To illustrate how a "systems thinking" approach adds dimension and depth to our understanding of student performance, allowing us to reimagine our data systems.

- To introduce the concept of "structural volatility" — and the new data tool that begins to capture this phenomenon.

- To suggest how the terms "at risk" and "early warning," despite the passion and urgency with which they were first introduced in "A Nation at Risk," have become the politic, specialist language Berry cautions against; and how, by shifting the framework through which we understand risk, we restore and reclaim them.

3 E.g., Balfanz, Herzog, & MacIver, 2007; Allensworth & Easton, 2005; Heppen & Therriault, 2008; Roderick, 1994; Kieffer & Marinell, 2012.

4 See: http://schools.nyc.gov/Accountability/tools/report/default.htm#Methods; Bolon, 2001; Darling-Hammond, 2003; Kane & Staiger, 2003.

5 E.g., Fairchild, Gunton, Donohue, Berry, Genn, & Knevals, 2011; Tucker, 2010; Neild, Balfanz, & Herzog, 2007; Halverson, Grigg, Pritchett, & Thomas, 2005.

Reimagining Our Early Warning Data Systems Through the Lens of Systems Thinking

New Visions' Progress to Graduation Metric is an example of first-order analytics,[1] the relatively blunt, yet fairly effective instrument we use to characterize different levels of achievement within and across our network of 75 schools serving 35,363 students. This four-category Progress to Graduation Metric[2] compares a student's point-in-time performance to New Visions' standard deemed "on progress toward graduation."

Using the four-color scheme as a starting point, there are at least two approaches to refining our metric: extend the current dimension and/or add a new dimension informed by systems thinking.[3] Extending the current dimension of our Progress to Graduation Metric is a simple matter of making it more discrete, e.g., turning the four-category system into a six-category system. The second approach adds other dimensions (e.g., volatility and direction) without collecting more data. Incorporating "volatility" and "direction" within the existing dimension of point-in-time performance compared to the standard refines the system in ways that more thoroughly characterize and visualize student progress.

STUDENT ACHIEVEMENT AS STOCK

At the student level, first-order analytics indicate moment-in-time categorical performance level as progress to graduation. At New Visions, we characterize this on a scale with four major categories:

1. On Track to College Readiness (Blue)
2. On Track to Graduation (Green)
3. Almost on Track (Yellow)
4. Off Track (Red)

Table 1. Summary of New Visions' Progress to Graduation Metric (see Appendix 1)

On track to college readiness	Meet "On Track to Graduation" requirements, plus 75s on the Regents for Math and ELA, and 4 additional Regents exams.
On track to graduation	Gain 1 credit per semester in each core subject; Gain 11 credits per year; Pass 1 Regents exam by end of freshman year, 3 by end of sophomore year, and 5 by end of junior year.
Almost on track	Gain credits in each subject, but maybe fewer than 1 per semester; Gain at least 8 credits per year; Begin passing Regents exams by junior year.
Off track	Fail to gain credits in a particular subject; Gain very few credits overall; Pass no Regents exams by end of junior year.

1 First order analytics describe the basic attributes of the spatial database (e.g., color, category, mean). Data attributes (e.g., standard deviation, volatility, direction) that describe patterns of movement and thereby expand the spatial database are second-order analytics.

2 See Fairchild, et al., 2011.

3 Though we are pursuing the addition of more discrete categories to the Progress to Graduation Metric, this paper focuses only on the addition of volatility and direction.

Using this conventional paradigm, indicators of risk are generally benchmarked by a student's end-of-semester category, as are many interventions. (Similarly, a classroom, a department, a school, or the network can be characterized using aggregations of these categories.) A student's moment-in-time achievement, represented by one of these four colors, is his or her "stock," which is an accumulation of achievement that has built up over time.[4] A school's stock is the proportion of students at any one time in each of the four categories. But as static, moment-in-time measures, stocks are insufficient to characterize context, to contemplate history, to capture feedback loops, or to visualize flow.

Take, for example, the comprehensive but two-dimensional table of progress to graduation categories by semester aggregated across a school (Table 2).

Table 2. Student (n=100) progress to graduation across 8 semesters of high school

	Sem 1	2	3	4	5	6	7	Diploma	
On track to college readiness	10	19	23	16	10	11	10	8	Advanced Regents
On track to graduation	54	31	29	13	54	29	38	58	Regents
Almost on track	30	40	32	52	30	42	35	19	Local
Off track	6	10	16	19	6	18	17	12	Still enrolled, IEP, GED
								3	Dropped out

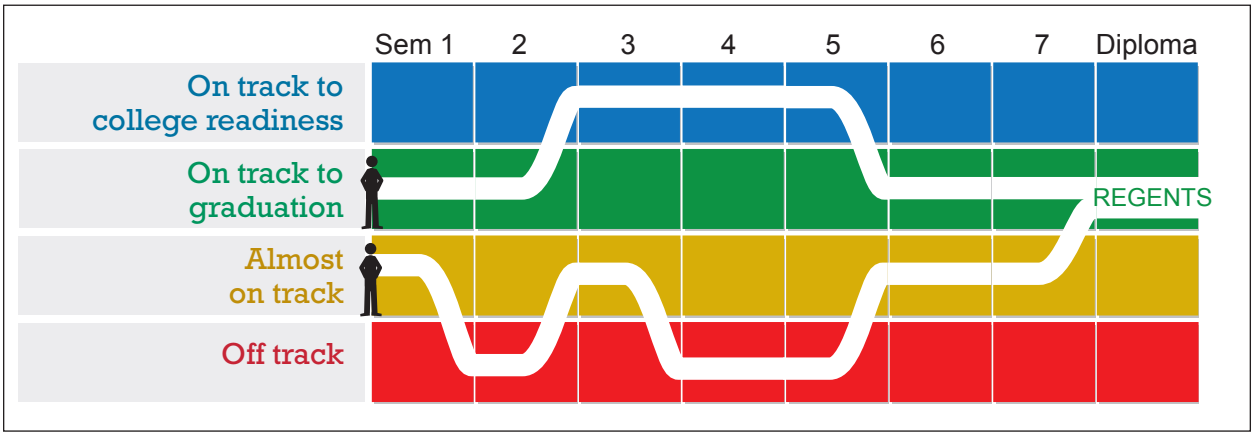
Much can be ascertained and interpreted from this table of stocks, such as the school's success in graduating college-ready seniors, or, with a bit more interpretation, the steady decline over time in the number of students who maintain college readiness over the high school years. But this two-dimensional matrix falls short of informing our understanding of history, trajectory, feedback, and flow.

4 Meadows, 2008.

STUDENT ACHIEVEMENT AS FLOW

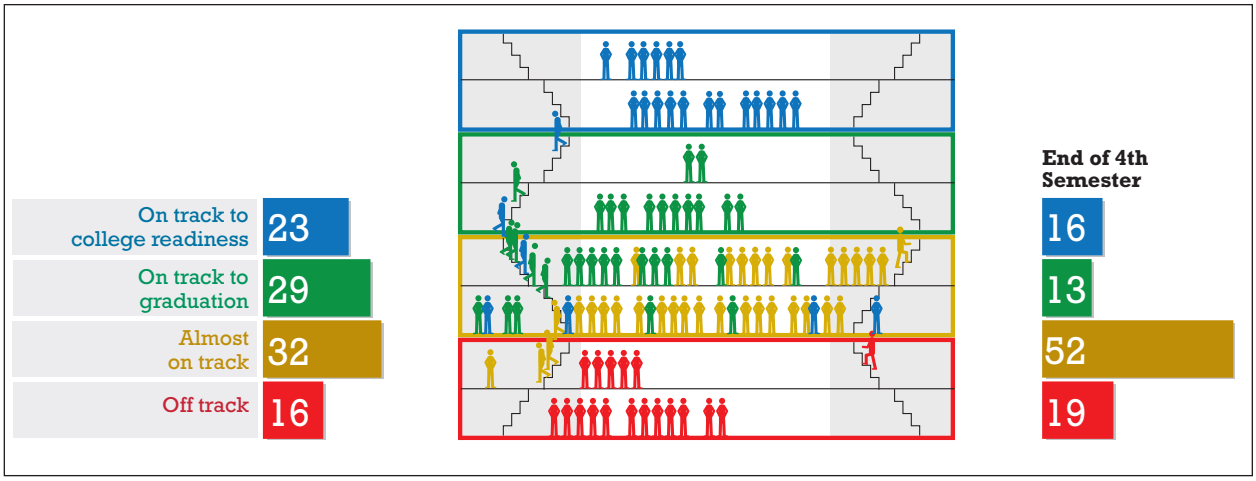
Systems thinking, and the analytics revealed by it, begins to address these deficiencies by adding data attributes that characterize movement. Intuitively, it is obvious that two students who end in the same place but have different progress pathways are not identical (see Figure 1).

Figure 1. Student pathways to a Regents Diploma



If stocks represent a specific moment such as student progress at the end of the third semester and those stocks change by the end of the fourth semester, then "flows" represent the dynamic quality or movement of student progress during semesters.

Figure 2. Stock and flow between 3rd and 4th semesters



In Figure 2, seven blue students, through lower performance during the fourth semester, flow into the green or yellow categories by the end of that semester. At the school level, flow is the "filling or draining"[5] process that changes the numbers of students within each of the progress to graduation categories. This inherent dynamism between stock and flow can be characterized and quantified in two ways: volatility and direction. Volatility is the amount of

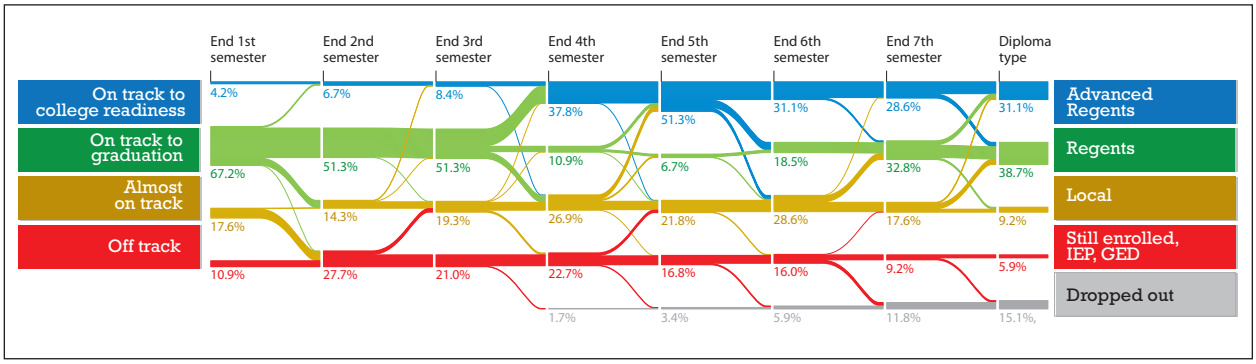
5 Meadows, 2008, p. 18.

variation or change in a student's progress to graduation status over eight semesters (i.e., how many times a student changes category).[6] Direction depicts whether movement is toward or away from on-target performance, as indicated by progress to graduation category (color). From one semester to the next, a student can move to a higher category, move to a lower category, or stay in the same category. It is important to remember, though, that "stability" (i.e., the lack of volatility, or staying in the same category from one semester to the next) still represents an accumulation of stock because a later semester implies a greater level of attainment than an earlier semester, even as color-coded performance is stable.

STUDENT ACHIEVEMENT AS STOCK AND FLOW

Putting it all together, stocks and flows produce an important perspective for schools. Figure 3 illustrates how New Visions' metric informs the stock-flow-stock patterns. This stock-flow-stock pattern shows the continuity and flow of student progress over time.

Figure 3. Progress to graduation key and school-level progress to graduation maps across 8 semesters



In addition to seeing the patterns of student progress in a school, when compared across school years, these progress to graduation maps may also reveal potential changes in school strategy and/or the effects of an intervention. By comparing different cohorts (i.e., by comparing two or more maps next to each other), schools can begin to identify where interventions are needed for the subsequent cohort. Thus, these graphics demonstrate that important feedback loops are driving student and school performance. Rather than the traditional x → y (causes) approach, systems thinking is "x → y, which in turn → x" — in this case, the feedback loops that represent the interdependencies between students and schools.

6 Volatility (V) is a cardinal variable that counts the number of category transitions between semesters. Any transition, no matter how distant the categories are from one another (e.g., red to yellow versus red to green), counts as one transition (the magnitude and direction of the transition are captured in the direction variable). Diploma type is a proxy for 8th semester, where Advanced Regents Diploma is blue, Regents Diploma is Green, Local Diploma is Yellow, and GED and Dropout are red. Volatility can range from 0 to 7, representing up to seven changes within the eight semesters of high school.

Structural Volatility: A New Concept in School Performance

Just as a school can shape its students, the students can shape their school. Key school attributes — strong leadership; professional capacity; rigorous, ambitious instruction; student-centered climate; and parent and community engagement* — exist not independent of students but often in response to them. We see not only that students can and do flow between higher and lower performance levels, but also that schools, through the structures listed above, can induce student volatility, whether purposely or inadvertently. The most obvious example of purposely induced volatility occurs when schools intervene to improve failing students' performance (e.g., a school may adopt intensive "credit recovery" for seniors who are at risk for failing to graduate on time). Less obvious, though, is the inadvertently induced volatility that may occur when those same schools — while focusing on interventions that help improve failing students — draw focus from students less obviously at risk. Likewise, student performance can induce structural changes in schools (e.g., a school may adopt a new math curriculum in response to poor performance on the math Regents exam).

Simply stated, school structures shape student performance, which shapes school structures. Structural volatility is a feedback loop representing the ways in which students and schools respond to each other. In other words, structural volatility reflects the way a school runs itself.[7]

Figures 4–6 capture the movement of four cohorts[8] of students (cohorts 2008, 2009, 2010, and 2011) across eight academic semesters (or four school years) in a single high school. These graphical illustrations demonstrate the highly variable, and even dramatic, patterns of student progress within a school. In

* New Visions has developed an overarching system to track, analyze, and refine our school-level interventions. Extensive studies and longitudinal research by the Consortium on Chicago School Research have shown that schools must focus on five key elements to increase student achievement: strong leadership, distributed professional capacity, rigorous instruction, student-centered learning, and parent and community engagement (Bryk, Sebring, Allensworth, Luppescu, & Easton, 2010). We believe that these five components form a comprehensive model for organizing, prioritizing, studying, learning from, and ultimately scaling best practices at the school level. Over the past year, New Visions has created an organizational Learning Framework based on these elements. We have worked extensively with our coaching and leadership development staff to define the principles (the key drivers of student improvement); categories (the systems and areas of focus within each principle); strategies (high-leverage skillful moves to ensure and exemplify high-functioning systems of support); and action items (articulated components of or steps toward achieving a strategy) that together comprise the specific steps a school leader or community must take to ensure that they are adequately addressing each of the five essentials to increased student achievement.

7 See Meadows, 2008.

8 The 2008 cohort for graduation rate accountability consists of all students, regardless of their current grade level, whose "First Date of Entry into Grade 9" (anywhere) was during the 2004–2005 school year (or four years prior, if a different graduating year).

each semester and for each cohort, we see high-performing students stumble and low-performing students rise. At the same time, even greater numbers of students remain in the same category, for better or worse. And perhaps even more interesting, when we look across cohorts, we see students with similar academic profiles (e.g., state test scores) performing differently. What accounts for these differences among students with similar levels of performance?

EXAMPLE OF STRUCTURAL VOLATILITY: PROGRESS TO GRADUATION

Figures 4–6 are progress to graduation maps for a cohort, or graduating class, of students. Each student, and, in the aggregate, each cohort, starts with an initial stock of high school readiness[9] (see "Before" in the first column of the graphics below). As they move through eight semesters of high school, ending with diploma earned, students flow through varying strata of progress to graduation. Each of the graphics below is not a point-in-time; rather, it is a longitudinal depiction of a cohort. It is critical to remember, however, that a school's reality is not a single cohort over four years, but four cohorts (freshmen, sophomores, juniors, seniors) simultaneously progressing through time. When we compare each of these graphics, this high school's unique student performance trajectory unfolds. We begin with Cohort 2008[10].

Cohorts 2008 and 2009

The progress toward graduation of the class of 2008, over the course of eight semesters or four years, is shown in Figure 4.[11] The map shows that 69.3 percent of students entered high school having achieved a level 2 proficiency on the eighth grade math and English Language Arts (ELA) exams.[12] By the end of the first semester in high school, approximately 48 percent of cohort 2008 were on track to graduate (green). But it is the dramatic movement of students flowing out of higher-performance categories into lower categories that characterizes the end of semester 1 to the end of semester 2. That is, Figure 4 shows a high rate of students draining from the on track to graduate (green) stock into the almost on track (yellow) stock. During that same time period, the off track (red) stock fills with almost on track (yellow) students.

By the end of second semester, the proportion of students categorized as on track to graduate (green) begins to erode. At this moment in time, 86 percent of cohort 2008 is almost on track or off track. This rather incredible yellow/red

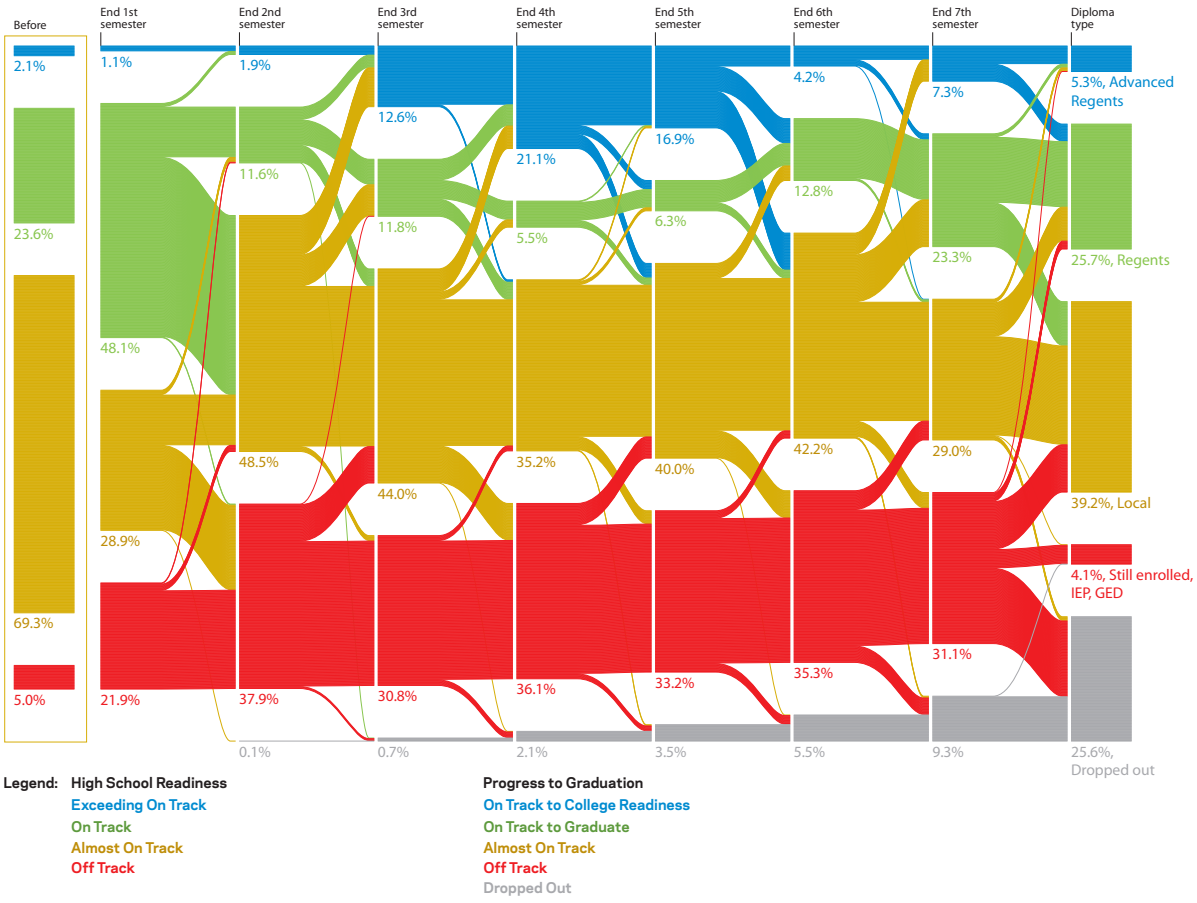
9 Students' high school readiness is calculated based on their ELA and math decile scores assigned to them by the city. The decile scores are determined by students' performance on the ELA and math eighth grade assessments when available; for students without scores, the deciles are based on the students' demographics (specifically, DOE weights are based on Black/Hispanic, free/reduced price lunch, special education status, English language learner status, and students with interrupted formal education). **On Track** students are generally already meeting expectations as they proceed into high school, with **Exceeding on Track** students well above those expectations. On the other hand, **Almost on Track** students are somewhat below expectations in one or both subjects, and **Off Track** students are significantly below.

10 New Visions became a Partnership School Organization (PSO) in 2007. Cohort 2008 was the first cohort from whom we have extensive student-level data.

11 The school-level progress to graduation maps are remarkably similar for cohorts 2008 and 2009. Therefore we only describe patterns of progress for cohort 2008.

12 New Visions four-point High School Readiness scale is similar to the four-point scale used to grade middle school tests in ELA and math. A score between 3 and 3.9 (green) is considered Proficient, and scores of 4 and above are Above Proficient. The 2–2.9 range is Below Proficient, and 1–1.9 is Well Below Proficient. See Appendix 2 for a detailed explanation of the calculation.

Figure 4. Cohort 2008 school-level progress to graduation map



To be considered on track to graduate (green) at the end of semester 1, a student needs 5 total credits, 1 in each core subject area. To be considered almost on track (yellow), a student needs 3 total credits. At the end of semester 2, a student must earn 11 total credits (2 in each core subject area) plus 1 Regents passed to be considered on track to graduate (green). A student must earn 8 total credits (1 in each core subject area) to be considered almost on track (yellow). Students fall from green to yellow because some first semester green students successfully earn credits but struggle with the Regents exams. Students will fall from yellow to red because earning 3 credits first semester is relatively easy, but earning a total of 8 credits including 1 in each subject area is considerably harder. In both cases, maintaining a progress to graduation category from first to second semester requires more work than students demonstrated in the first semester.

tide of students will flow from one semester into the next until the end of the sixth semester, when it finally begins to recede.[13]

In the third, fourth, and fifth semesters, we see the filling of the on track to college readiness stock (blue). But by seventh semester, the on track to college readiness gains have diminished to levels seen at the end of first semester. The drop occurs mostly at the end of the sixth semester, when students need to have passed both ELA and math Regents exams at high levels to be considered on track to college readiness. In this school, both recovery (lower-performing students becoming higher-performing students) and dropping out escalate during the sixth, seventh, and eighth semesters.

13 It is likely that 86 percent of cohort 2008 is almost on track (yellow) or off track (red) because schools tended to delay administration of the Living Environment Regents exam until the sophomore year. Many students did not attempt a Regents exam freshman year, hence the large swath of almost on track (yellow) students. We believe this yellow/red tide begins to recede by the end of sixth semester partly due to credit recovery as well as students' retaking the Regents exams they previously failed.

Cohort 2010

Two years later, cohort 2010 students (Figure 5) are similar to cohorts 2008 and 2009 students with respect to incoming eighth grade math and ELA state scores. But by the end of semester 1, this school has substantially more students on track to graduate (green) in cohort 2010 (61 percent) than in cohorts 2008 (48 percent) and 2009 (43 percent). Again, because the school population is relatively stable from one year to the next, this increase in the percentage of on track to graduate (green) students suggests that the school has implemented a new strategy targeting freshmen during their first semester in high school. But by the end of their freshman year (end of second semester), a large percentage of those green students flow into yellow — with percentages similar to that of the previous year. The student momentum from one semester to the next was not sustained in the on track to graduate (green) category. Still, the percent of on track to college readiness (blue) and on track to graduate (green) is higher than in previous cohorts, while the percent off track (red) is lower. This finding seems consistent with a freshman-year intervention that started or intensified after the class of 2008. The number of almost on track students who flow into the off track stock between the end of semester 1 and the end of semester 2 has been reduced compared to previous years. The patterns across subsequent semesters are similar to those of the previous two years — *with the important exception that each year, fewer students are flowing into lower-performance categories.*

Cohort 2011

From 2008 to 2011, the eighth grade math and ELA state scores of incoming freshmen are comparable; though the student population appears to be consistent on this measure across time, new patterns of student performance emerge. By the end of the second semester, fewer on track to graduate (green) cohort 2011 students (Figure 6)[14] flow into almost on track, more are stable greens from one semester to the next, and more flow into on track to college readiness. In fact, the on track to college readiness (blue) category swells to impressive percentages between the fourth and fifth semesters, but then drains down to end of semester 1 levels. Over time, this high school maintains a more robust on track to graduate (green) core. By the end of high school, more students graduate with a Regents Diploma. The pattern of students falling into the off track category has also changed from one cohort to the next. The larger percentages of students flowing from the almost on track to off track stock in early semesters in previous cohorts are markedly reduced in cohort 2011.

Because each graph plots the stocks and flows of a single cohort over four years, it may reflect changes within the school over time; but any such change in the school is obscured by the natural development and growth of the cohort itself, as the students progress to graduation. Comparing two or more of these figures across years, however, can begin to give us important insights regarding

14 Figure 6 displaying cohort 2011 data includes more missing data than cohorts 2008, 2009, and 2010. The missing data are disproportionately off track students, thereby inflating the percentage of students in the higher categories. The 2011 cohort data file has not yet been finalized. New Visions works with the Department of Education to acquire a finalized cohort of students. However, Figure 6 for cohort 2011 reveals new student performance trends that persist regardless of missing data.

Figure 5. Cohort 2010 school-level progress to graduation map

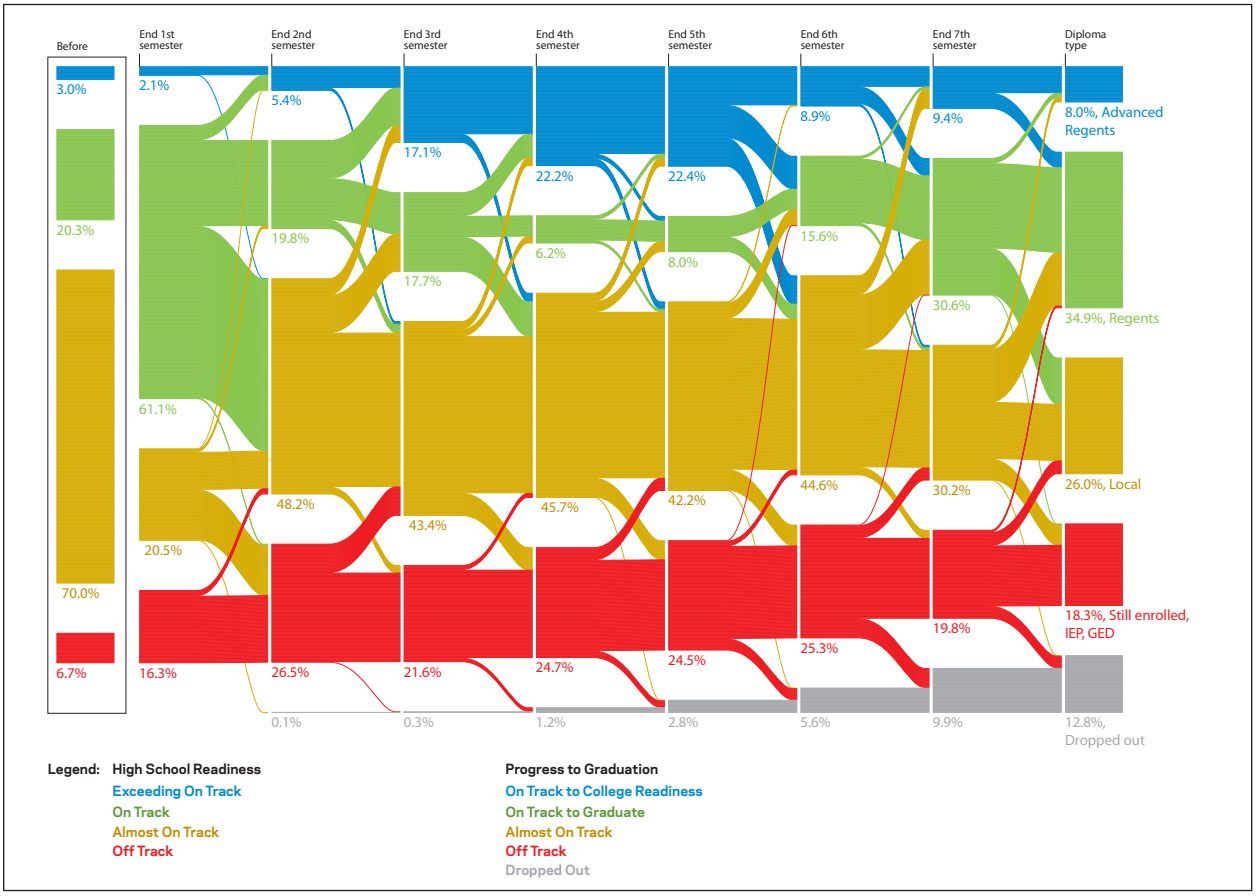
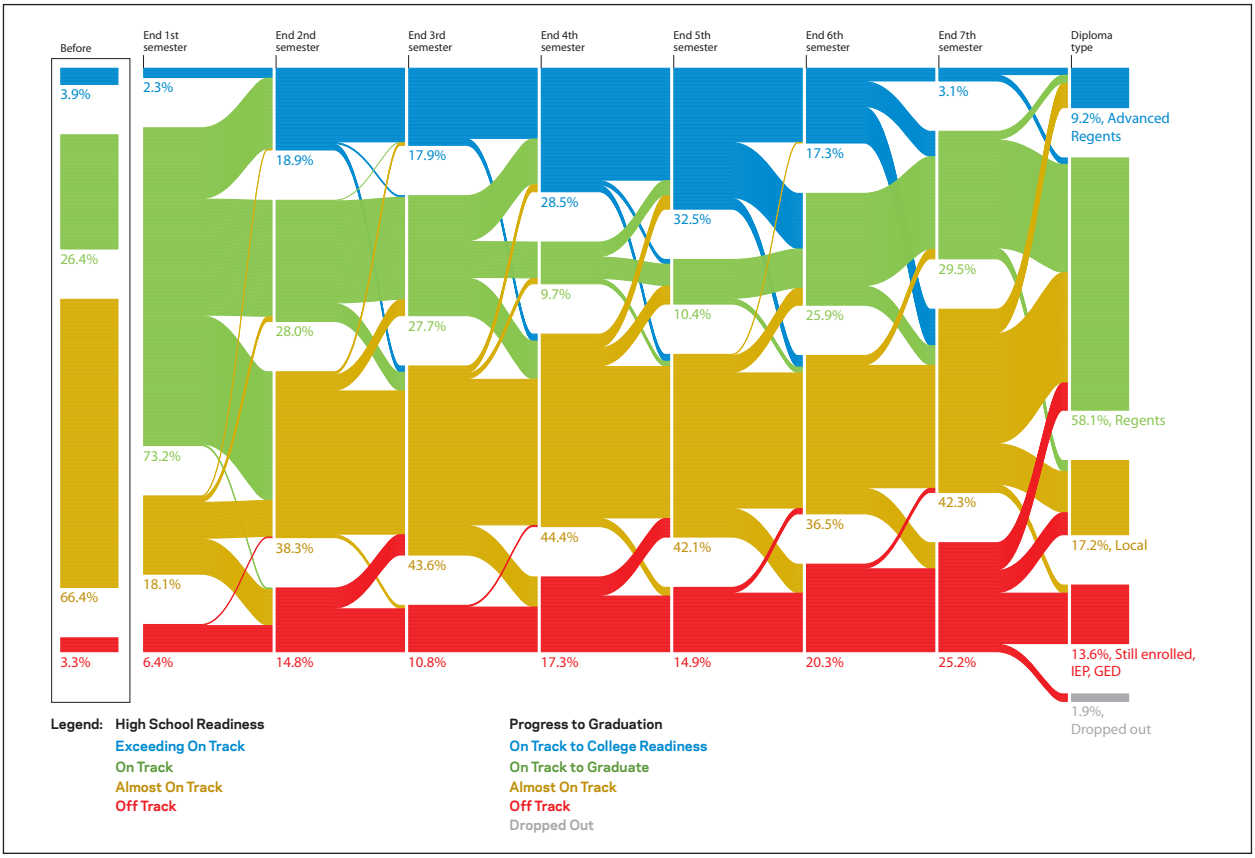


Figure 6. Cohort 2011 school-level progress to graduation map



structural volatility, e.g., how one cohort's influence induces change in the school to benefit the next cohort, or how one cohort's influence induces change in the school to the detriment of the other cohorts of students. These insights can become the foundation for a practical planning/reflection tool for schools. Schools get a visual sense of how major interventions have potentially shaped student performance.

Schools must ask themselves: Does the structural volatility captured in these graphs reflect proactive or reactive decisions and actions taking place in schools? These graphs will not answer this question definitively, but they support and advance the conversation. The answer to this question is paramount and is directly linked to a school's conceptualization of risk.

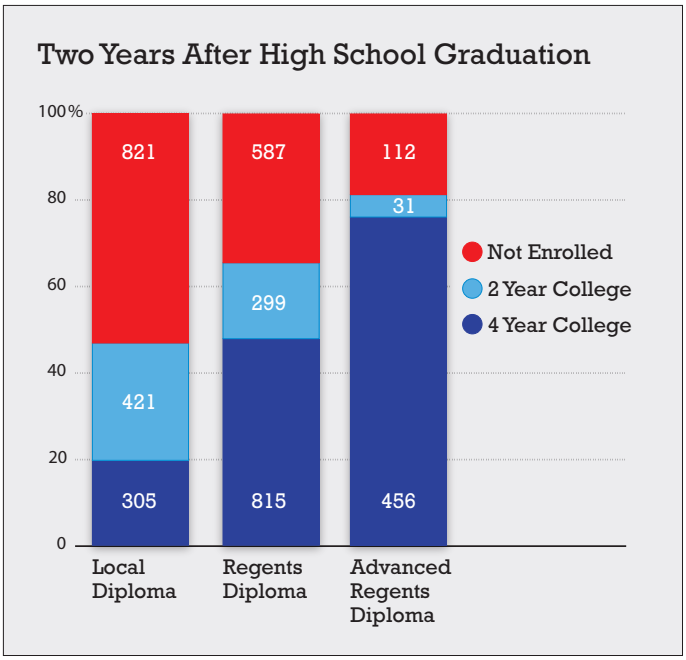
Risk: An Emergent Property of the School

Wendell Berry's call that we stand by our words requires that we first understand the phenomenon of risk. Individuals who chronically react to unfolding events in schools mean something different when they talk about "at risk" than individuals who are proactively anticipating events and looking into the root cause. According to Sterman, "complexity hinders our ability to discover the delayed and distal impacts of interventions, generating unintended 'side effects.'"[15] In other words, our linear-thinking minds are no match for the complexity that presents in our schools, and this is reflected in the way in which we use words like "at-risk" and "early warning." Traditionally, we use "at-risk" to describe a student at a moment when the symptom has presented. This focus on a student at a particular moment distracts us from perceiving the structures that are systematically producing risk. In systems thinking, this phenomenon where we aim to "fix" the immediate problems that present rather than focusing on the root cause is called "shifting the burden."

SHIFTING THE BURDEN

For example, even though increasing numbers of students in New York City are graduating with a Regents Diploma, not all Regents Diplomas are created equal. Beginning in 2009, New York State's Board of Regents began phasing out the less rigorous Local Diploma.[16] By 2012, all general education New York City students must meet the requirements for the Regents Diploma if they are to graduate from high school. When looking at the class of 2008 in 34 New Visions PSO schools with a graduating cohort, it is no surprise that those students graduating with a more advanced diploma had more successful college outcomes (Figure 7). Approximately 76 percent of students who graduated with an Advanced Regents Diploma were enrolled in a four-year college two years after graduating from high school, as compared to 49 percent of students who graduated with a Regents Diploma and 23 percent who graduated with a Local Diploma.

Figure 7. Persistence in college two years after high school graduation by diploma type for cohort 2008 students



15 Sterman, 2006, p. 505.

16 Prior to 2009, requirements for earning a Local Diploma included passing the five required Regents exams at 55 or higher and/or the six Regents Competency Tests with a pass rate and accumulating 44 credits. For cohorts 2012 and beyond, special education students are the only students eligible to receive a Local Diploma.

These diploma groups, however, are not homogeneous (see Figure 1, p. 7). When we look more closely at the students who have earned a Regents Diploma, we see how different student pathways shape postsecondary outcomes. For instance, in Figure 8, approximately 81 percent (n=167) of students in cohort 2008 who graduated with a Regents Diploma and who were on track to college readiness (blue) in their seventh semester persist in college two years after graduating from high school. Conversely, approximately 46 percent (n=100) of students who were off track (red) in their seventh semester and who earned a Regents Diploma persist in college.

Looking further back into students' high school progress to graduation history, we see that students who have an average progress to graduation score[17] of 3 (on track to graduation) or higher[18] for the fifth, sixth, and seventh semesters have stronger persistence rates in college (see Figure 9).

Figure 8. Persistence in college two years after high school graduation by seventh semester progress to graduation category for cohort 2008 students with a Regents Diploma

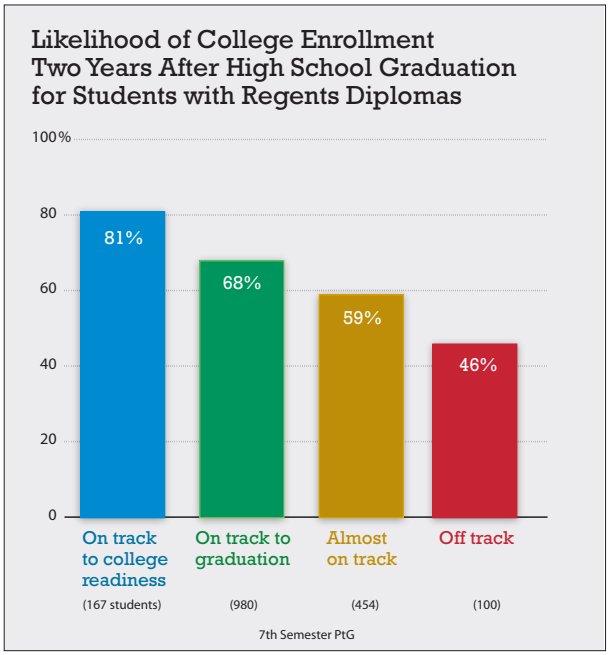
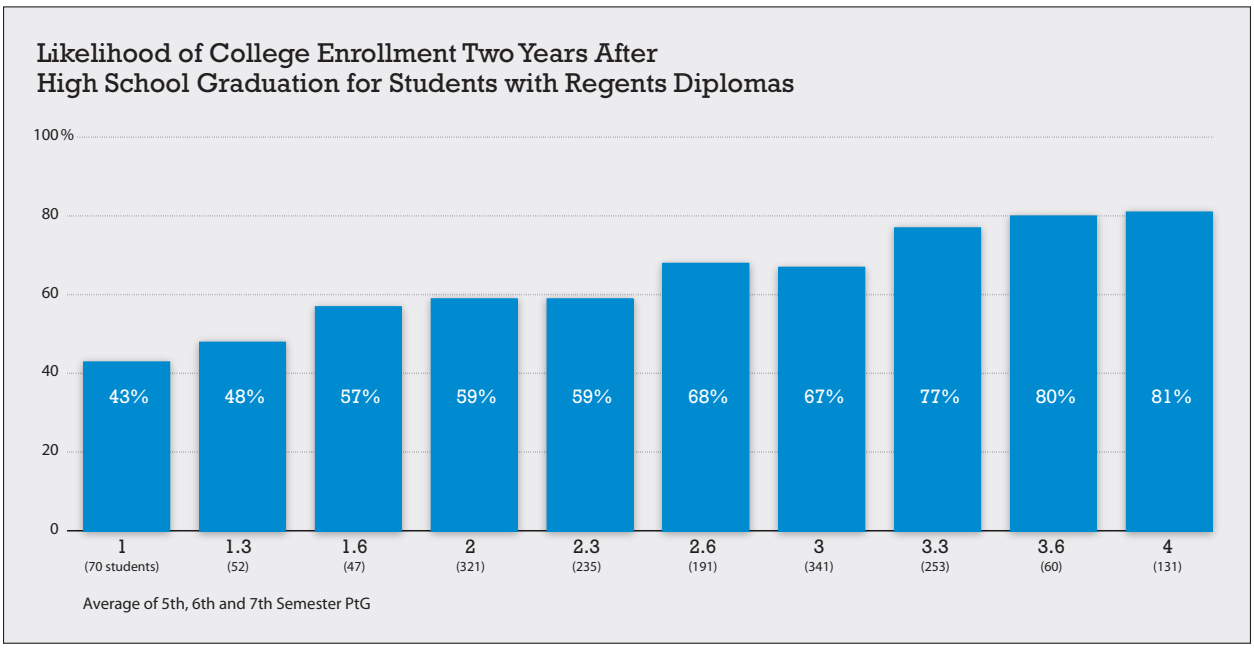


Figure 9. Persistence in college two years after high school graduation by average fifth, sixth, and seventh semester progress to graduation categories for cohort 2008 students with a Regents Diploma



17 Average progress to graduation is the average of x semesters where blue = 4, green = 3, yellow = 2, red = 1.

18 Students with an average progress to graduation score of 3 or more in the fifth, sixth, and seventh semesters are some combination of on track to college readiness (blue) or on track to graduate (green) during those three semesters.

The data tell us that, while schools may succeed at catching students just before they drop out or just before they fail to graduate, educators have not sufficiently addressed the skill and content vulnerabilities such that those students will succeed in college. In other words, the notion of student "risk" is shifted to postsecondary institutions. Policy makers' focus on diploma type, while no doubt necessary, does not address the inherent differences in past and future performance that exist among students who earn the same diploma. The accountability structures unintentionally reinforce shifting the burden patterns linked to the later problems of college enrollment and persistence.

Systems thinking takes a different approach; we observe the interactions between student and school to try to prevent — not simply to catch — failures. While our early warning data systems may not be able to implement or fully prescribe interventions that promote structural changes (e.g., leadership; rigorous, ambitious instruction; parent-community relations; student-centered climate; professional capacity), these data systems can begin to model complexity and identify potential points of high leverage.

SIMULATION — A STRATEGY FOR IDENTIFYING ROOT CAUSE

According to Sterman, "simulations provide low-cost laboratories for learning,"[19] allowing us to test our conceptual models and to see the implications of those models unfold. In the absence of simulation, learning generally happens via real-world feedback that is often delayed.[20] In fact, in Figures 4–6, we see the relatively slow arc of structural changes taking place in a school in response to student performance across subsequent cohorts of students.

We have developed a simple simulation program that models the relationship between school resources and progress to graduation outcomes.

19 Sterman, 2006, p. 511.
20 Sterman, 2000.

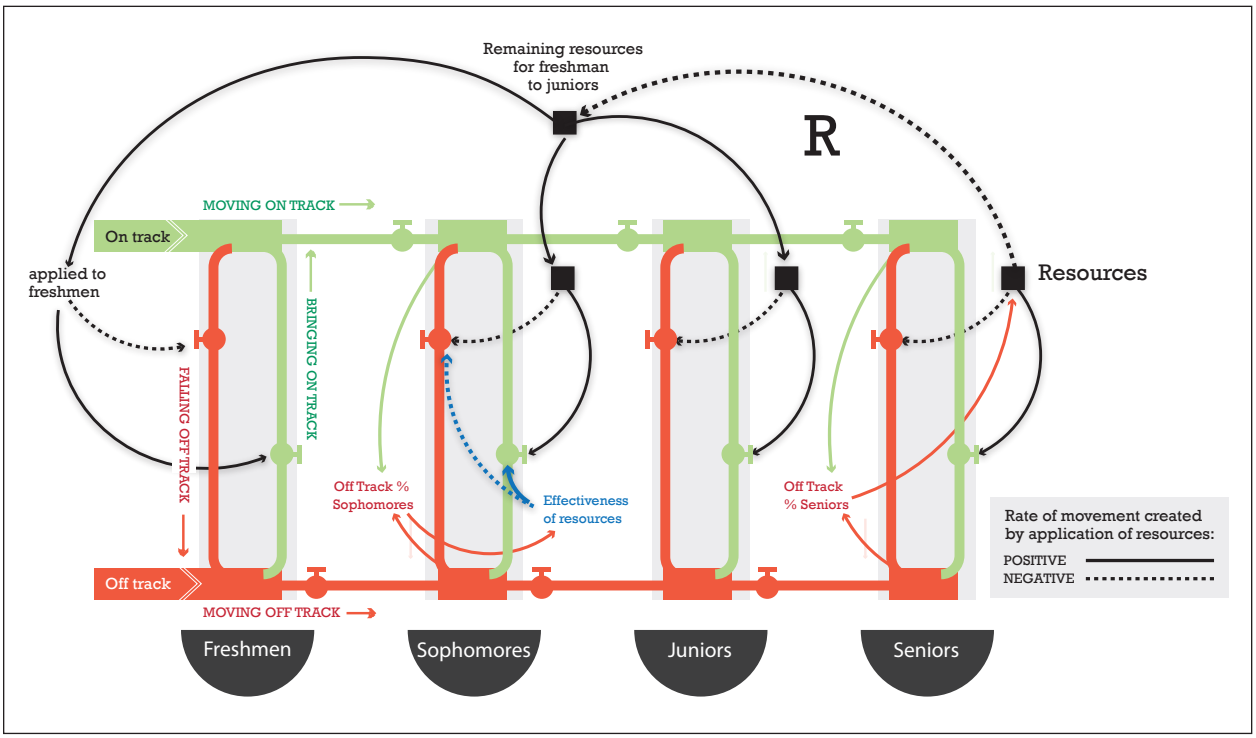
UNDERSTANDING THE MODEL

The first critical step in the simulation process is the articulation of our conceptual model (Figure 10).

In Figure 10, stocks are represented by the green and red bathtubs (e.g., "On Track Freshmen," "Off Track Freshmen," "On Track Sophomores"). The red and green pipes (the flows) that connect the stocks represent the filling and draining process. Students who were on track at the start may, over the course of the year, drop down into the off track stock. Conversely, some students who were off track at the beginning of the year may fill the on track stock. This pattern repeats across the four years of high school.

One of the objectives of the school is to increase the total number of on track students. To do so, schools apply resources to students. Resources include not only financial resources, but also time, the quality of adult talent, technology, and focused attention on an issue as a system priority. In Figure 10, if resources are applied to freshmen, off track freshmen may progress to a higher performance category while fewer freshmen drain out of the higher performance category. When more resources are appropriately applied, a greater percentage

Figure 10. Model of progress to graduation and relationship to resource distribution



We use the computer modeling and simulation software STELLA to (1) construct a dynamic model represented in Figure 10 above, (2) operationalize the model, and (3) simulate different scenarios by manipulating various parameters of the model. Important building blocks of our model include stocks, flows, and feedback loops — all of which are explained in this report. Figure 10, however, provides the reader with a slightly more technical view into the modeling process. Stocks (red and green tubs), flows (red and green pipes), connectors (black, red, and blue arrows), and converters (black, red, and blue text) represent important system parameters and interconnections that will define the behavior of the system. For instance, "Resources applied to freshmen" is a converter or a rate. Converters "open" and "close" the faucets on the flows (pipes). When we apply more resources in the freshman year, we open the faucet that allows more off track freshmen to flow into the on track freshmen stock. When we apply fewer resources in the freshman year, we open the faucet, thereby allowing more on track freshmen to flow into the off track freshmen stock. It is important to keep in mind that simulation is not an "exact" science. Rather, it explicates a structure (a set of causal assumptions) and from that structure generates behavior about a complex system.

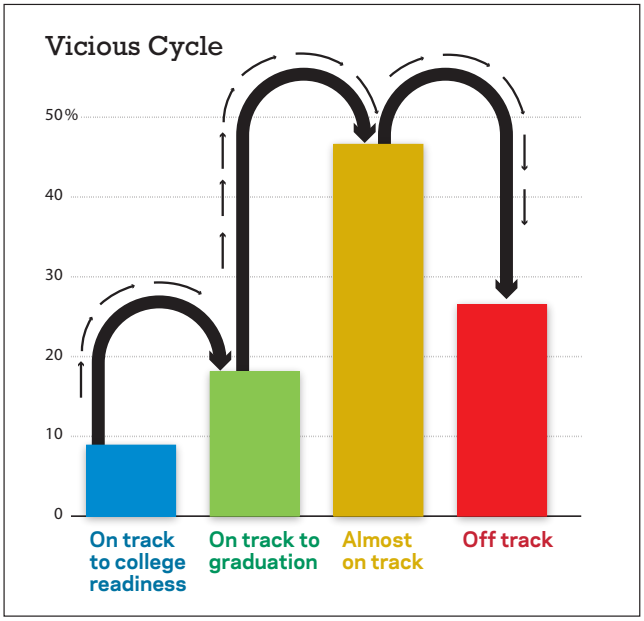
of students flow into higher-performance categories. (Resources may be badly used. The simulation can tell us much about the points in time to intervene and which student should be the object of intervention, but not what to do.) If fewer resources are applied to freshmen, then fewer numbers of freshmen progress. Resources are applied to each grade (year).

WHAT WE CAN SIMULATE THROUGH THE MODEL

Assume a school, through the best of intentions, decides to allocate intensive resources to intervene with those seniors on the cusp of not graduating. Assume also that there are limited total resources in a school. The more resources applied to the senior class, the fewer remaining resources can be applied elsewhere (i.e., to the earlier grades). As resources to freshmen are reallocated to seniors, the rate of moving off track freshmen into higher-performance categories is reduced; the fewer the resources, the more students can fall off track. This means that by the time freshmen become sophomores, there will likely be a higher percentage of off track sophomores than would otherwise have occurred. Assuming everything else stays the same (which is unlikely), the school will also end up with a higher percentage of off track seniors. This then leads the school to allocate *even more* resources to seniors.

This creates a vicious cycle, or a "reinforcing loop" (denoted with an "R"). Fewer students may enter their senior year prepared to graduate; and the more resources applied to seniors, the less likely it is that students will enter senior year on track in the future. However, this reinforcing loop could turn into a virtuous cycle. Assume that more resources are applied to freshmen. A greater percentage of on track students will have moved through the system, and the school will spend fewer resources on seniors.

Figure 11. Simulation results: resources and building foundational skills applied to senior year — the vicious cycle



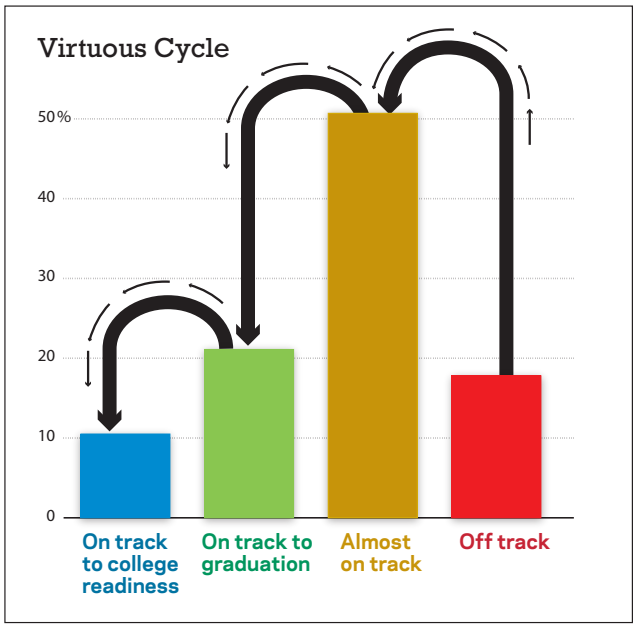
In addition, students on track earlier in their high school careers are more likely to have developed the foundational skills that allow them to progress. These students will be easier to teach, and the resources applied to them will be more effective. In fact, when we run this simple simulation, we see that our conceptual model holds.[21] Figures 11 and 12 capture results of the simulation. We begin year 1 with the average New Visions high school: 10 percent on track to college readiness (blue), 20 percent on track to graduation (green), 50 percent almost on track (yellow), 20 percent off track (red). If greater resources are applied in the

21 Refer to Appendix 2.

senior year and many of those seniors are relearning foundational skills, we increase the total number of off track (red) students in the school; we don't substantially increase the number of on track to graduation (green) students; and our on track to college readiness (blue) percentages decrease. Figure 11 represents the vicious cycle.

When we apply more resources during freshman year and those resources also build stronger foundational skills, we've increased the percentages of students in a school who are on track to college readiness (blue) and on track to graduation (green). We have also increased the percentage of almost on track (yellow) students. *But this increase is the result of moving the off track (red) students into the almost on track category — not the result of the on track to graduation (green) students slipping.* Figure 12 represents our virtuous cycle.

Figure 12. Simulation results: resources and building foundational skills applied to freshman year — the virtuous cycle



The immediate dilemma, of course, is what to do with the current cohort of upperclass students. Allocating more resources to freshmen means fewer resources for other students. This is a classic systems story of "worse before better." It creates a tension between the short term and the long term, and it is a common reason why systems do not improve. However, school-level progress to graduation maps (Figures 4–6) and simple simulations do allow schools to consider where resources can be better applied and where resources could be withdrawn without negative impact.

Conclusion and Future Directions

Within a systems thinking framework, all parts of the school are connected. The ninth graders who are at risk of dropping out, for example, and the seniors who are at risk of not graduating are not independent of one another. They vie for the same resources from the same administrators and teachers in the same school setting. In this respect, risk is not merely a specific student at a specific moment in time. Rather, risk is a property of the system that emerges from the interactions between the students and the school. This new conceptualization of risk necessitates the expansion and/or redesign of early warning systems that not only alert us to specific student events but that present risk as a system-wide phenomenon.

At New Visions, we are particularly invested in helping our schools to identify structures that are systematically inducing student flows from higher to lower performance categories. That is, for students flowing between different levels of performance, what accounts for the variance? To date, we have considered stocks and flows at the macro school-level. This level of aggregation undoubtedly hides departmental volatility. If a student is consistently almost on track (yellow), does that imply consistent (if mediocre) performance across all subject areas, or are these students catching up in some subjects while falling behind in others? Are off track students failing across the board, or are some passing most classes but have one subject where they cannot make headway?

The next step, then, will be to understand transitions from stocks through flows at the departmental level and to understand how the structural volatility described in this report is composed of several smaller departmental structures that influence students' progress through each subject and through the Regents exams. As some departments succeed in moving their students forward and others struggle, what is the effect on the school as a whole? How does a school respond when one or two subjects are largely responsible for delaying student progress, and how does that response limit a school's options in offering advanced classes or other means of advancing college preparedness?

And what about the students who seem to fall outside the structural volatility of a school — those students who cannot pass classes or even regularly attend school, no matter what interventions are applied? They may not be a particularly volatile group in terms of their own performance, but certainly they impact the structural volatility of a school. How are these students different from those making at least some progress, however little? How do schools respond when this block of students reaches a tipping point within the school?

Traditionally, the success of a high school is measured by its graduates, particularly those graduates who have earned a Regents Diploma or better. This, however, is not enough. *How* a student arrives at the Regents Diploma matters; and this not only has implications for postsecondary success but also suggests something about the way we conceptualize and manage risk in our schools. Only when we understand the interdependencies between students and schools can we design early warning systems that present risk holistically and that help us to promote "virtuous" volatility within our schools.

References

Allensworth, E., & Easton, J. (2005). *The on-track indicator as predictor of high school graduation*. Chicago, IL: Consortium of Chicago School Research.

Balfanz, R., Herzog, L., & MacIver, D. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grade schools: Early identification and effective interventions. *Educational Psychologist, 42* (4), 223–235.

Berry, W. (1983). *Standing by Words*. Berkeley, CA: Counterpoint.

Bolon, C. (2001). Significance of test-based ratings for metropolitan Boston schools. *Education Policy Analysis Archives, 9* (42), 1–57.

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.

Darling-Hammond, L. (2003). Standards and assessments: Where we are and what we need. *Teachers College Record*. Retrieved May 31, 2012, from http://www.tcrecord.org/content.asp?contentid=11109

Fairchild, S., Gunton, B., Donohue, B., Berry, C., Genn, R., & Knevals, J. (2011). *Student progress to graduation in New York City high schools: Part 1: Core Components*. New York, NY: New Visions for Public Schools.

Gardner, D. P. (1983). *A nation at risk*. Washington, DC: The National Commission on Excellence in Education, US Department of Education.

Halvorson, R., Grigg, J., Pritchett, R., & Thomas, C. (2005) *The new instructional leadership: Creating data-driven instructional systems in schools*. Paper presented at the Annual Meeting of the National Council of Professors of Educational Administration, July 25, Washington, DC.

Heppen, J. B., & Therriault, S. B. (2008). *Developing early warning systems to identify potential high school dropouts*. Washington, DC: National High School Center.

Kane, T., & Staiger, D. (2001, April). *Volatility in school test scores: Implications of test-based accountability systems*. Cambridge, MA: National Bureau of Economic Research.

Kieffer, M. J., & Marinell, W. H. (2012). *Navigating the middle grades: Evidence from New York City*. New York, NY: The Research Alliance for New York City Schools.

Meadows, D. (2008). *Thinking in Systems: A Primer*. Vermont: Chelsea Green Publishing.

Neild, R. C., Balfanz, R., & Herzog, L. (2007). An early warning system. *Educational Leadership, 65* (2), 28–33.

Roderick, M. (1994). Grade retention and school dropout: Investigating the association. *American Educational Research Journal, 31* (2), 729–759.

Roderick, M., Nagaoka, J., & Coca, V. (2009). College readiness for all: The challenge for urban high schools. *The Future of Children, 19* (1), 185–210.

Sterman, J. (2000). *Business Dynamics*. New York: McGraw-Hill.

Sterman, J. (2006). Learning from evidence in a complex world. *American Journal of Public Health*, 96 (3), 505–514.

Tucker, B. (2010). Putting data into practice: Lessons from New York City. *Education Sector Reports*, October 2010. Retrieved March 14, 2011, from http://www.oxydiane.net/IMG/pdf_Putting_Data_Into_Practice_RELEASE.pdf

Appendix 1

New Visions' Progress Toward Graduation Metric: 2 Term Schools

For a student to be considered On Track for College:

By the end of semester...	A student must earn this many credits:	Including this many credits in:													and	Regents passed at 65 and above[4]:	
		English	Math[1]	Social Studies					Science[1]			Foreign Language[1]	PE/Health	Art/Music			
				Total[2]	Global	American	Economics	Participation in Gov't	Total[3]	Physical Science	Life Science						
1	5	1	1	1	0	0	0	0	1	0	0	0	0.58	0		1*	*Including a 75 in Math **or** ELA
2	11	2	2	2	0	0	0	0	2	0	0	0	1.16	0		1*	
3	16	3	3	3	0	0	0	0	3	0	0	1	1.74	0		1*	
4	22	4	4	4	0	0	0	0	4	0	0	2	2.32	0		3*	
5	27	5	5	5	0	0	0	0	5	0	0	3	2.9	0		3*	
6	33	6	6	6	0	0	0	0	6	1	1	4	3.48	0		7**	**Including a 75 in Math **and** ELA
7	38	7	7	7	4	2	0	0	7	2	2	5	4.06	1		7**	
8	44	8	8	8	4	2	1	1	8	2	2	6	5	2		9**	

For a student to be considered On Track for Graduation:

By the end of semester...	A student must earn this many credits:	Including this many credits in:													and	Regents passed at 65 and above:
		English	Math	Social Studies					Science			Foreign Language	PE/Health	Art/Music		
				Total[2]	Global	American	Economics	Participation in Gov't	Total[3]	Physical Science	Life Science					
1	5	1	1	1	0	0	0	0	1	0	0	0	0.58	0		0
2	11	2	2	2	0	0	0	0	2	0	0	0	1.16	0		1
3	16	3	3	3	0	0	0	0	3	0	0	0	1.74	0		1
4	22	4	4	4	0	0	0	0	4	0	0	0	2.32	0		3
5	27	5	5	5	0	0	0	0	5	0	0	0	2.9	0		3
6	33	6	6	6	0	0	0	0	6	1	1	0	3.48	0		5
7	38	7	6	7	4	2	0	0	6	2	2	1	4.06	1		5
8	44	8	6	8	4	2	1	1	6	2	2	2	5	2		5

For a student to be considered Almost On Track for Graduation:

By the end of semester...	A student must earn this many credits:	Including this many credits in:													and	Regents passed at 65 and above:
		English	Math	Social Studies					Science			Foreign Language	PE/Health	Art/Music		
				Total[2]	Global	American	Economics	Participation in Gov't	Total[3]	Physical Science	Life Science					
1	3	0	0	0	0	0	0	0	0	0	0	0	0	0		0
2	8	1	1	1	0	0	0	0	1	0	0	0	0	0		0
3	12	1	1	1	0	0	0	0	1	0	0	0	0	0		0
4	20	2	1	2	0	0	0	0	1	0	0	0	0	0		0
5	23	3	1	3	0	0	0	0	1	0	0	0	0	0		1
6	30	4	2	4	0	0	0	0	2	0	0	0	0	0		2
7	34	5	3	5	0	0	0	0	3	0	0	0	0	1		3
8	38	6	4	6	0	0	0	0	4	0	0	0	0	1		3

1. College Readiness. The fourth year of math and science is required for College Readiness, but not for any diploma type. The 6 Foreign Language credits are required for being on track to an Advanced Regents Diploma and College Readiness, but not a regular Regents diploma. CTE students may receive an Advanced Regents Diploma with 2 Foreign Language credits, and 7 Regents passed at 65 or above. Advanced Regents Diploma is also available for students in an approved arts sequence.

2. Social Studies Total. The "Total" column includes the total number of social studies credits students must have earned by the end of each semester. In some cases, that includes specific credits in "Global," "American," "Economics," and/or "Participation in Goverrntment" courses.

3. Science Total. The "Total" column includes the total number of science credits students must have earned by the end of each semester. In some cases, that includes specific credits in "Physical Science" and/or "Life Science" courses.

4. Math Regents. Starting with the Class of 2011, students need 3 math exams (Integrated Algebra, Geometry, and Algebra 2/Trigonometry) to graduate with an Advanced Regents Diploma.

5. Students with an IEP eligible to take the RCT are considered On Track to Graduation or Almost on Track if they have acquired RCTs in place of Regents exams and have earned the number of credits needed for each term. Students with an IEP are not considered to be On Track to College Readiness unless they have fulfilled the appropriate Regents and credits needed for a particular term.

Appendix 2. Technical Notes

HIGH SCHOOL READINESS

New Visions High School Readiness (HSR) scale is based on the ELA and math deciles assigned to every student by the city. A decile score of 9 or 10 means the student is at or above proficiency and is expected to perform relatively well in high school; a decile score of 1 or 2 marks the student as below proficiency. When eighth grade test scores are available, the deciles are calculated directly from these scores. When they are unavailable – for example, for students not in New York for eighth grade – the city calculates a separate decile based on such factors as race, poverty, ELL, Sped, and SIFE status.

The HSR metric combines the ELA and math deciles and divides them into ranges that generally correspond to the proficiency levels on state tests. A decile score between 1 and 2 is labeled Off Track (Well Below Proficiency); between 2 and 6.5 is Almost on Track (Below Proficiency); between 7 and 9 is On Track (Proficient); and between 9.5 and 10 is Exceeding on Track (Above Proficiency).

STELLA SIMULATION MODEL

The simulation model, which employs empirical data from more than 5,000 students across eight semesters (more than 40,000 observations), is intended to demonstrate the downstream effects of complex, multifactorial systems as users manipulate inputs from two key dimensions (resources and fundamental skills). It is important to understand when using the simulation tool that this model more appropriately demonstrates directionality and relative magnitude than it does actual magnitude. The empirical data used to calibrate the model insufficiently captures certain of the inputs that would specify actual magnitude.

Based on individual school experience, users can calibrate two dimensions:

1. Relative resources invested in each grade level at a school. Assuming an unchanging stock of resources, users can manipulate the relative resources applied to each grade. Starting at a default baseline of 25 percent of resources in each of four grades, users can calibrate resources in any one grade between 0 and 40 percent of total resources, and the other grades will increase or decrease commensurately and proportionally.

2. Relative application of fundamental skills (within that resource allocation). Fundamental skills are core building blocks that are needed to progress. These fundamental skills take on more or less primacy at each grade level. Users can calibrate the application of these fundamental skills within a grade level based on school experience.

Once the user calibrates these two key dimensions, the simulation model uses empirical data to calculate a propensity score at each point-in-time measurement, which has a cascading effect on all future points in time. Therefore, by manipulating the model on two key dimensions at any or all points within the four years of high school, a school can optimize both its resource allocation and its application of fundamental skills to maximize student success.

NEW VISIONS
FOR PUBLIC SCHOOLS

www.newvisions.org


www.facebook.com/NewVisionsforPublicSchools


@NewVisionsNYC

ABOUT NEW VISIONS FOR PUBLIC SCHOOLS

New Visions for Public Schools, founded in 1989, is dedicated to improving the quality of education children receive in New York City's public schools. Working with the public and private sectors, New Visions develops programs, solutions, and strategies to energize teaching and learning and to raise the level of student achievement. As a Partnership Support Organization (PSO), New Visions is accountable for improving student achievement in 75 New York City public schools, serving more than 35,000 students. As a charter network, New Visions is opening two charter schools, with plans for a network of 18 charter schools over the next few years. As a laboratory, New Visions is researching and developing novel solutions for schools, teachers, and students. The overarching goal is to graduate all students ready and successful for college, career, and life.

New Visions for Public Schools • 320 West 13th Street • New York, NY 10014