Identifying Potential Dropouts: Key Lessons for Building an Early Warning Data System

A Dual Agenda of High Standards and High Graduation Rates

A white paper prepared for
Staying the Course: High Standards and Improved Graduation Rates,
a joint project of Achieve and Jobs for the Future,
funded by Carnegie Corp. of New York

Prepared for Achieve, Inc., by
Craig D. Jerald
President, Break the Curve Consulting
PART I. INTRODUCTION

Last year many of America’s political, business, and education leaders convened at a National Education Summit in Washington, D.C., to discuss ways to improve the nation’s high schools. Following the summit, 22 states joined the American Diploma Project Network, an initiative to better align high school graduation requirements with the skills young adults need to do well in today’s labor market and higher education.

High school graduates stand to benefit enormously from such policies. But increasing the value of the high school diploma will do little to help those who don’t earn one.

Therefore, Achieve and Jobs for the Future (JFF) have joined other organizations and foundations in calling for “dual goals” to guide high school improvement efforts. Policymakers must find ways to raise graduation rates even as they simultaneously work to raise standards for graduation.

Unfortunately, most dropout prevention programs conducted over the previous 30 years have garnered only disappointing results. And many people still consider the two goals to be in natural conflict with one another — believing that the cost of raising standards must inevitably be higher dropout rates, or that raising graduation rates can only be accomplished by lowering standards.

Recent experience suggests that worries about such zero-sum trade-offs might be misguided. During the last decade, a number of states and districts have raised graduation standards without seeing the plummeting graduation rates that critics of such policies often prophesize. But holding graduation rates steady cannot be the policy goal.

Thirty years ago, most teenagers who dropped out of high school could expect to find a well-paying job, and most who worked hard could expect to climb the economic ladder. But the world has changed. Today, high school dropouts face diminishing opportunities and a lifetime of financial struggle. In fact, the median earnings of families headed by a high school dropout declined by nearly a third between 1974 and 2004.

And wage declines are only part of the picture. Last year, Columbia University hosted a conference on the cost of educational failure that presented evidence that high school dropouts are more likely to be unemployed, receive public assistance, commit crimes and become incarcerated. At the same time, they are less likely to receive job-based health insurance and pension plans, be healthy and live as long, and vote and make other kinds of civic contributions.

That means the larger society stands to benefit from dropout prevention as well. Keeping all students in high school and graduating more young people with better skills will save millions of taxpayer dollars, greatly expand tax revenues, reduce crime, and improve citizenship.
Fortunately, recent research and experience convey three important lessons:

1) The dropout problem is not an inevitable, immutable feature of American education. Demographics matter, but what happens in schools has a great impact on whether students stay in school and graduate. Recent research suggests that, even for students who have difficult home lives, dropping out has much to do with how schools operate and the educational experiences students have within them. Moreover, the conventional wisdom that dropping out is a highly idiosyncratic process driven by entirely personal factors is not true for most students who leave school. Most dropouts follow identifiable pathways through the education pipeline.

2) We can do a much better job predicting which students are most likely to drop out. An ongoing study in Philadelphia can now identify half of all eventual high school dropouts as early as 6th grade. And a group of Chicago researchers can predict 85 percent of eventual dropouts in that city’s public school system based on just a few facts about the system’s 9th graders.

3) Finally, we know more than ever before about how schools contribute to high dropout rates and what educators can do to solve the problem. (See the companion JFF paper, Improving Outcomes for Struggling and Out-of-School Youth).

This white paper was prepared for Staying the Course: High Standards and Improved Graduation Rates, a joint project of Achieve and JFF funded by Carnegie Corp. of New York. Its goal is to provide policymakers with an overview of research about the dropout problem and the best strategies for building an early warning data system that can signal which students and schools are most in need of interventions.

As pressure mounts to do something about the dropout problem, many school systems will be tempted to skip questions about how to predict which students are most at risk of dropping out and simply begin with reforms meant to solve the problem. Leaders might assume that educators can do a pretty good job guessing which students are at risk based on subjective judgments. Or they might simply decide not to spend money on data systems but rather invest all of their dollars in interventions and reforms instead. As this paper attempts to demonstrate, however, such decisions can have a variety of negative consequences.

Indeed, the cost of building an accurate Early Warning System is relatively small compared with the cost of providing programmatic interventions or systemwide reforms meant to increase graduation rates. But the payoff of basing interventions on accurate data can be huge. A large school system that invests in better data to support dropout prevention can obtain much better results for hundreds of thousands or even millions of dollars less than a similar system whose leaders decide to skip that step.
PART II. WHAT KINDS OF STUDENTS DROP OUT OF HIGH SCHOOL, AND WHY?

Knowing how best to predict dropping out requires knowing something about the kinds of students who drop out of high school and the reasons they give for doing so. Researchers have identified three kinds of factors that put students at greater risk for dropping out:

1. STUDENTS’ SOCIAL BACKGROUND

Not surprisingly, a voluminous body of research has shown that factors related to students’ social and family background can increase or decrease the risk of dropping out:

- Students who are poor, who are members of minority groups, who are male, who transferred among multiple elementary and middle schools, and who are overage for their grade are more likely to drop out of high school.

- Students who come from single parent families, have a mother who dropped out of high school, have parents who provide low support for learning, or have parents who do not know their friends’ parents well also are placed at greater risk.

- Finally, studies have suggested that teenagers who take on adult responsibilities — becoming a parent, getting married, and holding down a job — are also more likely to leave school without a diploma.

However, several studies also have found that some social background indicators are neither good nor bad, per se, but instead influence dropping out in very complex ways. For example, one study found that mobility between schools before 8th grade increases the risk of dropping out, but mobility during early high school can be beneficial. Another found that holding a job outside of school can exert a negative, neutral, or even beneficial impact depending on a student’s gender and type of job he or she has.

2. STUDENTS’ EDUCATIONAL EXPERIENCES

For a long time, researchers investigating the causes of the dropout problem focused mainly on the social and demographic characteristics of dropouts. Partly as a result, many educators still believe that dropping out is caused exclusively by students’ personal and family characteristics — things that don’t have anything to do with education and that educators can’t control. As one school principal told a writer for the American School Board Journal, “If you want to know why some students drop out, look at their parents — they pass their low aspirations on to their kids.”
As often turns out to be the case, however, the conventional wisdom is wrong. A large body of more recent research has documented that while demographic factors are important, they don’t tell the whole story or even the most salient part of it. Students’ educational experiences play a large role in shaping the dropout problem, too.

This is reflected quite clearly in surveys of recent dropouts. For example, in 1990 a federal research project surveyed a representative sample of dropouts about why they had left school. Out of 21 possible reasons reflecting a wide range of educational and personal factors, 51 percent of dropouts reported, “I didn’t like school,” and 44 percent said, “I was failing.” A team of researchers who further analyzed the results found that “in each of the racial/ethnic as well as gender groups, school-related factors are the most cited reasons for dropping out.”

A survey of the nation’s high school students conducted in 2002 obtained similar results. When students who had ever considered dropping out of school were asked why, 76 percent said school was boring and 42 percent said they were not learning enough — responses that once again beat out a long list of other possibilities by a substantial margin.

More sophisticated studies have helped verify and clarify the education-related factors that play a role in fueling dropout rates. Researchers have identified two important categories of educational risk factors:

1) **Academic performance**: Students who struggle in the classroom and fall behind academically are more likely to drop out. Low grades, low test scores, Fs in English and math, falling behind in course credits, and being held back one or more times all have been linked to lower chances for graduation.

2) **Educational engagement**: Students who become disengaged from school and develop disciplinary problems are more likely to drop out. High rates of absenteeism or truancy, poor classroom behavior, less participation in extracurricular activities, and bad relationships with teachers and peers all have been linked to lower chances for graduation.

Education-related factors should be important to policymakers and educators concerned with reducing dropout rates both because they are practical and because they are predictive. First, such factors describe something that takes place inside of school rather than something that happens in the home or that students bring with them from outside. As such, they can be helpful in determining the kinds of interventions that students might need and that education systems can provide.

Second, educational experiences can be powerful predictors of dropping out, either along with demographic factors or, in some kinds of communities, even apart from them. Especially in urban districts, where high proportions of students have demographic risk
factors, very low academic performance and educational engagement can help discriminate the ones most likely to be on the path to dropping out.

The same is true in other more homogeneous settings. For example, one study conducted among white, French-speaking Canadians found that, “once school experience is known, family, behavioral, personal, and social experiences do very little to improve the predictive accuracy.”

3. School Characteristics

Beginning in the late 1980s, some researchers became concerned that their peers were focusing entirely on individual-level factors behind dropping out, wondering whether schools themselves could be contributing to the problem, too. Since then, a convincing and ever-growing body of research has validated their suspicions: School-level factors play a significant role in determining whether students will earn a diploma. Institutions matter as much as individuals, and attending a high school with certain characteristics can itself be a risk factor for dropping out.

That’s not to say that some schools literally kick students out (although that does happen), but rather that variations in the dropout rate across schools in a district or a state are not just a function of the students they teach. For example, the Consortium on Chicago School Research has documented that dropout rates vary widely across Chicago high schools — even after researchers adjust for a host of individual risk factors, including race, gender, prior academic achievement, family socioeconomic status and whether students are overage when they enter 9th grade.

What characteristics of high schools make them more or less good at “holding” teenagers until graduation? Several recent studies have found that, student risk factors being equal, high schools with smaller enrollments, better interpersonal relationships among students and adults, teachers who are more supportive of students, and a curriculum that is both more focused and more rigorous exhibit lower dropout rates. And the benefits are especially great for low-achieving, low-income students.

School effects can be quite strong. For example, researchers Valerie Lee and Robert Croninger found that high schools with highly supportive teachers cut the probability of dropping out in half. Lee and David Burkam found that schools that offered fewer math courses below the level of Algebra I reduced the odds of dropping out by 28 percent, and those that offered Calculus reduced the odds by 55 percent.
PART III. CAN WE PREDICT WHICH STUDENTS WILL DROP OUT?

WHAT DOESN’T WORK?

Knowing the research described in the last section is important, but it does not automatically lead to a practical, reliable system for predicting whether any particular student will drop out. Consider the lesson of past dropout prevention efforts, which experienced a surge in popularity during the 1980s and early 1990s.

Following publication of *A Nation at Risk*, a number of major counter-reports lamented lack of attention to graduation rates on the part of the burgeoning “educational excellence” movement. Foundation officials and others worried that raising graduation standards would help the college-bound, but leave other high school students stranded in the “rising tide of mediocrity.”

States and districts responded. “By the late 1980s, virtually every major school system had grappled with the issue of high school dropouts — on the one hand, by trying to identify the extent of the problem in their communities, and, on the other hand, by committing resources to address the problem,” researcher Melissa Roderick observed in 1993.

The federal government pitched in by contributing $214 million between 1988 and 1994 for a School Dropout Demonstration Assistance Program.

As programs proliferated, experts become concerned about the methods being used to identify potential dropouts for interventions. For example, a brief published in 1989 by the National Dropout Prevention Center at Clemson University reported that schools often used “checklists” to assess students’ risk of dropping out. Such instruments typically were developed by examining the research literature to find out what characteristics were associated with dropouts or by interviewing dropouts themselves. The authors warned that the checklist approach had several major weaknesses:

First, it is difficult to know how many characteristics must be checked before deciding if a student is at risk. Second, this gross approach to identification often leads to students being misclassified at risk and placed in prevention programs although they probably would not have dropped out had nothing been done for them. Third, educators frequently borrow instruments designed for use with students in program locales very different from their own, limiting generalization of characteristics on the checklist.

A study conducted by Phillip Gleason and Mark Dynarski of Mathematica Policy Research, Inc., suggests that such concerns were quite valid. They examined the effectiveness of widely used risk factors by examining data from a sample of program sites funded by the federal School Dropout Assistance Program. No single risk factor yielded an actual dropout rate of more than 28 percent. (That is, the most powerful risk
factor was not very predictive: More than 70 percent of students with that risk factor would have graduated anyway.)

Some experts contend that risk factors are cumulative in nature, such that the greater number of risk factors a student has, the higher the probability he or she will drop out. Therefore, Gleason and Dynarski examined whether prevention programs using a combination of risk factors would better identify students who truly required interventions. Again, however, no combination of risk factors yielded a dropout rate of more than 28 percent.

The Mathematica researchers concluded that “dropout prevention programs often serve students who would not have dropped out, and [they] do not serve students who would have dropped out.”22 In other words, knowing that an indicator has been “correlated” with dropping out might not automatically make it a great, or even very good, predictor of whether a particular student will graduate or not.

To understand why, consider a hypothetical example: If a district survey revealed that 40 percent of dropouts had average grades below C, compared with only 20 percent of graduates, and, further, that 35 percent of dropouts had been suspended, compared with only 7 percent of graduates, one could say that dropouts in the district “were twice as likely to have had low grades and five times as likely to have had behavior problems.” That certainly makes it seem like the district could use those risk factors to “predict” dropping out.

However, those same numbers would also mean that the majority of the district’s dropouts (60 percent) had grades of C or above, and approximately two-thirds of them had never been suspended. Prediction requires more than simply knowing which personal and educational characteristics dropouts are more likely to have.23

To that end, the 1989 Dropout Prevention Center brief encouraged educators to adopt an alternative to checklists. The authors described a process whereby programs could administer a questionnaire to two groups — dropouts and graduates — and conduct a statistical analysis that considered all of their characteristics at once (often called a “regression analysis”) to develop a “statistically generated prediction formula.”24

Gleason and Dynarski reasoned that some programs might be using such a method, so they tested whether that approach would have helped the federally funded dropout prevention programs better identify students who were in real danger of dropping out of school. The good news was that it did. The bad news was, not by enough. Once again, the Mathematica researchers found that even if dropout prevention programs employed statistical formulas, they were apt to “serve more students who do not need dropout prevention services than students who do need them.”25
WHAT WOULD WORK BETTER?

The answer lies in how Gleason and Dynarski tested the predictive power of indicators commonly used by prevention programs. Simply put, “longitudinal data allowed risk factors to be assessed.”

Longitudinal data, information collected and accumulated over time, permit analysts to follow the progress of individual students as they progress from grade to grade as members of a “cohort,” a group of students who start out in the same grade at the same time. Such data make it possible to observe what happens to students who develop risk factors at any point along the way, and thereby to paint a more detailed, nuanced portrait of the patterns and pathways students tend to follow as they move toward dropping out or graduating from high school.

Of course, research using longitudinal information dates back to before the 1990s. But many researchers studying dropouts have been confined to using national data sets that wash out important differences among different regions and communities. Some researchers now believe that the power of place can be as important as the power of time for unmasking the pathways students take to dropping out.

Since the early 1990s, several longitudinal studies zeroing in on individual school districts have shed considerable new light on how, in addition to why, students drop out of school. Below we describe each study in some detail, because, taken together, they provide a rough blueprint for how districts can better identify the highly predictive risk factors for students in their own systems.

The Fall River Study

In a pioneering research project conducted during the early 1990s, Roderick studied the educational career pathways of a cohort of beginning 4th grade students in Fall River, a small urban school district in Southeastern Massachusetts. She examined academic performance, school engagement, and social background factors both for students who eventually dropped out and for students who eventually graduated. Roderick made several important discoveries:

First, she discovered comparing all dropouts to all graduates masks important variations among subgroups. At first, her analysis seemed to indicate that dropouts were immediately distinguishable by warning signs as early as the 4th grade: On average, 4th graders who would eventually drop out had substantially lower grades compared with their peers who would graduate. However, when she probed more deeply, Roderick discovered that the gap was being driven by two subgroups:
• The average 4th grade performance of dropouts was being pulled down by a subgroup of very low performers — about one-third of all dropouts — who would eventually drop out before even making it to 10th grade.

• The average 4th grade performance of graduates was being pulled up by a subgroup of high performers who would graduate in the top two-thirds of their class.

Second, her analysis revealed that the district had two very distinct types of dropouts — those who left school between 7th and 9th grades (early dropouts), and those who left during 10th through 12th grades (later dropouts). Not only did members of these two groups leave school at different points in the pipeline, they also had very different educational careers — exhibiting different risk factors at different points in time:

• Early dropouts started out with much lower grades in elementary school.

• However, later dropouts — who made up 65 percent of all Fall River dropouts — had 4th grade attendance and grades that looked no different from students who would eventually graduate in the bottom third of the class.

Third, Roderick uncovered powerful evidence that the transition years — during which students move from elementary to middle school and from middle to high school — had a decisive impact on students who would later drop out. During the transition to middle school, academic performance declined for nearly all students. But the deterioration was much steeper for students who eventually dropped out. The same thing happened later during the transition to high school.

In effect, the two transitions were like hurdles that some students tripped over, causing them to fall behind their peers in the trek to graduation. Even though their grades did not decline more than graduates’ during the non-transition years, most dropouts never made up the ground they lost during 6th and 9th grades.

Roderick found that educational engagement diverged in middle school as well, but in a somewhat different pattern: Later dropouts experienced a dip in attendance during 6th grade that continued to accelerate until the end of middle school and into high school.

**The Chicago Study**

The Consortium on Chicago School Research conducts revealing analyses using a longitudinal database that tracks every student who moves into and through the Chicago Public Schools (CPS). That research has enabled Consortium researchers to construct new and valuable indicators for predicting which students will drop out of high school.

For example, by combining two of the four high school risk factors they had observed to be strongly correlated with graduation (attendance, GPA, credits earned, and failing
grades), Consortium members have developed an on-track indicator that signals when 9th graders are falling seriously off the track to earning a diploma. A student is considered on-track at the end of 9th grade if he or she has accumulated enough course credits to earn promotion to 10th grade while receiving no more than one F (based on semester marks) in core academic subjects.28

Further analysis has revealed the on-track indicator to be an especially good predictor of whether students will graduate. For example, among students entering 9th grade in 1999, the four-year graduation rate for on-track freshmen was 81 percent, compared with only 22 percent for those who fell off track as 9th graders. Another way to look at it: The on-track indicator was 85 percent successful in predicting which members of the freshmen class would not graduate.

The on-track indicator has revealed other important lessons about how students move through the CPS pipeline. Last year the Consortium issued a report showing that, not surprisingly, students who had high test scores in 8th grade were more likely to stay on track when they got to 9th grade. However, high 8th grade test scores were no guarantee that a student would weather the difficult transition to high school, and vice versa:

In the 2003–04 freshman class, for example, of the students who entered with very high 8th grade test scores (those in the top quarter of their class), almost one-quarter were off track by the end of their freshman year. On the other hand, of the students who entered high school in 2003–04 with very low 8th grade test scores (those in the bottom quarter of their class), more than 40 percent were on track by the end of freshman year.29

Why don’t 8th grade test scores tell the whole story about who will fall off track during 9th grade? One reason is that high schools themselves play a substantial role in determining whether students stay on track, independent of the risk factors students bring with them when they enter. The Consortium demonstrated that even after controlling for student-level risk factors related to race, gender, prior academic achievement, family socioeconomic status and being overage for the grade, high schools varied enormously in both their 9th grade on-track rates and in their graduation rates.30

Moreover, in a separate study, Roderick and Eric Camburn found that Chicago high schools also vary widely on another important measure — rates of recovery from 9th grade failure — even after controlling for a wide array of student risk factors.31

**The Philadelphia Studies**

Researchers also have begun to study the pathway to dropping out of Philadelphia’s public high schools. One series of studies by University of Pennsylvania researcher Ruth Curran Neild and colleagues has shed important light on the role of 9th grade as a critical year on the road to graduation.
First, success or failure in 9th grade greatly impacts a student’s chances of graduating, even taking into account a wide array of 8th grade risk factors related to social background, previous academic performance and school engagement. In other words, 9th grade was a “make-or-break” year for many dropouts — a finding that led the researchers to reject the notion that the timing of events related to dropping out is random. The researchers concluded, “Our ability to predict dropout … increases considerably when we know how students fare during their high school transition year.”

Second, 9th grade acts like a kind of quicksand for many Philadelphia students. More Philadelphia dropouts leave during 9th grade than in any other grade level. That does not mean dropouts spend only one year in high school, however. In fact, fewer than 5 percent of dropouts leave during their first year in high school. Because of very high 9th grade retention rates, the most common number of years dropouts spend in Philadelphia high schools is three.

Finally, certain risk factors that students bring with them into high school — being overage, having reading and math scores significantly below grade level, low 8th grade attendance, and having failed courses in middle school — increase the odds that students will fail 9th grade. However, low attendance during the first 30 days of 9th grade is a more powerful predictor than any 8th grade factor, including test scores, age and academic failure.

A co-author of one those 9th grade studies, Robert Balfanz of Johns Hopkins University, is currently working with Liza Herzog, a researcher at the Philadelphia Education Fund, to investigate the next obvious set of questions related to predicting dropouts in Philadelphia: How early in middle school could the district identify students at high risk for falling off the track to graduation? What middle school risk factors best predict dropping out of high school?

To answer those questions, the researchers built a longitudinal database to follow a cohort of students who entered 6th grade in 1996–97, tracking their progress up until 2003–04, one year after they should have graduated.

First, Balfanz and Herzog conducted an initial screen of about 20 student characteristics to identify high-yield risk factors — ones for which 75 percent of students with the characteristic do not make it to 12th grade on time. Next, they examined the impact of those high-yield risk factors on outcomes at critical points along the pipeline — 8th grade test scores, age in 8th grade, 9th grade success or failure, and graduation from high school in four years or five years.
Although that project is still incomplete, it already has produced several highly important findings:

- Four 6th grade risk factors proved to be powerful predictors of falling off the track to graduation — low attendance (80 percent or lower), a failing mark for classroom behavior, a failing grade in math, and a failing grade in English;

- Sixth graders with any one of those risk factors had only a 10 percent chance of graduating on time and only a 20 percent chance of graduating a year late; and

- Such students also were far more likely to perform poorly on state assessments, become overage during middle school and fail 9th grade — “often for several years.”

The results seem to support what the Fall River study found: The transition to middle school is as critical as the transition to high school, and 6th grade risk factors can be powerful predictors of dropping out. In fact, Balfanz and Herzog found that they could predict nearly half of all Philadelphia dropouts as early as 6th grade.
Table 1. Examples of Risk Factors that Significantly Increased the Odds of Dropping Out of High School from District-Level Studies

<table>
<thead>
<tr>
<th>Type of Risk Factor</th>
<th>Chicago</th>
<th>Philadelphia</th>
<th>Fall River</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Performance</td>
<td>• Receiving more than one grade of F in core academic courses or not earning enough credits to be promoted during 9th grade</td>
<td>• Earning an F in English or mathematics during 6th grade</td>
<td>• Significant drop in grade point average from 8th to 9th grade</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Failing one or more courses during 8th grade</td>
<td>• Being retained in any grade during K–8 or in high school</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Entering 9th grade with math or reading scores below 8th grade level</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Being retained in 9th grade</td>
<td></td>
</tr>
<tr>
<td>Educational Engagement</td>
<td>(n/a)</td>
<td>• Low attendance (80 percent or lower) during 6th grade</td>
<td>• Significant drop in attendance beginning in 6th grade and worsening in subsequent years</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Receiving a failing classroom behavior mark during 6th grade*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Low attendance during 8th grade</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Low attendance during the first 30 days of 9th grade</td>
<td></td>
</tr>
</tbody>
</table>

* Recent longitudinal studies typically have not had good access to other measures of behavior, such as number of office referrals, that might prove predictive in addition to or in place of teacher reports.
ESSENTIAL LESSONS FROM LONGITUDINAL COHORT STUDIES

1. While there is no single pathway that every dropout follows, there are common patterns, common crisis spots in the pipeline and common signposts, too.

The notion that dropouts can follow observable and predictable patterns on the journey to leaving school runs afoul of another bit of conventional wisdom. Just as many people still believe that dropping out is caused entirely by individual social factors, many also believe that every dropout follows a unique journey — in other words, that dropping out is such a highly idiosyncratic, personal and complex phenomenon that no common patterns could ever be identified.

Roderick contends that her Falls River study “challenges the assumption that dropping out is largely an individualized phenomenon [with] ‘many different routes.’ … [If] the path to dropping out could be best characterized as one that is different for every youth … we would not have observed any clear patterns at all in trends in late-grade dropouts’ school performance.” And the more recent Chicago and Philadelphia studies clearly support her claim.

At the same time, however, longitudinal studies also have revealed that — at least in the districts that have been studied so far — dropouts are not a monolithic group whose members all exhibit the same risk factors and follow the same pathway to dropping out. A district might find two or three distinct sub-groups among dropouts, each of which follows a distinct pathway, along with a much smaller subset whose members exhibit no risk factors prior to dropping out.

2. Longitudinal studies consistently identify a certain set of general educational risk factors as highly predictive of dropping out.

The Fall River, Chicago, and Philadelphia studies did not identify new and surprising kinds of educational risk factors, but rather confirmed the general findings of earlier studies: Students who experience low or declining levels of either academic performance or educational engagement are much more at risk of falling off track and dropping out of school.

3. However, it would be risky to simply guess at the exact measures of academic performance and educational engagement that best predict dropping out for students in a particular grade level in a particular district.

For example, Balfanz and Herzog found that, when it came to academics, classroom grades trumped standardized test scores in predicting which 6th graders would someday drop out of Philadelphia high schools. As for engagement, classroom behavior marks turned out to be a much better predictor than the number of suspensions a student received.
Of course, that does not mean that standardized test scores are never useful. For example, in a study of risk factors for failing 9th grade — the biggest single predictor of dropping out for Philadelphia students — Neild and Balfanz found that, “In comparison to students who score at the 6th grade level or below in mathematics, those who score at the 7th grade level have odds [of failing 9th grade] that are lower by almost a quarter; those at the 8th grade level have odds of non-promotion that are lower by 42 percent … and those at the 9th grade level or above have odds that are lower by almost 60 percent.”\footnote{39}

Moreover, there are a few factors — such as low attendance, being overage for the grade and failing 9th grade — that show up as important predictors of dropping out in nearly every single study.

4. **There might not be one single “leading indicator” that all dropouts exhibit first.**

It would be easier to predict dropping out if all dropouts exhibited a common risk factor that started them on the path to leaving school. Indeed, several earlier studies suggested that dropping out is a process that begins when students fall behind academically, after which they become frustrated and disengaged from school.\footnote{40}

However, Balfanz and Herzog found that Philadelphia dropouts can begin by exhibiting problems *either* with academic performance *or* with school engagement during early middle school. “We saw 6th graders with poor behavior and attendance who were not failing math and English,” says Balfanz. “And we saw students who were failing, but who did not have low attendance or behavior problems. That leads us to speculate that at the beginning of middle school, there might be an academic track and a non-academic track to failure. However,” he adds, “one set of problems can lead to the other, problems can compound, and the tracks can merge later on.”\footnote{41}

He points out, “That also means you can’t just use straight academic predictors alone at the 6th grade level, because you will be missing a group of kids for whom demographics and adolescence are creating problems that will lead to academic struggle later on.”

In other words, for many dropouts, “one thing can lead to another.” But it is not always the *same* thing.

5. **Dropouts differ not only in how they begin on the path to dropping out, but where they end it.**

As Roderick’s Fall River study found, dropouts in a district can “clump” into subgroups based on the grade level in which they leave school, and those subgroups can exhibit very different educational pathways to dropping out. For example, she found that Fall River dropouts who left before 10th grade exhibited clear warning signs as early as 4th grade, but later dropouts showed their clearest warning signs during the transitions to middle school and high school.
6. The transition to middle school and high school is a critically important time for students. Many dropouts begin to show warning signs and downward trends in 6th grade or 9th grade.

Students face many changes when they transition from elementary to middle school and from middle school to high school. In addition to having to negotiate a new and often larger institutional setting, students find that coursework is more intellectually demanding, teachers are less supportive, peer groups are larger, relationships are more complicated, and temptations become greater at the same time that they begin to experience more personal freedom.

Not surprisingly, many students, not just eventual dropouts, experience declines in academic performance or school engagement during the transition years. But students who will eventually drop out often experience especially big declines from which they never recover.

7. Transition-year failure often begins very early, and powerful risk factors can be observed even during the first month of school.

Neild and Balfanz found low attendance during the first 30 days of 9th grade to be a more powerful predictor of later failure for Philadelphia high school students than any risk factor they exhibited in 8th grade. And Elaine Allensworth and John Easton note that the correlations between 9th grade risk factors and dropping out in Chicago are only slightly less predictive when measured at the end of the first semester as they are when measured at the end of the year.

8. Students with poor prior achievement and behavior are more likely to fail during transition years. However, it can be hard to predict transition-year failure based solely on information collected the year before students change schools.

In her Fall River study, Roderick found that “students for whom difficulty following school transitions was a defining event were not necessarily those who were doing the most poorly in school before the school move.” For example, when they were 4th graders, later dropouts had grades and attendance comparable to graduates. However, the decline in their academic performance during 6th grade was almost three times as big as the decline experienced by their peers who went on to graduate.

Similarly, Consortium researchers found that low 8th grade test scores do not automatically doom students to 9th grade failure, and vice versa:
Almost one-quarter of students who began high school with high 8th grade achievement (those in the top quarter of their class) were not on track at the end of their freshman year. Although these students graduated from [8th grade] with strong test scores, only 37 percent of these off-track students graduated from high school within four years.  

9. However, that does not necessarily mean that educators should wait until 9th grade to begin looking for warning signs.

Balfanz believes the results of the current Philadelphia study call for a secondary school response to the high school dropout problem. Among Philadelphia dropouts who exhibited warning signs prior to 10th grade, the largest group by far was the group whose members exhibited them in 6th grade.  

Moreover, additional research conducted by Roderick in Fall River suggested that “better performance during the transition to middle school would substantially reduce the likelihood of dropping out.”  

10. School factors have an independent influence on whether incoming freshmen will graduate from high school or fall off track on the way to a diploma.

Researchers affiliated with the Consortium on Chicago School Research have demonstrated that it is possible to calculate “adjusted” versions of several important high school outcome measures — staying on track in 9th grade; being promoted to 10th grade; recovering from 9th grade academic difficulty; and four- and five-year graduation rates.

Since high schools still vary widely even in their adjusted outcomes, districts can use such data to identify how many and which high schools have an especially negative impact on whether students stay on track to graduate, and vice versa.
PART IV. PUTTING IT ALL TOGETHER: ADVICE AND KEY POLICY QUESTIONS FOR BUILDING AN EARLY WARNING DATA SYSTEM

The previous two sections provided an overview of the lessons that long experience and recent research have taught us about the feasibility of identifying potential dropouts. This section addresses how policymakers can put those lessons into action.

We now know that research is as important as development for building a good early warning data system. Skipping the research step can severely undermine development of an effective and cost-efficient system to identify dropouts for intervention. As the authors of one of the Philadelphia studies put it, “We argue that longitudinal studies of dropout are enormously useful in identifying particular points in students’ educational careers where school problems worsen or begin to manifest themselves.”

Therefore, school systems should start by replicating the kind of research project currently being conducted in Chicago and Philadelphia — before undertaking the development of a data system to provide early warnings to educators and before implementing intervention strategies.

Of course, that does not mean starting from scratch or working entirely alone. As described below, the existing research base provides a very solid foundation to build on. And there is no reason to suspect that groups of geographically and demographically similar districts cannot collaborate on research and development, share findings, and borrow from one another.

In general, however, building an effective and efficient early warning data system will require a two-phase process:

- Phase I — Research: Conducting a cohort-based, longitudinal study aimed at identifying the most effective and efficient risk factors for a given school system, uncovering patterns in how those risk factors play out for students over time, and assessing how to measure the impact of schools.

- Phase 2 — Development: Leveraging the knowledge gained in Phase I to create an electronic data system to inform intervention efforts.

That might sound as if doing this right will require a sizeable up-front investment in time and money. However, as we describe below, the research phase can take a very short amount of time and require relatively few resources.
PHASE I. CONDUCTING A LONGITUDINAL COHORT STUDY

Many education officials have the mistaken impression that conducting a longitudinal cohort study will require first developing a sophisticated student tracking system, after which they will have to wait for six or seven years as the system follows a cohort of students through the pipeline. In other words, it would take a lot of up-front time and money before you could begin to learn anything.

Fortunately, that’s not the case. The methods used by researchers conducting longitudinal analyses in places like Fall River and Philadelphia reveal a much quicker, cheaper alternative. Those researchers collected existing information on previous cohorts whose students had already moved through the school system by engaging in a kind of “paper chase” — pulling data from the student records that every school district maintains in paper files or electronic databases.

That means districts can save time and money by examining data on past cohorts to identify good predictors of what will happen to students in future cohorts. For example, a system that wanted to begin building a cohort data set to examine next fall would obtain information on the group of students who began as 6th graders during 1998–99. Those students should have graduated from high school in 2004–05. Data from 2005–06 will provide information on any of them who took an additional year to graduate.

Step 1: Collect the Necessary Information

Where Do We Obtain the Information?

Data for conducting the Phase I analysis will be much easier to obtain in districts that have been using a centralized, electronic database for maintaining comprehensive records on individual students since the mid-1990s. However, while some districts began moving from paper records to electronic records a long time ago, others have yet to take the first step.

In those districts, the information necessary to conduct the Phase I study will be spread across paper files maintained by multiple central office departments. In that case, analysts will have to spend time “going from door to door” to obtain the data. However, the information-gathering process need not take forever even under those circumstances. Balfanz reports that it took only about a month to accomplish it in Philadelphia, even though they had to obtain the information from multiple files:

- A basic demographic file with race, gender and special ed status;
- An enrollment file with information on whether, at the end of the prior year, students were still enrolled in school, had actively withdrawn or had graduated;
• A transcript file with information on classroom grades;
• An attendance file; and
• A test score file.

Which Starting and Ending Points in the Pipeline Should We Look At?

Analysts could theoretically pick any beginning grade level to define a cohort. Indeed, several studies have linked student achievement and behavior in 1st grade to dropping out during high school. However, the recent longitudinal studies described above suggest that the best starting point is probably the transition into middle school. In most districts, that will be 6th grade.

Note that Roderick found that declines in grades and attendance from late elementary school to middle school also revealed important information. That need not mean beginning with a full cohort of 5th graders. Districts can collect a subset of previous-year academic, attendance, and behavior data for their cohort of 6th graders to answer questions about “how they were doing” before they got to middle school.

Defining the end point might seem obvious. However, consider the Philadelphia studies suggesting that in some big city school districts, students spend more than one year, and sometimes up to three years, in 9th grade. Moreover, a federal cohort study tracking a national sample of 8th graders found that dropping out is seldom as simple as leaving and never returning. A large number of dropouts reenter the system, sometimes drifting in and out several times.

Therefore, analysts should place the ending point for collecting data on a cohort at least a year after the students in that cohort should have graduated, if only to look for differences in risk factors based on four-year (on-time) and five-year graduation rates. Of course, they might not find big differences, and that would be important to know as well: The Consortium on Chicago School Research found that its 9th grade on-track indicator worked just as well at predicting four-year graduation rates as five-year graduation rates.

This is not an argument for what kinds of dropout rate definitions should be used for public reporting and accountability, which is beyond the scope of this white paper. Indeed, policymakers might find it advisable to use a strict, on-time, four-year graduation rate for such purposes. The overriding goal of the Phase I study is simply to learn as much about a district’s dropouts as possible.
How Many Cohorts Should We Examine?

Systems will have to analyze data on at least one additional cohort to verify the findings of the Phase I study. This is especially important to keep in mind when planning information collection in districts where data must be obtained from paper records maintained by different departments.

What Pieces of Information Should We Include? How Many Possible Risk Factors Should We Test?

As discussed above, the current Philadelphia study suggests that it’s advisable not to make untested assumptions about possible education-related risk factors before conducting the Phase I study. For example, one might assume that, among measures of academic performance, standardized test scores would reveal as much or even more about risk for dropping out as classroom grades. Conveniently, standardized test scores are often easier to obtain since they are more often centrally maintained.

However, Balfanz and Herzog found just the opposite: Classroom grades are much better at predicting which Philadelphia 6th graders will drop out than are standardized test scores. The vast majority of Philadelphia students perform poorly on standardized tests, Balfanz reasons, but it is hard to actually fail 6th grade since teachers in middle schools are still fairly forgiving and often pass students based on effort. Therefore, classroom grades might be a better discriminator for identifying students who are truly falling off the track to earn a diploma.

Similarly, when it came time to measure the predictive power of factors related to educational engagement, Balfanz and Herzog found that marks given by teachers for classroom behavior were much more predictive than the number of suspensions a student received. Behavior marks are cumulative, reflecting student attitudes and actions over the course of a semester or year, while suspensions reflect sporadic occurrences. Also, behavior marks reflect what is taking place inside the classroom — whether students are paying attention and are engaged in learning — while suspensions are often given for serious infractions that happen outside the classroom (before the first bell, between classes or after the last bell at the end of the day).

Such surprising, counterintuitive findings reveal why it is important to begin the analysis by examining a large pool of potential indicators. For example, Balfanz and Herzog conducted a preliminary screen of about 20 student characteristics to obtain a final list of only four high-yield risk factors.
Table 2 provides a list of high-priority data elements to obtain for a Phase I analysis in each of the three categories of individual risk factors — social background, academic performance and educational engagement. Note that this is a “birds-eye view” of possible data elements. While we have attempted to provide examples of specific data to look for under broad categories like “discipline” — such as classroom behavior marks or office referrals — information maintained on students can vary widely from district to district. A district that does not have any of the examples listed might still find alternative pieces of information that measure the same thing.
Table 2. Top-Priority Student-Level Data Elements To Obtain for Phase I Analysis

<table>
<thead>
<tr>
<th>Social Background</th>
<th>Academic Performance</th>
<th>Educational Engagement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Middle School</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. SES — FRPL eligibility, family income, etc.</td>
<td>1. Grades in academic subjects including at least English and math by end of quarter, semester and year</td>
<td>1. Attendance — number of days or percentage of days absent</td>
</tr>
<tr>
<td>2. Race/ethnicity</td>
<td>2. Failing grades in math and English</td>
<td>2. Discipline problems — indicators of poor behavior, including, for example:</td>
</tr>
<tr>
<td>3. Gender</td>
<td>3. Scores on standardized assessments in at least reading and math, including grade-level and benchmark assessments</td>
<td>a) classroom behavior marks</td>
</tr>
<tr>
<td>4. Mobility — number of schools enrolled</td>
<td>4. Number of times retained in grade during elementary and middle school</td>
<td>b) number of office referrals</td>
</tr>
<tr>
<td>5. Years overage for grade</td>
<td></td>
<td>c) number of counseling referrals</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) number of suspensions</td>
</tr>
<tr>
<td><strong>High School</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. SES — FRPL eligibility, family income, etc.</td>
<td>1. Grades in core academic subjects, by end of quarter, semester and year</td>
<td>1. Attendance — number of days or percentage of days absent</td>
</tr>
<tr>
<td>2. Race/ethnicity</td>
<td>2. Number of courses failed and passed in core courses, by end of quarter, semester and year</td>
<td>2. Discipline problems — indicators of poor behavior, including, for example:</td>
</tr>
<tr>
<td>3. Gender</td>
<td>3. Number of credits attempted by semester, by year and cumulatively</td>
<td>a) number of office referrals</td>
</tr>
<tr>
<td>4. Mobility — number of schools enrolled</td>
<td>4. Number of credits earned by semester, by year and cumulatively</td>
<td>b) number of counseling referrals</td>
</tr>
<tr>
<td>5. Years overage for grade</td>
<td>5. GPA by semester, by year and cumulatively</td>
<td>c) number of suspensions</td>
</tr>
<tr>
<td></td>
<td>6. Ninth grade “on-track indicator” equivalent to or adapted from measure developed by Consortium on Chicago School:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a) earned enough credits to be promoted and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b) received not more than one semester F in core academic subject, with same calculated for subsequent grade levels</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7. On-time promotion to 10th grade</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8. Scores on standardized assessments, including grade-level, end-of-course, benchmark assessments and exit exams</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9. Dropped out previously and re-enrolled</td>
<td></td>
</tr>
</tbody>
</table>

Note: The numbering of data elements in this table is not meant to imply rank ordering on the basis of priority. All elements in this table are high priority. The relative predictive power of any element will vary by location.
Step 2: Analyze the Information To Identify High-Yield Risk Factors

The goal of this step is to identify a group of high-yield risk factors, ones that do the best job — and a good enough job — predicting which students will drop out. In their Philadelphia study, Balfanz and Herzog decided to define “high yield” as a 75 percent or greater chance of not graduating from high school on time (i.e., 75 percent or more of the students who had the risk factor did not graduate on time).

The analysis at this point is not complicated. Analysts simply follow students’ time, examining outcomes for subgroups within the cohort that exhibit various characteristics, determining what proportion drop out or graduate, as well as what proportion of students who do not exhibit a given risk factor drops out or graduates.

However, if that method does not identify high-yield risk factors, analysts can then begin to look for them in more complex configurations of data (and, to hone the analysis, they might want to examine more complex configurations anyway):

a) Combinations of Indicators. For example, analysts can examine what happens to students who exhibit different combinations of multiple risk factors. At this point, analysts should examine not just whether educational characteristics (academic performance and educational engagement) are higher yield in combination, but also whether educational risk factors are more predictive for males than for females or for some racial/ethnic groups more than others. For example, they might examine combinations, such as the following:

- Failing grade plus low attendance
- Failing grade plus male
- Failing grade plus low attendance, male and Hispanic

b) Changes in Indicators. Since several studies have found steep declines in academic performance or educational engagement to be good predictors of falling off course for graduation, particularly during transition years, analysts might want to examine changes in student characteristics from one year to the next. For example:

- Difference in GPA (or some kind of mathematical average reflecting overall academic performance) between 5th and 6th grades
- Difference in attendance between 5th and 6th grades

Note that the bulleted examples above are just that — examples of possible configurations of data to examine. In other words, there is not one formulaic set of steps for examining student characteristics to identify high-yield risk factors. Rather, analysts must take a more exploratory approach. However, as Balfanz points out, “If you have done a good job compiling a robust set of social background data, academic performance
data, and behavioral data, you can create any variations or combinations you might want to explore.55

**Step 3. Conduct a Pipeline Analysis Based on the High-Yield Risk Factors Identified in Step 2**

Analysts should observe the pathways that students with high-yield early warning signs follow as they progress through the pipeline. Are there any especially troublesome grade levels — such as the transition years — where indicators of academic performance or educational engagement begin to plummet? What is the impact of high-yield risk factors in determining intermediate outcomes along the pipeline before students drop out, such as staying on track or failing the 9th grade?

For example, after Balfanz and Herzog winnowed their list of potential indicators down to four “high-yield” risk factors, they next examined the impact of those four risk factors on key outcomes throughout the rest of the pipeline, including 8th grade test scores, ages at 8th and 10th grade promotion, and on-time and five-year graduation.

Analysts also should look for patterns in the grade level and year that students drop out of high school. Are there distinct subgroups of early leaving and late leaving dropouts? How large is each group? How do their educational pathways and warning signs differ?

Note that Philadelphia studies found that the year in which students left high school might not be obvious from looking at the grade in which they left high school. Therefore, examining both together might yield additional knowledge about a school system’s dropouts.

The goal of this stage is to paint a vivid portrait of what the pathways to dropping out look like in a given school district. That analysis will help them more clearly understand how early risk factors work over time, as well as helping them make recommendations to district leaders for how the early warning data system should function and at which grade levels the school system should provide intensive interventions.

**Step 4. Assess How Effective and Efficient Highest-Yield Indicators Are, as well as the Potential Payoff from Interventions Triggered by Them**

Recall that any given risk factor or combination of risk factors only measures the likelihood that a given student will drop out. In other words, a risk factor is never a true prediction, but rather an assessment of how much risk an individual student faces based on his or her membership in a group whose educational outcomes we can predict much more accurately based on past cohort data.

Therefore, policymakers and system leaders can use hypothetical scenarios to assess what various high-yield risk factors might produce if used as predictors to trigger interventions.
This process begins with knowing the answers to several questions about a risk factor: What percentage of students with the risk factor dropped out? What percentage graduated? What percentage of students without the risk factor dropped out? What percentage graduated? Using that information, system leaders can assess what proportion of dropouts would have been correctly targeted for interventions if a given risk factor or combination of risk factors had been used as the trigger.

This is especially important for planning purposes because it helps leaders assess both the potential effectiveness and potential efficiency of various risk factors:

The lower the number of “false negatives” (students a risk factor did not identify but who did drop out), the more effective the risk factor is. Zero false negatives means a district would have identified and provided interventions to all students who eventually would have become dropouts. While it’s not actually possible to get the number of false negatives down to zero in practice, the higher the dropout rate for students that have a risk factor, the more effective the risk factor.

Similarly, the lower the number of “false positives” (students a risk factor would have identified but who did not drop out), the more efficient the risk factor is. Zero false positives means a district would not have provided interventions to any students who would have graduated anyway, thus saving dollars.

Figure 1. How Risk Factors Identify False Negatives and False Positives

Consider the following hypothetical example based on a real risk factor — the 9th grade on-track indicator created by the Consortium on Chicago School Research. The Consortium reports that 81 percent of students who were on track by the end of 9th grade completed high school in four years, compared with only 22 percent of 9th graders who finished off track. Based on that information, it’s possible to calculate that the on-track indicator correctly predicted 85 percent of eventual dropouts among that group of entering freshmen.

Now assume that the CPS decided to provide targeted assistance to all off-track students at the end of 9th grade. Further assume that 30,000 freshmen enter Chicago high schools each year and about 58 percent of them fall off track, which means the program would provide interventions to 17,370 students. The risk factor would fail to identify 2,400 false
negatives, students who, despite being on track at the end of 9th grade, would eventually fail to graduate. Those students would not receive the help they need.

Conversely, the indicator would identify 3,821 false positives, off-track students who would have recovered and graduated anyway. Assume that the intervention program costs an average of $350 per student. That would make the total cost of the program $6,079,500, of which $1,337,490 (or 22 percent) would have been spent on false positives, students who didn’t need the extra help because they would have graduated even without it.

However, if the program is even halfway effective (i.e., if it achieves a 50 percent success rate for getting off-track students back on track to graduate), the CPS would:

- Reduce the number of dropouts from 15,948 to 9,174 (a 42 percent reduction);
- Increase the number of graduates from 14,052 to 20,826 (a 48 percent increase);
- Reduce the district’s dropout rate from 53 percent to 31 percent; and
- Increase the district’s graduation rate from 47 percent to 69 percent (i.e., to about the national average).57

While it is not perfect, the on-track indicator clearly is a potentially very powerful tool for curbing dropout rates. And it is not the only powerful risk factor at the district’s disposal.

Such evaluations of the effectiveness and efficiency of different risk factors can be conducted with very few pieces of information — the dropout rate for students who have the risk factor, the dropout rate for those who do not, the proportion of students who exhibit the risk factor and the number of entering freshmen (or 6th graders).

Step 5. Conduct a School-Level Analysis To Better Understand Where Individual Risk Factors Are Most Concentrated and Which Schools Put Students at an Even Greater Risk of Dropping Out

Once the most powerful student-level predictors of dropping out are known for a particular school system, the district can conduct a school-level analysis. For example:

a) Concentration of risk factors: Analysts should determine which middle schools and high schools receive especially high numbers of students with particular high-yield risk factors.

b) Contribution to risk factors: Analysts should identify middle and high schools where transition years are most problematic. For example, they can identify schools where:
• High numbers of students exhibit declines in academic performance or educational engagement during transition years;

• High numbers of students who previously showed no warning signs exhibit such declines during transition years; and

• Students in general experience especially steep declines.

c) **Impact on outcomes:** Analysts should evaluate how much individual high schools themselves contribute to the dropout problem independent of student-level risk factors. They can do that by using regression analyses to “control for” the risk factors that students carry with them when they enter 9th grade to determine which high schools have higher-than-predicted or lower-than-predicted outcomes. They also can calculate adjusted rates on important outcomes for each individual high school:

• Adjusted rate of freshmen staying on track (the on-track indicator);

• Adjusted 10th grade promotion rate;

• Adjusted rate of recovery among students who fall off track in 9th grade; and

• Adjusted four- and five-year graduation rates.
PHASE II. DEVELOPING AN EARLY WARNING DATA SYSTEM

What Kind of Database Will We Have To Buy or Build To Do This?

To keep track of which students develop risk factors, when they develop them, and how they fare throughout the pipeline, school systems will have to build electronic record-keeping systems, or data warehouses, that maintain information on individual students and schools over time. Many states and districts have already begun to invest in the development of such systems, especially to meet the reporting requirements of No Child Left Behind. Ideally, the early warning data system should piggyback on such efforts.

In general, such databases should include:

- A unique student identifier that allows an individual student to be tracked by grade level and from elementary to middle to high school;
- Accurate enrollment information on each student, including entry and exit by school attended;
- Student demographic information, including eligibility for the federal free and reduced-price lunch program, race/ethnicity, gender, and age;
- Student transcript information, including courses attempted, courses completed, grades, credits earned, and any instances of being retained in a grade;
- Student attendance;
- Student behavior grades or discipline records; and
- Student graduation and dropout information.

Districts will be in a much better position to construct such a database in states that have created statewide longitudinal data systems. For example, states can do a better job tracking whether a student who left a district dropped out entirely or enrolled in another district elsewhere. A publication entitled Creating a Longitudinal Data System, available online at www.achieve.org, provides greater detail on the necessary and desirable features of statewide longitudinal databases.

Not surprisingly, the number of commercial database software programs on the market and the number of vendors offering customized database assistance have increased greatly over the last few years. District leaders who would like a general overview of such software, as well reviews of more than a dozen of the more popular commercial...
packages, can consult the Data Use Web site maintained by Jeff Wayman of the University of Texas at Austin, http://edadmin.edb.utexas.edu/datause/index.htm.

**Which Warning Signs Should the System Monitor and Report On?**

The answer to this question should be determined primarily by the results of the Phase I analysis. System leaders should be wary of taking shortcuts. The primary value of an early warning data system rests in its level of its predictive power. For example, analysts might find that grades are more highly predictive than test scores, but be told by department heads and budget officers that “test scores are just so much easier and cheaper to include.”

Policymakers should examine the long-run trade-offs associated with following that advice. It might turn out to be much more cost-effective over the long run to include indicators that are more expensive to obtain and maintain in any given year but do a better job predicting which students are more likely to drop out.

Furthermore, it might seem odd to include risk factors that rely on human judgment and that psychometricians, for example, would consider less “statistically reliable” because they are “more subjective” and “less standardized.” However, it is important to keep in mind that the primary purpose of an early warning data system is prediction and prevention, not measurement per se. Thus, the value of any piece of data lies not in its general “reliability” but in its specific validity — it’s *predictive* validity.

**Which Grade Levels Should the System Monitor and Report On?**

The answer to this question should be driven by two factors — the grade levels identified by the Phase I study as the most useful for predicting dropouts and the grade levels at which it is feasible to provide interventions. In general, systems should target intervention to the grade levels that are the weakest points in their secondary pipeline, the “crisis spots” that set students on the road to dropping out.

Many districts are providing extra assistance to 7th and 8th graders to help them improve their standardized test scores. There can be little doubt that this is a positive development. Research by the Council of the Great City Schools and the National Center for Educational Accountability has found that monitoring student performance and providing extra academic assistance — as early and often as possible — can be an effective strategy for raising student achievement and closing achievement gaps between groups. Moreover, given that low academic performance generally puts students at greater risk of dropping out, such efforts might help reduce dropout rates as well.

However, policymakers should not assume that such programs automatically constitute the best possible approach to reducing dropout rates or that they can necessarily “kill both birds with the same stone.” For example, Roderick found that, in Fall River, “better performance in non-transition years would have a small impact on the conditional...
likelihood of dropping out. [But] better performance during the transition to middle school would substantially reduce the likelihood of dropping out.” And Balfanz and Herzog found that far more Philadelphia dropouts first showed warning signs in 6th grade than in 8th grade, and that classroom grades were better predictors of whether 6th graders would eventually drop out than were standardized test scores.

**What If a District Chooses Not To Undertake the Phase I Analysis?**

As discussed above, analyzing which precise risk factors are most predictive of dropping out is relatively quick, easy, and — best of all — inexpensive. And the long-term payoffs, both in terms of economic efficiency and graduation outcomes — can be huge. However, districts that cannot undertake such an analysis for some reason should still collect information to analyze student- and school-level needs before spending significant resources on programs meant to improve graduation rates. The major factors that put students at risk of dropping out — low academic performance and educational engagement — are well established, and a district can hedge its bets by collecting multiple kinds of information to measure each of those factors (for example, classroom grades and standardized test scores). In such case, Table 2 on page 24 can serve as a helpful summary to the basic kinds of research-based information to collect on students.

**How Frequently Should the System Monitor and Report on Risk Factors?**

If feasible, districts should monitor and report on warning signs in every middle and high school grade level on an annual basis. But they should consider more frequent monitoring and reporting if certain grade levels turn out to be particularly important sticking points on the road to graduation.

For example, longitudinal studies have revealed two important additional findings about transition years: First, few students recover from failure during transition years. Second, academic failure and disengagement occur very early during the transition years. Therefore, system leaders should consider configuring the early warning data systems to make multiple reports during transition years — at the end of every quarter or at the end of every semester.

However, policymakers should keep in mind that longitudinal studies also have taught us important lessons about the dangers of making untested assumptions and generalizing based on data from other places. A district’s Phase I analysis could conceivably reveal different pathways than those identified in the Fall River, Chicago and Philadelphia studies. Indeed, some national studies have found that the chance of dropping out increases from 9th through 12th grades — a quite different pattern than in Philadelphia, where three in four dropouts leave before ever reaching 11th grade.

Therefore, the decision about when and how often to report and act on warning signs should be driven primarily by what a system finds when it conducts its own Phase I pipeline analysis.
What Kinds of Data Should We Report to Whom?

An early warning data system can report many kinds of information helpful for reducing dropout rates. First of all, of course, the system can report which individual students exhibit risk factors to school or program staff. But reports of aggregate risk factors by school can be helpful to school and district leaders as they plan and budget for intervention programs. And reports of adjusted outcome measures by school can be useful to school-level committees working on improvement plans and to district officials considering schoolwide interventions and reforms.

Readers should consult the companion JFF paper, *Improving Outcomes for Struggling and Out-of-School Youth*, for a full overview of research on the effectiveness of various student intervention programs and whole-school reforms for reducing dropout rates, as well as recommended state and local policies for supporting such efforts. However, here are several examples that illustrate how data from the Early Warning System might be useful for making such decisions:

a) **Intervention: Extra help for certain groups of students who share a particular risk factor.** If many students exhibit risk factors related to low academic performance, for example, policymakers might decide to provide accelerated instruction to such students in the form of catch-up courses. Johns Hopkins University has developed such courses as part of its Talent Development model for middle school and high school reform, and recent research suggests such courses are particularly effective at helping students get back on track academically and, along with other student support, can increase credit accumulation, 10th grade promotion and graduation rates. Other secondary school reform models have developed such courses as well; for example, America’s Choice offers courses such as Ramp Up to Literacy and Ramp Up to Algebra.

b) **Intervention: Especially intensive or personalized help for individual students.** Students who exhibit warning signs related to educational engagement — for example, very low attendance or very poor behavior — might require one-on-one counseling. Schools can target students for counseling based on data from the Early Warning System. (However, if high proportions of students exhibit such warning signs in some schools, the district might consider schoolwide interventions instead or in addition to one-on-one counseling. Moreover, some schoolwide reforms make it much easier to ensure that adults can provide individual support to students instead of simply expecting often overworked counselors to do so. See C below.)

c) **Prevention: School-level reforms that can reduce student risk factors and dropping out.** Providing interventions to students who develop risk factors can help improve graduation rates. But district leaders can also intervene on a
schoolwide level to create conditions that help prevent students from developing risk factors in the first place, and that reduce the negative impact of some schools on graduation rates. For example, are there middle schools and high schools where transition years are especially difficult for students? Do students who enter such schools exhibit big declines in academic performance, educational engagement or both? If academic performance, school or district leaders might consider curriculum changes, professional development, or carving out more time for math and literacy. If educational engagement, leaders might consider restructuring those grade levels into small learning communities, instituting adult advocate or mentor programs, restructuring schedules to allow teachers more time to interact in supportive ways with individual students, etc.
<table>
<thead>
<tr>
<th>Uses of Student- and School-Level Information Generated by an Early Warning Data System</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Why School Staff Members Need To Know</strong></td>
</tr>
<tr>
<td>1) Risk factors by individual student</td>
</tr>
<tr>
<td>2) Aggregate risk factors by school and type of school</td>
</tr>
<tr>
<td>3) Sizes and rates of declines in academic performance and educational engagement during transition years</td>
</tr>
<tr>
<td>4) School-level outcomes, raw and adjusted for student-level risk factors</td>
</tr>
<tr>
<td>• On track by HS grade</td>
</tr>
<tr>
<td>• Off track recovery rates</td>
</tr>
<tr>
<td>• Graduation rates</td>
</tr>
<tr>
<td>5) Systemwide analyses of student characteristics, risk factors and outcomes</td>
</tr>
</tbody>
</table>
Data produced by the system can have uses beyond direct intervention. For example, system leaders should strongly consider investing the resources to provide middle grades educators with reports showing how their graduates fare in high school — how many stayed on track as they progressed through high school and how many dropped out. The Consortium on Chicago School Research has offered detailed information about 8th grade graduates to Chicago’s elementary schools (Chicago uses a K–8 structure) since 1999. Student outcomes are broken out by various factors, including 8th grade test scores and passing rates in math and English. Finally — and perhaps most important — the reports employ compelling and easy to understand graphics to convey complicated information (see Figure 2).

Such reports might help bring middle school educators on board as systems seek to upgrade the mission of middle schools to focus on preparing all students for success in rigorous high school coursework. For example, surveys by the Southern Regional Education Board have found that fewer than half of middle school teachers in the region see their primary mission as preparing students for high school, and only 12 percent see it as preparing all students with the academic knowledge and skills needed to be successful in college-preparatory classes during high school.62
Note that the interventions and other data uses discussed above are not meant to represent an exhaustive list of solutions to the dropout problem, but only to provide several examples of interventions that can benefit highly from the kind of data system discussed in this paper. In large urban districts where risk factors are pervasive and more than half of the student population drop outs, for example, individual interventions and even aggressive institutional reforms might not be enough to adequately address the problem. Such systems might need to invest in large-scale, systemwide strategies as well, such as creating multiple institutional and non-institutional pathways to obtaining a diploma and a portfolio of flexible second-chance options for students who already have dropped out. More information on such strategies can be found in the companion JFF paper, *Improving Outcomes for Struggling and Out-of-School Youth*.

**Some Final Words of Advice**

✓ Build capacity to do this over the short term and long term. Small and rural districts with fewer resources and smaller, or nonexistent, research departments should consider working collaboratively to share the cost of research and development. Federally funded assistance centers, regional education service agencies and institutions of higher education all can help by providing...
technical expertise and advice. (Even systems with midsize to large research departments might want to consider partnering with researchers at nearby colleges and universities who are interested in the dropout problem.) And states can help by providing resources and by investing in statewide data systems that offer an informational backbone that local efforts can build upon.

✓ Districts should incorporate fields into the database that allow them to keep track of which students receive which interventions. That will allow them to evaluate how well the interventions are working, for whom and where.

✓ Districts should keep in mind that high-yield risk factors might change over time. For example, the demographic characteristics of a district’s student population could change, or the school board might enact policies that cause the risk factors or the timing of the risk factors to shift. However, once the electronic data system is in place, repeating the Phase I analysis will not be difficult or time consuming.

✓ Districts should add more information about students as it becomes available. For example, some school systems are using standardized quarterly benchmark tests to monitor student performance, and many of them maintain the results electronically. Failing to add new information to databases and analyses could mean missing out on identifying even more powerful predictors of dropping out.

And Some Future Actions To Consider

✓ School system leaders should consider analyzing other high school outcome variables in addition to promotion and graduation. For example, another goal of high school reform is to increase the proportion of graduates who complete a sequence of rigorous college prep courses and score at college-ready levels on state assessments. What factors predict whether students are more or less likely to do so? Can school systems create risk factors and on-track indicators related to college-and-work readiness, too?64

✓ Finally, state leaders should consider going “beyond the basics” to build a more expansive longitudinal data system. Some states have built P-16 databases that track individual students through the entire educational pipeline and even out into the workforce. The potential payoffs to such investments are tremendous. See Creating a Longitudinal Data System, available at www.achieve.org, for additional information.
PART V. CONCLUSION

Because of the rapidly changing American economy and a new commitment on the part of state leaders to raise graduation standards, solving the dropout problem has become more important than ever before.

At the same time, we know that dropout prevention efforts of the past have yielded poor results. Part of the problem is that such efforts did not provide at-risk students with effective interventions (see the companion JFF paper, Improving Outcomes for Struggling and Out-of-School Youth), but another reason is that they did not accurately target the right students to receive extra help in the first place.

Knowing which students are at greatest risk for dropping out and which schools most exacerbate the problem is the first step to reducing dropout rates. Fortunately, today’s education leaders have better research and data than were available 20 years ago.

If policymakers heed the most current research, avoid the mistakes of the past, and invest sufficient up-front “research and development” dollars, they can build data systems to identify a good many students on the path to dropping out early enough to make a difference. And district administrators can intervene in schools that contribute the most to the dropout problem, changing them from institutions that “push students out” into challenging and supportive environments that keep teenagers in school and on track for a diploma.
SELECTED BIBLIOGRAPHY


ENDNOTES


16 Croninger, R.G. & Lee, V.E. (2001). “Social Capital and Dropping Out of High School: Benefits to At-Risk Students of Teachers’ Support and Guidance.” Teachers College Record, 103, 548–581. Student/Teacher Relations was characterized as an aggregate measure of school social organization based on student-level standardized composite score. Student-level items were on a four-level agree/disagree scale, and include: Teachers are interested in students; teaching is good at this school; most teachers listen to me; when I work hard, teachers praise my effort; students get along well with teachers; and discipline is fair at the school.


23 This example was inspired by similar analyses in Roderick, M. (1993). *The Path to Dropping Out: Evidence for Intervention*. Westport, CT: Auburn House (p. 28).


27 Personal interview, December 14, 2005.


37 Balfanz, R. & Herzog, L. (March 2005). Keeping Middle Grades Students on Track to Graduation: Initial Analysis and Implications. Presentation given at the second Regional Middle Grades Symposium, Philadelphia, PA.


41 Personal interview, December 14, 2005.


46 Personal interview, December 14, 2005.


52 On the other hand, states and districts that have decided to use a four-year dropout definition for accountability purposes might want to pay special attention to that definition in the Phase I analysis for determining high-yield risk factors. The analysis might reveal that certain factors are better at predicting failure to graduate on time. Leaders in such systems can consider whether to give those factors primacy when they build their permanent early warning data system during Phase II.

53 Balfanz and Herzog examined three additional cohorts.

54 Personal interview, December 14, 2005.

55 E-mail communication, December 19, 2005.

56 Of course, whether policymakers would have “wasted” dollars providing interventions to such students is another question. Providing extra assistance to students who would otherwise graduate with low grades and skills might not be a waste of resources.

57 For comparison, if the intervention turned out to be 100 percent effective, it would decrease the dropout rate from 53 percent to 8 percent and increase the graduation rate from 47 percent to 92 percent. No single intervention ever proves that effective, however.
However, that does not mean that students who fail 9th grade should be viewed — or labeled — as “hopeless causes.” Researchers have shown that some students can and do recover, and that such resilience is related to both school and personal factors. See, for example, Catterall, J.S. (1998). “Risk and Resilience in Student Transition to High School.” *American Journal of Education, 106*, 302–333.


For example, many studies have shown that students who take 8th grade algebra are more likely to complete higher-level courses in high school, no matter their social background and prior achievement. And some states and districts have begun to use ACT’s EPAS system to monitor whether students in grades 8, 10 and 12 are on track for college readiness. (See http://www.act.org/epas/ for more information.)