

**Early Warning and Intervention Systems:
Promise and Challenges for Policy and Practice**

Draft-Oct 20, 2008

Prepared for National Academy of Education and National Research Council
Workshop on Improved Measurement of High School Dropout and Completion
Rates

Robert Balfanz
Center for Social Organization of Schools/Everyone Graduates Center
Johns Hopkins University

This paper reviews the work we have done to date on early warning and intervention systems, primarily in the middle grades. It then draws from our findings and experience to illuminate some critical issues and challenges states and school districts will need to address in order to be able to use on and off track indicators for high school graduation in effective and productive manners. The first section examines using off track indicators to build early warning systems. The second section looks at using early warning systems to support successful interventions to keep students on the path to high school graduation. The final section looks at how off track indicators might be used to identify which middle and high schools face the greatest educational challenges, distribute resources based on the degree of educational difficulty and evaluate the relative success of schools in meeting their educational challenges.

Using Off Track Indicators to Build Early Warning Systems

Foundational Studies and Developing Our Approach

Our exploration of off-track indicators-variables which signaled that absent intervention students have high odds of not graduating from high school- proceeded in stages. We began by examining in detail the educational demographics of students in the ninth grade in Philadelphia's non-selective neighborhood high schools. We wanted to establish the level of challenge different high schools faced in educating ninth grade students. This research indicated that most students in Philadelphia's neighborhood high schools appeared by common metrics to be in need of sustained intervention in order to succeed. Across the 21 neighborhood high schools most students in the 9th grade were either repeating the grade for the second or third time, over-age for grade, attended the 8th grade less than 80% of the time, or had math or reading skills below the 7th grade level. In the most challenged high schools less than 20% of the 9th graders were in the 9th grade for the first time, on-age, attended 8th grade at least 81% of the time and had math and reading skills at the 7th grade level or higher (if we set the bar at grade level skills the percent of students fell to single digits in the most impacted schools) (Neild & Balfanz 2006a).

This data showed that many high schools had extremely high degrees of challenges; up to four out of five ninth graders might need additional and sustained supports beyond those typically provided by high schools in order to succeed. Upon reflection we noted that some of these variables, like 8th grade attendance, middle grade test scores and age, were variables high schools in theory could know about students on day one, while others reflected behaviors that played out over the ninth grade. This led us to question whether a concise set of variables based primarily on middle grades data could be developed so school districts (and perhaps even the state and federal government) could know which high schools were going to need additional resources and of what type in order to meet the educational challenges they faced in the ninth grade. We also thought it would be useful for schools themselves to know who their students were on day one and what types and levels of supports they might need to succeed.

Finally, in a second research study we examined the characteristics of high school students who became involved with the juvenile justice system, in particular those who received out of home placements and were removed from high school. We observed that most of these students were incarcerated in the 9th grade and that the majority of them had a common set of 8th grade outcomes-failing half or more of their classes and attending schools less than two-thirds of the time. We also noted that there were an equally large number of 8th graders with these same characteristics who did not become involved in the juvenile justice system but who had the same graduation outcome as students that did. Namely students who failed half their classes in the middle grades and came less than two thirds of time had graduation rates of less than 10%, rather or not they were incarcerated in high school (Balfanz et al 2003). This led us to observe that clearly there could be 8th grade indicators which were highly predictive of graduation outcomes.

Based on these two studies we formulated our next studies in which we asked-how early could we reliably identify students who absent intervention likely would not graduate. In preparing for this we read and strongly agreed with Gleason and Dynarski (2002) on the need for indicators that were both reliable and valid and collectively had a high yield-meaning that students with the indicators had high odds of not graduating (absent intervention) and that a significant number of likely dropouts could be identified with them. In addition we believed it would be important for future utility for the variables to be based on data that was readily knowable and accessible at the school level. Finally, based on our reading of the youth development literature and our own experience working on school reform in high poverty middle grades schools we decided to focus our inquiry on the middle grades. Our operating theory based on the literature and our experience was that the interaction of early adolescence, living in neighborhoods of concentrated poverty, and attending an urban, high poverty middle schools produced a number of factors and forces which could promote student disengagement from school and result in large numbers of students becoming off-track for high school graduation (Balfanz et al 2007).

This then led to a refinement of our question, which became “how early in the middle grades can we reliably identify students who, absent intervention likely will not graduate, using data that is readily accessible and knowable to schools?” In practice this meant we would use school records data, rather than surveys to establish the indicators. We also establish some a priori definitions of reliability and high yield. We set a decision rule that we wanted to find indicators which signaled that presently 75% of students with the indicator did not graduate from the Philadelphia school district within one year of on-time graduation. We wanted the indicators to at least collectively identify 25% or more of future non-graduates and we would only include indicators in the final set that had independent and additive effects on a student’s odds of graduating. Following the work of Rumberger and others who have shown that transfers are sometimes simply delayed dropouts, we decided to initially focus on non-graduates rather than dropouts. A second reason for this decision was our experience with data from large urban districts and the difficulty of accurately identifying dropouts, in particular interpreting non-drop withdrawal codes like “whereabouts unknown”. We conducted secondary analysis to demonstrate that the indicators were not biased by transfers and in later studies conducted analysis with and without transfers. .

Initial Findings from Philadelphia

Using our decision rules we examined a wide range of school performance and status variables, selected based on prior research and our experience in schools. We did this by conducting a cohort analysis of all students enrolled in the sixth grade in the School District of Philadelphia in 1995-96-following them through to one year past on-time graduation in 2003-04. For each variable under examination we constructed a progression table which showed student's progression grade by grade through the school district from 1995-96 till 2003-04. In short this let us show for every 100 sixth graders in 1995-96-with a particular school performance or status variable (i.e. poor attendance or in special education) what grade they were in for each of the following years, how many ultimately received a diploma (within one year of on time graduation), how many were still enrolled in the school district in what grade, and how many had left the school system.

We then used our decision rule of looking for variable in which at least 75% of students with the indicator do not graduate within one year of on time to fine tune the potential indicators. So for example, we learned to reach the 75% threshold attendance in the sixth grade had to be below 80%. This does not mean that students who attended 81 to 89% of the time graduated in high number, but more than 25% did. It was only when the attendance bar was lowered to 80% that the 75% non-graduation rate threshold was achieved. Similar approaches were used with other variables-so for example-failing math or English in 6th grade reached the 75% threshold but failing any one class or getting a D did not, even though students with one failure and with D's did not graduate in large numbers.

Finally, we used logistic regression to identify indicators that both met our decision rules and had independent and additive impacts on yield (i.e. the percent of future non-grads that could be identified in this case in sixth grade). This identified our final set of variables or off-track indicators. They were:

- Attending Less than 80% of the time
- Receiving a Poor Final Behavior mark in a course
- Failing Math
- Failing English

Less than 25% of sixth graders with any one of these indicators in the 1995-96 cohort graduated from the school district of Philadelphia within one extra year of on-time graduation. Collectively these four indicators identified forty percent of the students in the cohort of over 14,000 sixth graders who did not earn a diploma from the school district (Balfanz et al 2007).

Follow Up Studies in Philadelphia and Replication Studies in Other Cities

Since the initial study we have conducted follow up studies in Philadelphia and conducted similar analysis (either directly or via consulting with local school districts and

research groups) in six additional school districts-Boston, Indianapolis, Mobile, Pueblo Col, and Baltimore (Balfanz & Boccanfuso 2008a, 2008b).

Across the follow up and replication studies the basic finding has held. We were able to identify between 40 and 50% of eventual non-grads or dropouts (in some of the replications we removed transfers from the sample) from sixth grade cohorts. Indicators based on attendance, behavior, and course performance were highly predictive at the 75% level or higher.

Working with a range of data sets and answering an assortment of school district queries about the data and indicators further help to illuminate important points regarding, individual variables, interactions between variables, and correlations or substitutions between variables.

Attendance

Across the different districts highlighted we did not find a common attendance threshold for establishing a highly predictive and high yield indicator. In Philadelphia students had to attend school less than 80% of the time in sixth grade to have only a 1 in 4 chance in graduating from the school district. In Boston the same outcome was obtained for students who attended school less than 90% of time. This led us to speculate that part of the impact of attendance on graduation odds is not just missing a certain number of days but also being in the lower part of the attendance distribution. So it may be both how many days you miss and how many days you miss relative to other students. But the key takeaway finding is that it may be important for each school district to examine its attendance distribution and its impact on graduation odds in order to determine where key thresholds are.

Behavior

Philadelphia, it turns out, had a relatively rare practice when we examined its sixth grade cohort. Middle grade students were given behavior marks, as well as academic grade in each course. In each course, during each grading period, students received either an excellent, satisfactory, or unsatisfactory behavior mark. The behavior grade for each marking period was then used to establish a final behavior grade for the course. We found that receiving a final unsatisfactory behavior grade in a core course in sixth grade was highly predictive of not graduating from the school district. Because of its strong predictive power and because many more students received poor behavior grades than were suspended and very few students who were suspended did not receive at least one poor behavior grade, we ultimately dropped suspensions from our final model because including it did not significantly add any more predictive power and a poor final behavior grade produced a much higher yield of future non-graduates. Philadelphia has subsequently dropped behavior grades and replaced them with a series of numerically coded comments and in all the other districts we have since worked with the only behavior measure available has been in and out of school suspensions. Absent behavior grades, suspensions have typically proven to be predictive of not graduating, though the number and type needed to reach our 75% threshold has varied. Our Philadelphia work, however, indicates, the absence of a measure of mild, sustained misbehavior (and/or a

teacher's perception that this is the case) may mean that a subset of kids who could be identifiable as likely not to graduate absent intervention may be missed when suspensions are the only behavior variable available.

Course Performance

Consistently across the districts we have looked at, we have found that course failure is the outcome which is highly predictive and produces a high yield of future non-graduates or dropouts. In most cases, we have not had sufficient information to create consistent grade point averages but we have found that while students who gets D's but not F's have lower graduation rates than students with higher grades, D's not produce the same consistent 75% or more of students not graduating, as do F's. We speculate that in part this might reflect that it is relatively difficult to fail in the middle grades, we have observed that teachers tend to give students at least some benefit of the doubt and factor effort and improvement into final grades. Thus students who receive F's maybe students who are literally failing all their tests and/or consistently not turning in assignments and are not viewed as either trying sufficiently hard or seeking extra help opportunities. In short, these are students who, absent sustained intervention, are likely to continue failing courses and hence not graduate.

Test Scores

In the original Philadelphia analysis, once we controlled for attendance, behavior, and course failure test scores were not significant predictors of non-graduation. In other districts, however, we have found that when significant numbers of students had very low test scores-below the 15th percentile on nationally normed tests-then very low test scores did add some predictive power to the models. It should be noted, however, that the district in which test scores had the strongest relationship was the school district in which we had the most minimal course passing data-and had to construct a composite measure that was based on failing any course in the middle grades 6-8, not just in the sixth grade.

Overage Age for Grade

Much like test scores, being overage for grade did not add predictive power in our Philadelphia models once attendance, behavior, and course performance was controlled for. In two other districts, however, being two or more years overage-in the middle grades-did add predictive power. Again this may in part have to do with the fact that we looked across the entire middle grades experience not just the sixth grade in these districts and that these districts had policies in place which led to relatively larger numbers of students being two or more years overage in the middle grades. So it maybe that overage for grade only has independent effects when there are large numbers of middle grade students who are two or more years overage.

Special Education and ELL Students

Across the school districts we found that once attendance, behavior, and course performance was controlled for, special education and ELL status had no predictive value. It should be noted, however, that we found special education coding in particular to often be incomplete or hard to comprehend in district data files.

Grade of Onset of Off-Track Indicators

The most common question we got asked by the school districts we were working with and one that interested us deeply as well was whether the grade in which a student first exhibited an off-track indicator impacted their graduation odds. Because of data and resource limitations we were not able to explore this question in all the districts we have worked with. Analysis of both Philadelphia and Boston data, where we were able to look at it, thus can only be considered speculative at this point. What both the Boston and Philadelphia data indicated was that the earlier in the middle grades or high school that a student exhibited poor attendance, behavior, or course failure, the lower their odds of graduating from the school district. In Philadelphia, for example, we needed to lower attendance to 70% in 9th grade to reach the same predictive yield (i.e. 75% of student with the characteristic not graduating) as 80% or lower attendance in sixth grade (Neild and Balfanz 2006b). In Boston, students who had an off-track indicator in ninth grade, but did not have one in the middle grades graduated at nearly twice the rate, students with off-track indicators in sixth grade.

The impact of off-track behaviors prior to sixth grade was also only observable in two school districts. In both Baltimore and Philadelphia we observed that students who exhibited poor attendance, behavior, or course performance in the sixth grade did worse than students who exhibited these outcomes in the fifth grade but not the sixth grade. Students who exhibited these outcomes in fifth and sixth grade did about the same as students who exhibited them for the first time in sixth grade. The biggest difference, overall, however, was that many more students exhibited poor attendance, behavior, and course performance in sixth grade with the transition to middle schools than did students in the 5th grade who were still in elementary schools. Although most of the districts we studied had a combination of K-8 and middle schools, we did not have sufficient data to conduct a good comparison of outcomes across these school types.

Number of indicators

A consistent finding across the school districts was that in the sixth grade and wherever we examined it throughout the middle grades most students had only one or two off track indicators. However, students who had multiple indicators had extremely poor graduation outcomes. In some cases, only a few percent of students who, for example, had poor attendance, got suspended and were failing a course or two in sixth grade ultimately graduated. The norm was dropout for students who were failing either math or English or at most two courses but not more and had either poor attendance or poor behavior.

Implications for Building State and District Data Systems

Our work with middle grades off track indicators has several potential impacts for building state and district data systems. Our most common experience in working with a half a dozen of school districts is that currently the necessary data is often housed in multiple locations and sometimes in different data systems within school districts. Although all of the key data is first collected in schools-they are usually the ones recording students' attendance, behavior infractions, and course grades-to create and examine off track indicators we needed to work with school systems to pull data from demographic files, attendance files, course grade files, and test files and then merge using student ID numbers. In fact, most of the time spend working with the districts is spent preparing the data set to be analyzed. The other challenge was that in to create historical cohorts, we often had to work across a change in data systems within the school district during the time period being analyzed.

In order to facilitate the development and fine tuning of early warning systems, state data systems will need to be able to capture students' grades, behavior measures, and attendance along with basic demographic, test score, and status variables (i.e. special education, ELL etc.). Since different districts often have different report card intervals at the state level it makes sense to focus on final grades. Also, it may be sufficient to focus on core academic courses. It will be important to have notations for length of course i.e. full year versus semester, and to note rather it is a core course or an extra help course. Since the full potential of using course grades for both on and off track indicators has not been fully explored and it may well be possible that in different setting D's become highly predictive it would not wise to only notate course failures. Rather, the data system should collect all final grades in core academic courses. Attendance data needs to be reported so categories of absences can be created i.e. missing 5 or fewer days, 10-19, 20-29, 30 or more etc, not just overall percents of days attended (without providing information on days enrolled). Behavior will be the biggest challenge since suspensions are the only common behavior measure available and policies over their use vary across districts. Thus, it may be important to at least create a depository of district suspension policies so that difference across districts in predictive outcome can be better understood.

Finally, it is very important to consider that to date, all the analyses we have been involved in have been with high poverty school districts. If concentrated poverty has an interactive impact with school performance in the middle grades in general and the sixth grade in particular in determining graduation outcomes, as we surmise it might, then it is important not to generalize our findings beyond high poverty, urban environments. Analysis of on and off track indicators at the state level is required to gain a better sense of how predictive reliability, validity, and yield may vary across urban, suburban, and rural, higher and lower poverty school districts.

Using Early Warning Systems to Enable More Effective Interventions

The idea of on and off-track indicators and early warning systems for potential high school dropouts is gaining notice. The state of Louisiana is working to spread indicator systems in the 9th grade. The Alliance for Excellent Education has recently

released a policy brief on early warning systems (Pinkus 2008). The National High School center has unveiled an on-line data gathering system for high school indicators. The school districts of Chicago and Philadelphia are moving to incorporate on and off track indicators throughout their school district data and accountability systems. Recent federal legislation has been introduced which uses attendance, behavior and course performance indicators to track the progress of reform efforts in the middle grades and high school. The real potential of early warning systems, however, will not be realized unless they are used to develop, provide, and target effective interventions at school and student level. This is where much of our work is currently focused.

Philadelphia Pilot and Demonstration Efforts

For the past two years we have been working on a pilot and demonstration effort with two Philadelphia middle grades schools, as well as consulting with similar efforts in New York City and Jacksonville, Florida. At the heart of this work has been an effort to merge early warning systems with an adaptation of a response to intervention model and the multi-tiered prevention and intervention systems often found in public health efforts (Balfanz et al 2007). The goal is to create a system which enables schools to get the right intervention, to the right students, at the right time with the required intensity. Teams of teachers use what might be called early, early indicators to try to identify students as close to the moment when they begin to fall off the path to graduation as possible. These include metrics like missing three or more days of school in a month, getting two or more office referrals for behavior in a month, or failing two or more quizzes, not turning in two homework assignments or failing a major test or project. The goal is to prevent as many students as possible from reaching off-track indicator status. Students who do reach this status, despite prevention efforts, then signal themselves as students in need of more intensive, often case managed interventions. Borrowing from the response to intervention methods typically used in elementary grades, a core idea behind the keeping students on the graduation approach is to reserve the most intensive and hence usually most expensive interventions only for students for whom less intensive supports have not proven effective.

Initially we worked with the pilot schools to weave and organize existing student supports into a more coherent and organized system. We observed that many of the support programs in the schools did not communicate with each other or were unknown to many teachers in the building. For each of the key off-track domains we worked to identify school-wide prevention, moderate intensity supports targeted at small groups of students, and intensive case managed supports for the neediest students. We also worked with the school faculty and administration to develop decision rules around when a student would move from one level of intervention to the next. During this period, our local partner the Philadelphia Education Fund also worked closely with the school district and the pilot schools to integrate early warning system and intervention data into the school district's Schoolnet data system. The result is that in pilot schools and soon more broadly across the school district, teachers can pull down classroom level data on key off track and intervention indicators (see Appendix A for an example).

The school district system could only accommodate ten fields which led to in depth discussion with teachers about ultimately what data would be most useful in

designing and targeting interventions. The consensus was teachers wanted to see prior year attendance, behavior, and course performance, and the same data for the most recent marking period for course performance and behavior, with year to date data for attendance (continuously updated). In addition teachers wanted to see test score levels in mathematics and reading and most recent reading level data (in Philadelphia students reading levels are tested several times a year). One informative outcome of pairing test score and reading level data with off track indicators is that it highlighted a number of students who had high test scores but one of more off track indicators, as well as students who appeared to have very low test scores but were doing ok in their courses, coming everyday and behaving. These stark juxtapositions led to teachers thinking about interventions in different ways (for example, it indicated that not every student failing a class may be in need of tutoring, especially if its targeted at basic skill development).

We also learned that while at first the schools seem to have a wide range of interventions and programs on paper in their school improvement plans, when the actual capacity and/or success of the interventions and programs were examined significant holes in the schools response system to off track behaviors became apparent. In part this was driven by the size of the school, it has over 800 students in grades six to eight, and many of the students it served were high needs students. This resulted in hundreds of students exhibiting behaviors-e.g. missing two days of school in a month, receiving two office referrals, or failing two quizzes-which required moderate intensity, targeted supports. Consequently, it became clear that more adults would need to be brought into the building to be able to meet the scale of student support demands that existed.

This year the pilot school with a large student body is using its early warning and early, early warning indicators to help assign students supports from two external organizations that have been brought in to provide the necessary person power to reach all the students in need. City Year, a national service organization is providing a team of fifteen seventeen to twenty-four year olds who are at the school four days a week from before it opens through an after school program they run. City Year corps members are assigned to homerooms and follow their classroom of student throughout the school day. Within each homeroom the corps members are given a target group of 8 to 12 students identified by the early warning system as students in need of moderate intensity supports including, attendance monitoring, homework support, assignment review, mentoring, after school enrichment and support activities and sometimes behavior contracts. Because the corps members spend the whole day with the class they know instantly if the student is not in school and can call to see where they are, they can see if they need a gentle reminder to pay attention or be respectful in class, they can check that they understand their assignments and if they need help completing their homework after school, the corps member knows what their assignment is and can support them. Collectively the corps is able to reach over 200 students. Additional case managed intensive supports is provided by Communities in School which provides both an additional social worker and site manager who works to connect students in need with external support providers and monitors the effectiveness of the interventions.

A final enhancement this year, is more formalized bi-weekly grade group meetings built into the school schedule which enables the teams of teachers which share common students, the City Year corps members who support them, the Community in Schools social worker, and a wide range school support personnel to check every two weeks if

students who are receiving supports are improving, showing signs of improvement, or staying the same or getting worse. The team then decides if the existing intervention should be continued for two more weeks, modified, or if another more intense or alternative intervention should be tried. A coding system has been developed to enable teachers in the main to have conversations about students, with only one member tasked to fill out documentation. The use of common codes will also enable quarterly reviews of the global effectiveness of the interventions, enabling the school determine over time, for example, when, under what conditions, and for what types of students mentoring or behavior contracts turn out to be effective interventions.

Work in Other Cities

Over the past year middle grade and high schools in several cities have piloted themes and variations of a multi-tiered response system linked to early warning indicators. One element which has emerged from these efforts is how status variables-e.g. over-age, special education status, ELL etc-can provide information which supports effective interventions. While, the general finding of off-track indicators to date has been that school performance variables are much more predictive than status variables, it appears that status variables can help schools interpret and target effective responses to off track indicators. For example, one middle school in the Bronx observed that nearly all their students requiring intensive behavior supports where two or more years over-age for grade. This led them to hypothesize that the only effective intervention for these students was to provide them a path to catch up and rejoin the peers. They developed a highly supported, intensive acceleration program that used extend day to enable students to cover two grades in one year. Another school found that most of its most chronically absent students were ELL students which led them to investigate the factors behind this and to devise targeted solutions based on their findings.

Implications for Policy and Practice

The emerging efforts to use early warning systems to support more effective interventions have several implications for policy and practice. First, the strategy of using early, early indicators to prevent students from reaching off track status (e.g. missing more than two days in month, getting two or more office referrals, failing two quizzes) suggests a possible middle ground between waiting for fully operational state wide early warning systems that enable classroom level analysis and having each school try to design its own system. States or districts could focus on establishing valid and reliable high yield predictors for different school populations and then provide updates to schools at least every marking period. Schools could then use simple early, early warning indicators to identify students in need of additional support. The ultimate effectiveness of the approaches used by the school could then be established by tracking the extent to which the number of students reaching off-track status declines across marking periods.

The second and perhaps most significant learning which emerges is that considerable effort will need to be invested in training and supporting teachers to use off track indicators to develop better interventions and more effectively target them. This

will include both mission building and cultural change work as teachers will need to be re-orientated from an outlook which says their primary job is to get their students to meet grade level objective and pass to the next grade, to seeing part of their responsibility as keeping students more globally on track to high school graduation. In practical terms, the challenge is to create the conditions and supports and rationales that promotes rapid intervention after the first signs that students are in danger of falling off track, rather than assuming or hoping that what might be seen as relatively small struggles-missing more than two days of school in a month, getting sent to the office few times, or failing a quiz or two will self correct. Support and efforts will also have to be provided to enable teachers and administrators to take a more proactive role in evaluating the effectiveness of interventions. Often the default response is to wait until the end of the year or later to examine rather than an intervention worked and then typically since it was one of many things going on, to attribute success or failure to everything that was tried that year.

Finally, the utility of status variables (i.e. age, special education status, ELL etc) in designing intervention indicates that this information should be accessible in early warning systems. Moreover, it suggests that states should take the lead in integrating information for agencies which work with youth outside of the school system. In particular the agencies which oversee juvenile justice, foster care, and child protection (abuse and neglect). It will be useful for both the school system and the agencies to know the extent to which students involved with the agency exhibit off-track behaviors and the extent to which students demonstrate these behaviors prior to involvement with the agency. This could lead to more integrated and effective student supports for students who are agency involved. It could also potentially provide the basis for the agencies to help fund and hold accountable school based prevention efforts designed to keep students from exhibiting off-track behaviors. In Philadelphia we were able to work with an integrated data set and it revealed three important facts. First, the students involved with the juvenile justice, foster care, and child protection agencies, as well as females who gave birth during high school had very high dropout rates, ranging from 70 to 90%. Second, although these outcomes were very bad, agency involved youth only accounted for a relatively small portion of all dropouts. Finally, we found that two-thirds of the males that would be incarcerated in high school and two-thirds of the females who would give birth in high school had an off-track indicator in sixth grade (Neild and Balfanz 2006b).

Using On and Off Track Indicators to Target School Support, Resources, and Accountability Systems

To date most of the emerging work on early warning and intervention systems has focused on student supports. It is possible, however, to use on and off-track indicators at a more aggregate level to support schools, as well as to identify those who are performing well and those that are not. This in essence gets us back to the original question we began with can we identify a school's degree of educational challenge? At some level the most basic analogy is Olympic sports like diving, in which a participant's total score is combination of technical skill and the degree of difficulty of the dive they attempted. Currently most funding, staffing, and accountability systems assume that all schools face the same degree of difficulty and implicitly only focus on technical skill-(i.e. did you

make AYP?) Even the exceptions including Title 1 funding and most of the emerging weighted student funding formulas operate using status, rather than school performance or behavior variables to distribute additional funds or differentially fund schools. Title 1 money is allotted based on the number of low income students in the school and weighted student funding formulas might take into account the number of special education, ELL, and perhaps even overage students as well. Finally, some systems consider the number of students with low or high test scores. This approach in essence is analogous to the dropout predictor strategies based on status variables which has been found to be not highly predictive of who dropouts out and hence not highly effective in targeting supports to the students who need them (Gleason & Dynarski 2002).

One could imagine funding and staffing formulas, particularly at the high school level, that take into account the number of students who enter the school having exhibited off-track behaviors in the middle grades. As we have shown there will be wide differences across high schools in some the vast majority of students enter high school, essentially already off-track, whereas in others is could be only 10 to 15% of entering students. The most sophisticated systems might even factor in the types, intensity, and concentration of on and off-track behaviors among students and tailor the additional resources large numbers and concentrations of students with middle off-track behaviors would bring to strategies and interventions which have proven effective in reversing off-track trajectories in high schools.

There might also be a role for on and off track indicators in accountability systems across middle and high schools and potentially even elementary schools, analogous to value added or growth modeling for test scores. Graduation rates are a key and perhaps the central outcome of high schools, but given the central importance of the ninth grade year in determining student's graduation outcomes an effective reform will likely take at least four years to produce a substantial impact on graduation rates. Currently and in the future, however, accountability systems don't work on four year cycles. If states could verify a set of on and off track indicators which work across school populations in their state or identify sub-sets of indicators for different populations, then these indicators could be used on an annual basis to demonstrate if a school is likely making progress towards raising graduation rates. In simplest terms if half the students in a school continue to have poor attendance, fail their courses, and not earn promotion to the next grade, it is highly unlikely that the school will see a large increase in graduation rates.

Conclusions

Over the last five years we have been able to answer many of the questions we initially had. It is possible to identify the degree of educational challenge different high schools face in educating their ninth graders. It also possible to identify significant proportions of future non-graduates or dropouts as early as the sixth grade using data readily available to schools. The variables that appear to have high reliability and validity (in high poverty, urban areas) and yield significant proportions of future non-grads and dropouts are school performance variables-student attendance, behavior, and course performance, or perhaps ironically the ABC's. In order for these off-track indicators to

have a positive impact on student outcomes they need to be effectively linked to successful intervention systems. This will require considerable investment in teacher and administrator training and support, as it will necessitate a sea of change in how teachers and administrators define their role with students in many schools. One factor working in favor of school level implementation is that school's can use readily available data-essentially the information in their grade books, in their daily attendance sheets, and in their office referrals to develop early, early indicators which signal that a student may be headed to off-track behaviors and use this information to intervene rapidly when students first begin to show signs of off track behaviors. Over time, however, state departments of education will need to make investments in good data systems that will both validated on and off track indicators for different student populations in the state, and provide teachers with easily accessible classroom level data on these behaviors and outcomes. Finally, to have a positive impact resources and staffing supports will have to follow the outcomes of early warning systems. Fundamentally, in order to effectively respond to the information provided by early warning systems schools will need the right ratio of skilled and committed adults to students in need.

References

- Balfanz, R., K. Spridakis, R. Neild and N. Legters. 2003. "High-Poverty Secondary Schools and the Juvenile Justice System: How neither Helps the Other and How that Could Change". *New Directions for Youth Development* 99(fall): 71-89.
- Balfanz, R., L. Herzog and D. MacIver. 2007. "Preventing Student Disengagement and Keeping Students on the Graduation Path in Urban Middle-Grades Schools: Early Identification and Effective Interventions". *Educational Psychologist* 42(4): 223-235.
- Balfanz, R. and C. Boccanfuso. 2008(a). "Falling off the Path to Graduation: Middle Grade Indicators in Boston". Working Paper.
- Balfanz, R. and C. Boccanfuso. 2008(b). "Falling off the Path to Graduation: Middle Grade Indicators in Indianapolis". Working Paper.
- Gleason, P. and M. Dynarski. 2002. "Do we Know Whom to Serve? Issues in Using Risk Factors to Identify Dropouts". *Journal of Education for Students Placed at Risk* 7(1): 25-41.
- Neild, R. and R. Balfanz. 2006a. "An Extreme Degree of Difficulty: The Educational Demographics of the Ninth Grade in an Urban School System." *Journal of Education for Students Placed at Risk* 11(2): 123-41.
- Neild, R. and R. Balfanz. 2006b. *Unfulfilled Promises: the Dimensions and Characteristics of Philadelphia's Dropout Crisis, 2000-2005*. Philadelphia: Philadelphia Youth Network.
- Pinkus, L. 2008. *Using Early-Warning Data to Improve Graduation Rates: Closing Cracks in the Education System* (Policy Brief). Washington, DC: Alliance for Excellent Education.

Appendix A: Sample of Classroom Level Data on Key off Track Indicators

Sample Class Feltonville School of Arts and Sciences

Student Name	2007-2008: Days Absent	2008-2009: Days Absent	Negative Behavior Comments	Math Grade 3/1/2008	Math Grade 6/1/2008	Literacy Grade 3/1/2008	Literacy Grade 6/1/2008	Reading Level 6/1/08	PSSA 2008 Math	PSSA 2008 Reading
Student A	53	0	10	D	D	F	F	5	Proficient	Basic
Student B	36	2	7	D	D	D	D	6	Basic	Basic
Student C	14	0	1	C	B	C	C	6.5	Basic	Proficient
Student D	5	1	6	C	B	D	C	7	Basic	Basic
Student E	18	0	7	C	C	D	F	5.5	Below Basic	Below Basic
Student F	29	2	1	D	C	D	D	6	Below Basic	Below Basic
Student G	6	0	8	D	D	F	D	5.5	Below Basic	Below Basic
Student H	46	2	3	B	B	D	F	5.5	Basic	Below Basic
Student I	41	0	4	D	C	D	D	3.5	Below Basic	Below Basic
Student J	17	0	1	B	B	C	D	2	Below Basic	Below Basic
Student K	61	4	7	C	F	D	C	7	Below Basic	Basic
Student L	24	0	10	F	F	C	D	6.5	Below Basic	Basic
Student M	18	0	2	B	D	D	C	3.5	Below Basic	Below Basic
Student N	3	0	6	B	B	B	C	7	Basic	Basic
Student O	2	1	5	C	D	D	D	5.5	Basic	Basic
Student P	15	1	4	D	D	F	D	5.5	Basic	Below Basic
Student Q	15	1	10	C	D	D	D	6.5	Below Basic	Below Basic
Student R	6	0	1	D	D	D	D	3	Below Basic	Below Basic
Student S	16	1	4	D	D	D	D	5	Below Basic	Below Basic
Student T	15	0	7	C	F	D	D	6	Below Basic	Basic
Student U	18	0	6	C	D	D	D	6.5	Below Basic	Below Basic
Student V	23	0	7	C	F	C	F	6	Below Basic	Below Basic
Student X	16	0	6	C	F	D	D	6.5	Basic	Basic
Student Y	18	1	3	B	C	D	D	6.5	Basic	Basic
Student Z	4	0	7	C	C	D	D	6.5	Proficient	Below Basic
Student AA	42	2	1	D	C	D	D	5.5	Below Basic	Below Basic
Student AB	13	0	2	D	D	D	C	4	Below Basic	Below Basic
Student AC	8	0	2	D	D	D	D	2	Below Basic	Below Basic
Student AD	22	1	8	C	F	D	D	6	Below Basic	Below Basic
Student AE	50	1	0	D	D	C	C	4.5	Below Basic	Below Basic
Student AF	18	0	6	C	C	F	D	5	Below Basic	Below Basic
Student AG	1	0	3	NG	D	NG	D	6	Below Basic	Basic