Data Driven Identification and Selection Algorithms for At-Risk Students Likely to Benefit from High School Academic Support Services

A paper presentation
during the

2011 Annual Conference of the
American Educational Research Association

on

April 9, 2011
New Orleans, Louisiana

by

Warren E. Lacefield
E. Brooks Applegate
Pamela J. Zeller
Nancy Van Kannel-Ray
Shelly Carpenter

Western Michigan University
Kalamazoo, Michigan

Data Driven Identification and Selection Algorithms for At-Risk Students Likely to Benefit from High School Academic Support Services

April 9, 2011

Warren E. Lacefield, E. Brooks Applegate,
Pamela J. Zeller, Nancy Van Kannel-Ray
and Shelly Carpenter
Western Michigan University

ABSTRACT

This study describes a well-defined data-driven diagnostic identification and selection procedure for choosing students at-risk of academic failure for appropriate academic support services. This algorithmic procedure has been validated both by historical quantitative studies of student precedents and outcomes as well as by current qualitative comparisons with existing school procedures and efforts to accomplish the same goal through committee work and recommendations. Results indicate it is both possible and feasible using readily available school student information system data to identify who appears to be at substantial academic risk, what some of those risks are, and who appears likely to benefit from specific academic support service interventions.

OBJECTIVES OR PURPOSES OF STUDY

Nationally, graduation rates for students in urban high schools are at 53 percent. In 2008, more than one-quarter million students did not receive a diploma (Swanson, 2003, 2008). Studies show retention problems may begin in middle school (Jackson & Davis, 2000). Therefore, as part of a major GEAR UP intervention serving almost 3,500 students, we developed academic case management programs targeting students in middle grades (Kretovics, 2005-2011). Since these programs were effective at reversing downward trajectories of at-risk children on track for failure in middle schools (Van Kannel-Ray, Lacefield, & Zeller, 2008, 2009), this model was revised, as graduation coaching, incorporating strategies appropriate to developmental needs of students progressing into high school.

A subsequent study following students through 10th grade in a large urban high school demonstrated the effectiveness of graduation coaching (Lacefield, Zeller, & Van Kannel-Ray, 2010). The goal was increased academic performance and improved retention rate. Results indicated some at-risk students receiving graduation coaching performed better than expected academically and appeared more likely to remain in school. However, both the middle and high school programs suffered from a problem affecting many other similar support service interventions. Specifically, significant difficulties associate with attempts to diagnose individual student weaknesses and identify those who may be at-risk of academic failure and who may actually benefit from targeted coaching or other academic support programs.

This study describes a well-defined data-driven diagnostic identification and selection procedure we have developed for our projects during the 2010-2011 academic year; one validated by historical quantitative studies of student precedents and outcomes and by current qualitative comparisons with existing school procedures and efforts to accomplish the same goal through committee work and recommendations.

PERSPECTIVES, THEORETICAL FRAMEWORK, AND CONNECTIONS TO LITERATURE

Successful schools have personnel who use data to help students achieve educational success (Armstrong, J. & Anthes, K., 2001). It becomes prudent to refine interventions such as graduation coaching by not only collecting but also using data to inform decisions about caseloads. As Bernhardt (2009, p. 27)) suggests, it is important to "look for leverage points" to refine the data used to improve a practice. Because student demographics have become diverse and show gaps between low and high achievers, informed educational decisions through the use of data are critical (Cromey, A, van der Ploeg, Masini, B, 2000).

We focus here on difficulties faced by school-based personnel who have responsibility for identifying students who might benefit from supplemental services and then providing those services to that group over extended time periods. Individual student services are expensive. Our experience across multiple districts and school sites has been that our case managers, coaches, and other intervention personnel based in-school find themselves often overwhelmed with "off-caseload" duties. Using daily activity report logs, we found one-third of their time was cafeteria and hallway supervision, school bus and pickup line duty, standardized test monitoring, truancy detection and follow-up, and other similar duties.

More importantly, we relied upon experience and training of principals, teachers, counselors, and coaches themselves to be able to select and serve those students most in need of educational supports and attention. However, when we compared actual caseloads of students being served to our own data-driven identification system results, we observed that the actual caseloads included many students we would not classify as at-risk based upon their past academic record. We needed to better understand the selection criteria being used for defining intervention caseloads.

It is legitimate to ask why this happens. Our coaches were employed by the schools and tried to do good jobs. They had full caseloads but still had time budgeted to participate in the school community, hopefully by talking to teachers and being good team players. But the data shows requests for essentially administrative assistance began to consume their resources excessively and they had no way to contest this. Teachers came to coaches frequently with intermittent daily student problems: e.g., "Johnny is missing 3 assignments and may do poorly this marking period. Can you help him?" Retrospective analysis of caseload logs strongly suggested that this was one way so many "successful" students got added to caseloads.

Without clear evidence to prove otherwise, coaches and teachers must rely on their best judgments or the directions given them by others. Obtaining that evidence is very difficult sometimes, even in an environment where performance data abounds and could be readily available. This paper describes a quantitative method and a practical procedure for aggregating available data, making responsible decisions regarding populations of students to be served, selecting students who may benefit from interventions, and monitoring their progress during the intervention, in such a way so that students who improve receive the most attention and students who do not are replaced by others on waiting lists for caseload services.

BACKGROUND AND METHODOLOGY

Study Site and Sample: This study is part of a five-year longitudinal project focusing on multiple factors including academic performance and student behavior as a result of academic support interventions. Located in rural southeast Michigan, the target district examined here currently enrolls 278 students in middle school and 441 in high school, 64% of whom qualify for free/reduced lunch. About 38% are from minority groups. Academic performance data (grades) are available for this study from 2004-2005 through 1st quarter, 2010-2011.[1] Since this study covers multiple years, we define a cohort according to the year when a student was in 8th grade. Thus, this study covers 3 student cohorts (N=310): those in 8th grade in 2007 (N=104), 2008 (N=114), and 2009 (N=92). A new cohort of 8th grade students in 2010 has now entered the 9th grade and high school in 2011 and have completed their first quarter of study at the time of this writing. These (N=106) students have not previously been a part of the GEAR UP project cohort and have not received direct GEAR UP services. They have, however, benefitted from sustained GEAR UP partnership with their schools and school districts.

Key Questions: Student grade changes in core courses from one grading period to the next over each academic year are of particular interest for determining eligibility and need-based interventions. Core courses include English/language arts, science, math, and history/social studies. Of primary interest is whether students who are likely to be at-risk of academic failure can be clearly identified? Second, can various types of "risk" be identified and students be so classified? Third, can this method of identification and classification be extended to specific core academic content areas for further diagnostic and selection purposes? Fourth, can this kind of information be organized such that educators and school personnel can make clear decisions regarding allocation of scarce and expensive academic support services?

Data Sources, Selection, Analytical Methods, and Use of Evidence to Support Conclusions: This is not an experimental or theoretical study. Rather it is a retrospective study and a demonstration of a practical and to some extent heuristic procedure. The historical comparison study; content, concurrent, and predictive validity studies; and the current intervention study share a similar overall research design. Repeated measures, generalized linear models, and non-parametric tests are employed to examine differences due to main effects and interactions. The criterion for statistical significance for this study is alpha = .05. The dependent variable of primary interest is grading period GPA. This is the average of grades in core coursework (usually 4 or 5 courses) posted at the end of each grading period.

[1] We will refer to an academic year: e.g., 2007-2008, according to the ending date: e.g., 2008, throughout the rest of this paper.

To illustrate how the student identification and selection process works, grading period GPA in core academic coursework is the dependent variable of interest for this study. What GPA means and how valid and reliable such measures are can be argued indefinitely. We understand GPA here as a measure of the degree to which a student (or a student's performance) satisfies a teacher's expectation regarding mastery of a particular content studied within a particular period of time. Our GPA data is about the degree to which students steadily over time "continue to satisfy these expectations."[2] What is believed and recorded to be real is very often real in its consequences. And consequences like failure or drop-out or belief in being "not good" in a subject are very real.

General Classification Procedures: We began the identification process with the longitudinal database of individual student academic performance as measured by average grading period GPA in core coursework from 6th through 8th grade. This results in GPA time data vectors corresponding to the grading periods used by the schools during the particular academic years in question. These vary from 2 semesters to 4 or 6 marking periods to 4 quarters to 3 trimesters. To construct a meaningful timeline, we convert these to time points: e.g., 6 grade marking period 2 is 6.50, 8th grade marking period 4 or semester 2 is 9.00, etc.

Our focus is on individual classification. Initially, we produce longitudinal "spaghetti" plots of spline-smoothed individual student core GPA trajectories over time.[3] By examining individual plots, we can select students exhibiting characteristic trajectories which, if extrapolated into 9th grade and beyond, would predict prompt and total academic failure.

Historical Validation Studies: A previous study (Lacefield, Zeller, & Van Kannel-Ray, 2010), applied this method to data available for a historical cohort of N=88 students in this rural school district in 8th grade in 2008 and 9th grade in 2009, none of whom ever received academic support services.[4] Two groups were identified: N=22 students who appeared to be on a clear

[2] Some may argue that, strictly speaking, GPA data is not about "growth." To use latent growth curve modeling, for instance, you have to have a reasonable theory of something that "accumulates" over time: e.g., achievement test scores. An achievement test is constructed of items that sample a very large domain of, say, knowledge or applied knowledge. A person is theorized to know practically nothing and then to "acquire" more and more "knowledge" .. defined as domain coverage understanding .. as he or she goes along over time through exposure, study, and/or practice. Eventually that person could master the domain. So achievement tests provide measures of sufficient bandwidth (i.e., broad information curves) that can be used multiple times at different "points of growth." That data can be modeled as a logistic curve over a baseline of time.

It lies beyond the scope of this paper to argue whether or not GPA represents a broader concept that would meet the requirement of 'accumulated growth." However, arguing that a necessary condition to satisfying teacher expectations is in fact a "growth" of some kind, then evidence of a stable level of satisfaction over time (in courses that represent increasing breath of knowledge, level of difficulty, or deeper understanding) does reflect (perhaps the logarithm of) growth. For example, a C in arithmetic, followed by a C in pre-algebra followed by a C in algebra followed by a C in calculus or trig clearly requires and demonstrates "growth." This student's scores each year on MATH ACT would increase, probably following a logistic curve. The same argument could be applied to language arts or science or social studies ... or to music or shop. With appropriate domain-referenced measures, it could also apply to "enrichment" courses that actually enrich as well as to "advanced" courses that actually advance (or even "remedial" courses that actually repeat).

[3] Using SAS Proc TRANSREG with a smoothing factor of .80.

[4] A slightly expanded version of this same cohort appears in this current study as the 2008 cohort.

downward trajectory and very much at-risk of academic failure and a peer comparison group of N=66 other students appearing relatively well prepared for high school.

We then "went forward" to see what actually happened to these students in the 9th grade in the absence of coaching. Most grade trajectories of students previously identified as at-risk continued to plummet, well below a GPA < 1.0 which we chose as an indicator of academic failure and likely eventual drop-out. Approximately 77% of the 8th grade students identified as at-risk of failure did fail in the 9th grade. Likewise, 86% of the students expected to pass did so (McNemar's S1 = 21.6216, $p < .0001$). This pilot seemed to validate the method as an appropriate tool for identifying at-risk students. The odds ratio for passing was estimated more than 22:1 in favor of students who are NOT identified as at-risk. Here we present below an extended replication of that study using two additional cohorts (2007 and 2009 as well as re-analysis of 2008) with complete outcome data through 9th grade and beyond (thru 10th grade for 2008 and thru 11th grade for 2007). Results are similar.

Finer-Grain Classification Procedures: This study also demonstrates more fine-grained analysis of academic performance in terms of type of "risk assessment" both in various content areas as well as all core courses. Before we classified students as likely to be **"Successful"** or **"At-Risk."** We expanded our student identification and selection system to further classify "at-risk" students into three types:

- **At-Risk Falling:** Student GPA trajectories demonstrate a characteristic accelerating decline in academic performance in middle school, especially in 8th grade.
- **At-Risk Rising:** Student GPA trajectories which either began very low in 6th grade or fell during 6th and 7th grades but are rising toward or above 2.0 (a "C" average) during the 8th grade year.
- **At-Risk Failing:** Student GPA trajectories which show consistent failure in almost all courses during 6th, 7th, and 8th grades but who have been promoted anyway and are entering high school. (These students often need more extensive remediation and relearning than most academic support services can provide.)

A fourth classification **"Successful"** represents student GPA trajectories that may be rising, falling, or flat, but generally at or above GPA=2.0.

All students in all 3 cohorts were classified in this manner using specially designed data entry "dashboards" displaying both actual data trajectories and a smoothed performance curves for each student and allowing a classification code to be recorded. We used several "professional raters" (authors of this study) to make these classifications and are exploring various statistical procedures and machine learning algorithms that can do the same with high reliability. Some of this work and results of various validity tests are reported below.[5]

[5] The argument for using knowledgeable people is that decisions about students are constantly and properly made by school personnel. "Data-driven" procedures are tools to be used by those responsible, not to replace those people.

RESULTS, CONCLUSIONS, AND POINTS OF VIEW

Results were similar for all three cohorts. Figure 1 presents three pairs of "spaghetti plots" showing smoothed academic performance trajectories of students identified as "successful" (blue curves) and "at-risk" (red curves) at the end of 8th grade. The first set represents the relative performance of individual students in each cohort based on core academic data from 6th grade through 8th grade when classifications were made.[6] The second set represents 9th grade actual outcomes. (The latter trajectories we hope to change for the new 2010 cohort through provision of appropriate support services in the 2011 school year.)

Statistical Models: One way to describe and model our data statistically is to use repeated measures and generalized linear models with AR1 correlations. One such model is a 2B1W design with between groups fixed factors COHORT and STATUS, while using TIME as a within groups repeated measure. This allows us to address questions regarding various main effects and interactions. It turned out that COHORT and COHORT by STATUS interactions were not significant and rather uninteresting. We present a simpler 1B1W model combining and, hence averaging over COHORT. Figure 2 presents means plots for analyses based on both the raw and the smoothed data values for students in various status groups over time. Wald statistics show highly significant interaction effects.

Validity Studies: ***Predictive Validity:*** It seems very clear that the students our classification system identify as "At-Risk" are in fact highly likely to fail in high school. Using the same criteria of GPA < 1.0 as an indicator of academic failure employed in our previous study and considering the entire sample of N=265 students with data examined here, Figure 3 shows that approximately 74% of the 8th grade students identified as at-risk of failure did fail in the 9th grade. Likewise, 94% of the students expected to pass did so (McNemar's S1 = 7.11, p < .0077). The odds ratio for passing is more than 45:1 in favor of students who are NOT identified as at-risk.

Face Validity: There is a question regarding how reliably can educators and others use graphical data to classify students as potentially "Successful" or "At-Risk Falling, Rising, or Failing". Of course, decisions similar to these are constantly being made by college admissions officers, job interviewers, and others who ask for transcript data. As a pilot study, several of us independently classified all 307 students in this study using a simple "dashboard" form presenting both the raw data and the smoothed trajectory line for one student at a time. Inter-rater reliability for a 4-way classification was consistently high (Kappa~.80).

Misclassifications are of particular interest in this study. When educators agree on classifications, they are accepted and acted upon. When disagreements occur, educators look more closely at each of those particular students to reach consensus and to decide whether some sort of intervention might be of value.

Construct Validity: Computers can "recognize patterns" and classify data graphs as well, in some cases being able to provide probabilities of "correct" classifications. These tools are

[6] Note, for some new students joining a cohort in 8th or 7th grade, prior data was not available. Obviously, only available data can be used for predictive purposes. However, where data is not available does not mean that it is "missing" and certainly not "missing at random."

necessary when dealing with large numbers of students, especially when the data is disaggregated by, for instance, subject matter. When not just core GPA but also math, science, language arts, and history/social studies GPA performance trajectories over time can be examined, detailed diagnostic information becomes available. With computer analysis comes ability to "pre-screen" data and create initial classifications which then can be examined in detail by school personnel and confirmed or changed. For instance, if automatic classification identifies a particular student as "At-Risk" in core subject matter or math or science, then educators would look more closely at that student's performance (and other relevant data such as behavior patterns or attendance or personal problems) to decide if intervention could improve the situation.

We continue to explore computer algorithms and pattern recognition methods for classifying student performance trajectory data. One of us (Lacefield) has now examined and classified almost 1000 such curves for students in this school and in several others. A discriminant function analysis (DFA) of the data in this study for N=307 students produces a 95% agreement between the prior classifications and the DFA results (Figure 4). This outcome can be viewed as a simple indicator of "rater internal consistency" which appears to be quite high. However, using these same functions and their calibration and applying them to large sets of other pre-classified data from large urban school student populations has consistently resulted in classification agreements ranging from 85% upwards. Figure 5 reports results using random samples of performance data from a 2008 cohort of 8^{th} grade students in a large urban school district. A calibration DFA was run based on a random sample of N=325, followed by a second test DFA based on a second random sample of 170. The calibration run achieved 90% agreement with prior classifications, but more important, the test run using those discriminant function calibrations also achieved 88% agreement. This degree of replication even in pilot studies shows that DFA and other pattern matching methods could have substantial utility as computer diagnostic algorithms for educational progress.

Concurrent Validity: Michigan legislature mandated an additional year of algebra and science in state high schools, effective 2010-2011. One response of the rural high school examined here has been to establish a 9th grade "Algebra Support" course to be taken concurrently with Algebra I. The authors were asked to identify incoming 9th grade students who appeared at-risk in math preparation at the close of 8^{th} grade and who may benefit from this supplemental course. Out of 92 incoming students, N=16 were identified as "at-risk" and likely to benefit. If space is available, 14 more students should be considered. Interestingly, the school also employed its normal teacher recommendation and committee selection/referral procedure to identify possible students for this new course. The agreement was only about 50%. This example validates the need for understanding and agreeing upon selection criteria and using data to aid decision making.[7]

[7] (personal communication) This example illustrates some of the problems schools face when attempting to identify students for additional support. For instance, misbehavior in math class was sometimes offered as grounds for recommending supplemental math work. Furthermore later on, something very unfortunate was discovered. The extra math study was provided in the form of a scheduled course. When constrained to place 15-20 students together in that course on certain days and times, the school scheduling software scheduled those students together in ALL courses. This resulted essentially in a form of tracking with many at-risk students placed together in a sub-cohort group. Further, the supplemental math course was graded and students complained that now they failed in two math courses instead of one while being prevented from taking an elective course in which they were more interested and likely to make a better grade.

Dashboards for Decision Making and Educational Diagnosis: First, we developed programs to aggregate and display longitudinal data for individual students graphically using MS ACCESS (readily available in most schools) and SAS (a much more powerful statistical, analytical, and visual display tool). This was followed by "historical dashboards" based on permanent student information system data we wanted to make available to graduation coaches, counselors, and academic support personnel in the target school prior to the beginning of the school year for initial classification. We now present a periodic version of a "real-time dashboard" that presents the historical data and extends it automatically as new permanent grades are recorded each grading period. This dashboard also serves a diagnostic function by displaying "history" and "progress" not only in aggregate core courses but also in math, science, language arts, and history/social studies subject matter areas. Figure 6 presents examples of student dashboards selected to illustrate the variety of individual trajectories and classifications generated. These tools are designed to help decision-makers identify at-risk students and select academic support caseloads or class enrollments. Another similar "real-time dashboard" based on school electronic gradebook data entered constantly by classroom teachers could form a daily or weekly individualized information resource available to coaches and support personnel to monitor and advise the students they work with closely.

Individual and Intervention Assessment: We ourselves are currently involved in an intervention and promising practices demonstration project in this rural school district with funding provided by the national GEAR UP program through the Research Triangle Institute (RTI) of North Carolina (Van Kannel-Ray & Carpenter, 2010-2011). Approximately 36 students in the incoming 9th grade class (the 2010 cohort) were classified as potential beneficiaries of "graduation coaching" support services. Since the beginning of the school year, several new students have been added and several have either withdrawn or did not re-enroll. These students meet weekly with their coach, an experienced school counselor now completing a doctoral program at our university, and receive a variety of services and advice. The coach has daily access to school services, teachers, parents, the school's electronic gradebook, as well as to the periodic "dashboard" reports. Certainly someone in the school is paying careful attention to these students.

Based on the arguments in this paper, we have pretty good evidence in the form of at least group, if not individual, predictions about may happen in 9th grade. What our RTI project is trying to do is to provide a coach with resources, including periodic estimates of the final outcome, hoping that that coach can do something with those students to prevent such a prediction from coming true. Early evidence has been positive.

EDUCATIONAL AND SCIENTIFIC IMPORTANCE OF THE STUDY

To promote student retention, success, and perseverance to graduation, the goal of the classification and selection process discussed here is to identify students who will likely benefit from additional academic support and coaching services. Most rural and urban schools have adequate electronic student records for longitudinal comparative data analysis of individual student academic and social performance through middle and into high school. Demonstrating how this information can be used to make data-driven decisions regarding which students might benefit from additional resource allocation is a goal of our current research. The ability to make valid forecasts and diagnostic classifications by type of risk as well as subject area is a first step toward individualized educational intervention and individualized progress monitoring.

Figure 1: Incoming 9th grade smoothed academic performance trajectories.

Students in RED are predicted "At Risk" of academic failure in high school, although some may be "Rising". Unfortunately, most of these students did in fact fail in 9th grade.

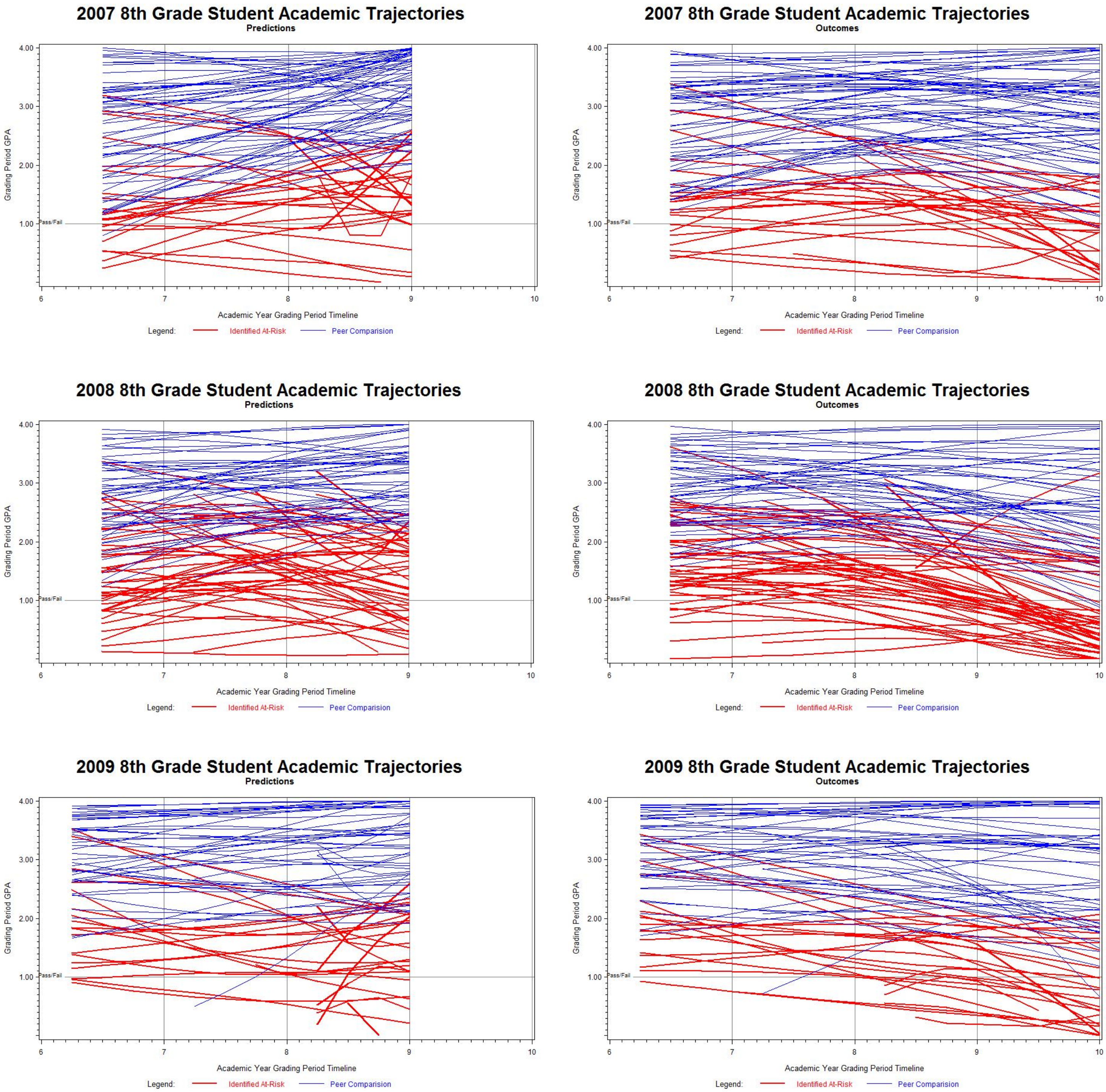
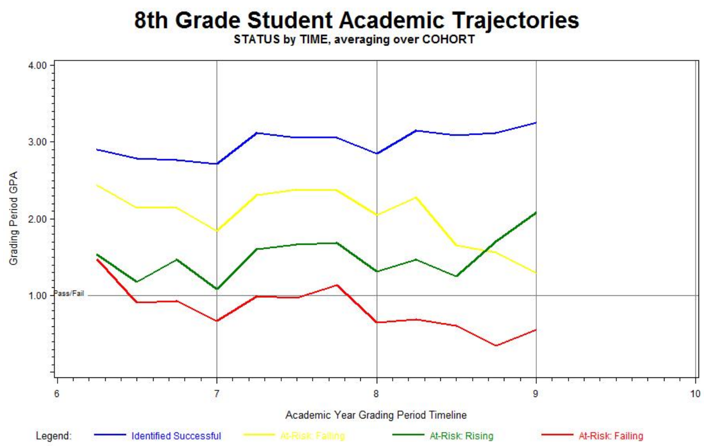
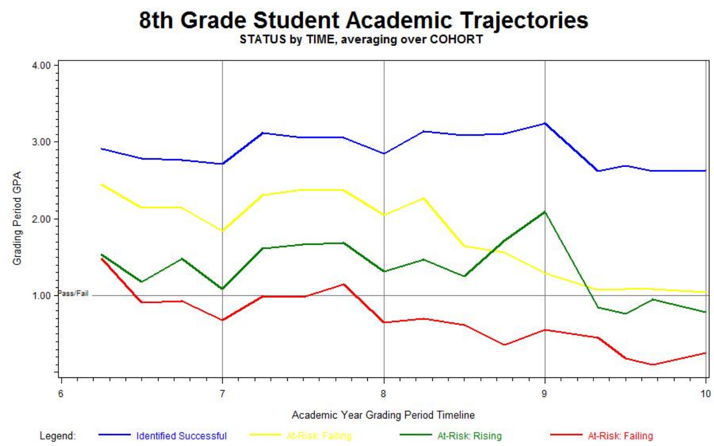


Figure 2: Means plots for academic performance over time, all three cohorts.

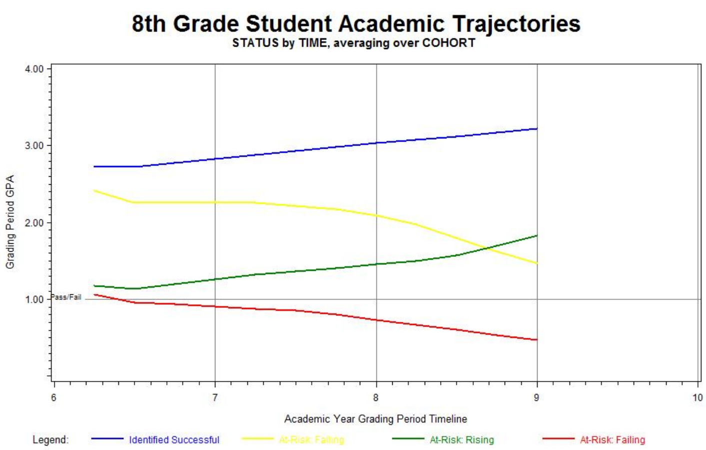
Plots are shown for raw grading period GPA data (row 1) as well as for smoothed data (row 2). In this and other similar studies, many students experience an almost .5 GPA drop from middle school to high school. This is sometimes referred to as a "hard transition." Note here that, unfortunately, even At-Risk students who were "Rising" in 8th grade continue on to fail in 9th grade in the absence of intervention.



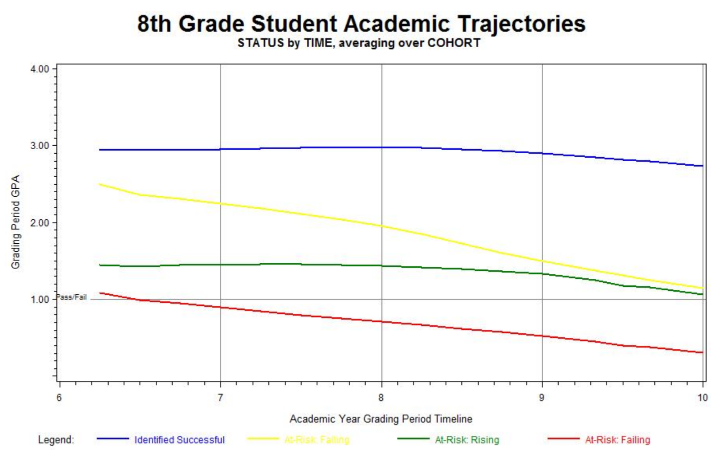
STATUS * TIME - Raw data through 8th grade			
Wald Statistics For Type 3 GEE Analysis			
Source	**DF**	**Chi-Square**	**Pr > ChiSq**
status	3	635.38	<.0001
time	11	145.16	<.0001
status*time	33	466.28	<.0001



STATUS * TIME - Raw data through 9th grade			
Wald Statistics For Type 3 GEE Analysis			
Source	**DF**	**Chi-Square**	**Pr > ChiSq**
status	3	763.91	<.0001
time	15	482.8	<.0001
status*time	45	547.96	<.0001



STATUS * TIME - Smoothed data through 8th grade			
Wald Statistics For Type 3 GEE Analysis			
Source	**DF**	**Chi-Square**	**Pr > ChiSq**
status	3	657.91	<.0001
time	11	62.49	<.0001
status*time	33	466.03	<.0001



STATUS * TIME - Smoothed data through 9th grade			
Wald Statistics For Type 3 GEE Analysis			
Source	**DF**	**Chi-Square**	**Pr > ChiSq**
status	3	788.57	<.0001
time	15	227.62	<.0001
status*time	45	442.23	<.0001

Figure 3: 8th grade classifications based on academic performance versus 9th grade outcomes in absence of intervention.

"Passing" 9th grade is defined here as an average core GPA => 1.00 for the last two quarters of 9th grade.

Students Predicted to:	Students Who Actually: Failed 9th grade	Passed 9th Grade	Row Totals
Fail in 9th Grade	N = 75 74%	N = 26 26%	N = 101 38%
Pass in 9th Grade	N =10 6%	N = 154 94%	N = 164 62%
Column Totals	N = 85 32%	N = 180 68%	N = 265 100%

Chi-Square = 133.28	p < .0001
McNemar's S1 = 7.11	p < .0077
Simple Kappa =	0.703
Odds ratio =	**45 to 1**

Figure 4: Discriminant function analysis (DFA) classification results based on student performance data from the current study.

Classification based on either choice of priors resulted in about 95% similarity for 2 levels of GROUP and in better than 85% similarity for 4 levels of STATUS. This indicates a high degree of rater consistency.

DFA Results Based on Equal Prior Probabilities

Number of Observations and Percent Classified into Group

From Group	Successful	At-Risk	Total
Successful	130	7	137
	94.89	5.11	100
At-Risk	2	65	67
	2.99	97.01	100
Total	132	72	204
	64.71	35.29	100
Priors	0.5	0.5	

Error Count Estimates for Group

	Successful	At-Risk	Total
Rate	0.0511	0.0299	0.0405
Priors	0.5	0.5	

Number of Observations and Percent Classified into Status

From Status	Successful	AR-Falling	AR-Rising	AR-Failing	Total
Successful	124	4	9	0	137
	90.51	2.92	6.57	0	100
AR-Falling	0	21	1	4	26
	0	80.77	3.85	15.38	100
AR-Rising	0	3	23	3	29
	0	10.34	79.31	10.34	100
AR-Failing	0	1	0	11	12
	0	8.33	0	91.67	100
Total	124	29	33	18	204
	60.78	14.22	16.18	8.82	100
Priors	0.25	0.25	0.25	0.25	

Error Count Estimates for Status

	Successful	AR-Falling	AR-Rising	AR-Failing	Total
Rate	0.0949	0.1923	0.2069	0.0833	0.144
Priors	0.25	0.25	0.25	0.25	

DFA Results Based on Proportional Prior Probabilities

Number of Observations and Percent Classified into Group

From Group	Successful	At-Risk	Total
Successful	131	6	137
	95.62	4.38	100
At-Risk	7	60	67
	10.45	89.55	100
Total	138	66	204
	67.65	32.35	100
Priors	0.67157	0.32843	

Error Count Estimates for Group

	Successful	At-Risk	Total
Rate	0.0438	0.1045	0.0637
Priors	0.6716	0.3284	

Number of Observations and Percent Classified into Status

From Status	Successful	AR-Falling	AR-Rising	AR-Failing	Total
Successful	132	2	3	0	137
	96.35	1.46	2.19	0	100
AR-Falling	2	20	1	3	26
	7.69	76.92	3.85	11.54	100
AR-Rising	2	3	23	1	29
	6.9	10.34	79.31	3.45	100
AR-Failing	0	3	1	8	12
	0	25	8.33	66.67	100
Total	136	28	28	12	204
	66.67	13.73	13.73	5.88	100
Priors	0.67157	0.12745	0.14216	0.05882	

Error Count Estimates for Status

	Successful	AR-Falling	AR-Rising	AR-Failing	Total
Rate	0.0365	0.2308	0.2069	0.3333	0.103
Priors	0.6716	0.1275	0.1422	0.0588	

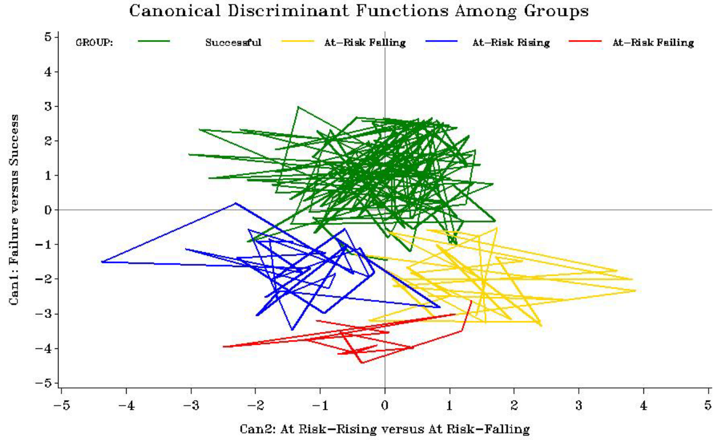


Figure 5: Discriminant function analysis (DFA) classification calibration and test results for similar academic performance data from a large urban school district.

Academic performance data for a 2009 cohort of approximately 600 8th grade students from a large urban school district was randomly sampled into two groups of complete data sets. A DFA involving the first group (N=325) was used to calibrate a set of discriminant functions. These were then used to classify students in the second group (N=170). The calibration results show approximately 90% similarity to prior classifications. This indicates a high degree of rater consistency. However, the test results show approximately 88% similarity to its own prior classifications. This degree of replication shows that the DFA is robust and generalizable and could have utility as a computer diagnostic algorithm.

DFA Results Based on Random Calibration Sample (N=325)

Number of Observations and Percent Classified into Status

From Status	Successful	AR-Falling	AR-Rising	AR-Failing	Total
Successful	197	6	3	1	207
	95.17	2.9	1.45	0.48	100
AR-Falling	4	59	2	1	66
	6.06	89.39	3.03	1.52	100
AR-Rising	10	3	22	0	35
	28.57	8.57	62.86	0	100
AR-Failing	0	1	3	13	17
	0	5.88	17.65	76.47	100
Total	211	69	30	15	325
	64.92	21.23	9.23	4.62	100
Priors	0.63692	0.20308	0.10769	0.05231	

Error Count Estimates for Status

	Successful	AR-Falling	AR-Rising	AR-Failing	Total
Rate	0.0483	0.1061	0.3714	0.2353	0.10
Priors	0.6369	0.2031	0.1077	0.0523	

DFA Results Based on Random Test Sample (N=170)

Number of Observations and Percent Classified into Status

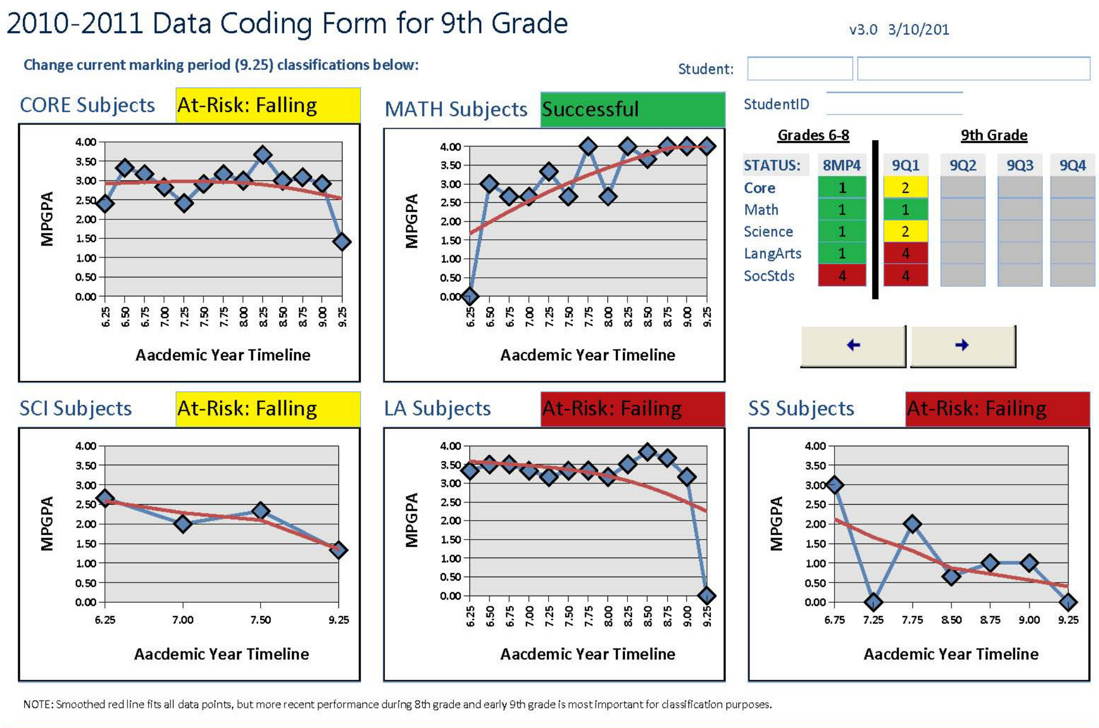
From Status	Successful	AR-Falling	AR-Rising	AR-Failing	Total
Successful	104	5	1	0	110
	94.55	4.55	0.91	0	100
AR-Falling	3	24	0	0	27
	11.11	88.89	0	0	100
AR-Rising	9	3	12	0	24
	37.5	12.5	50	0	100
AR-Failing	0	1	0	8	9
	0	11.11	0	88.89	100
Total	116	33	13	8	170
	68.24	19.41	7.65	4.71	100
Priors	0.63692	0.20308	0.10769	0.05231	

Error Count Estimates for Status

	Successful	AR-Falling	AR-Rising	AR-Failing	Total
Rate	0.0545	0.1111	0.5	0.1111	0.12
Priors	0.6369	0.2031	0.1077	0.0523	

Figure 6: Historical and Diagnostic Progress Dashboards

A)



B)

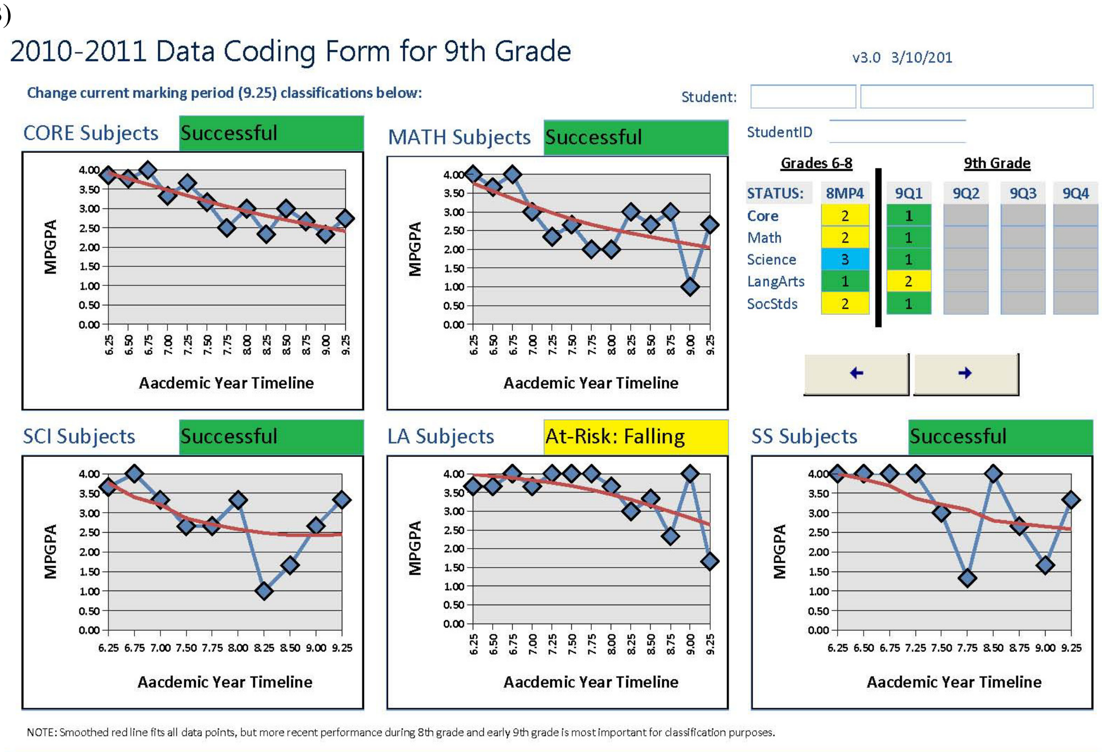
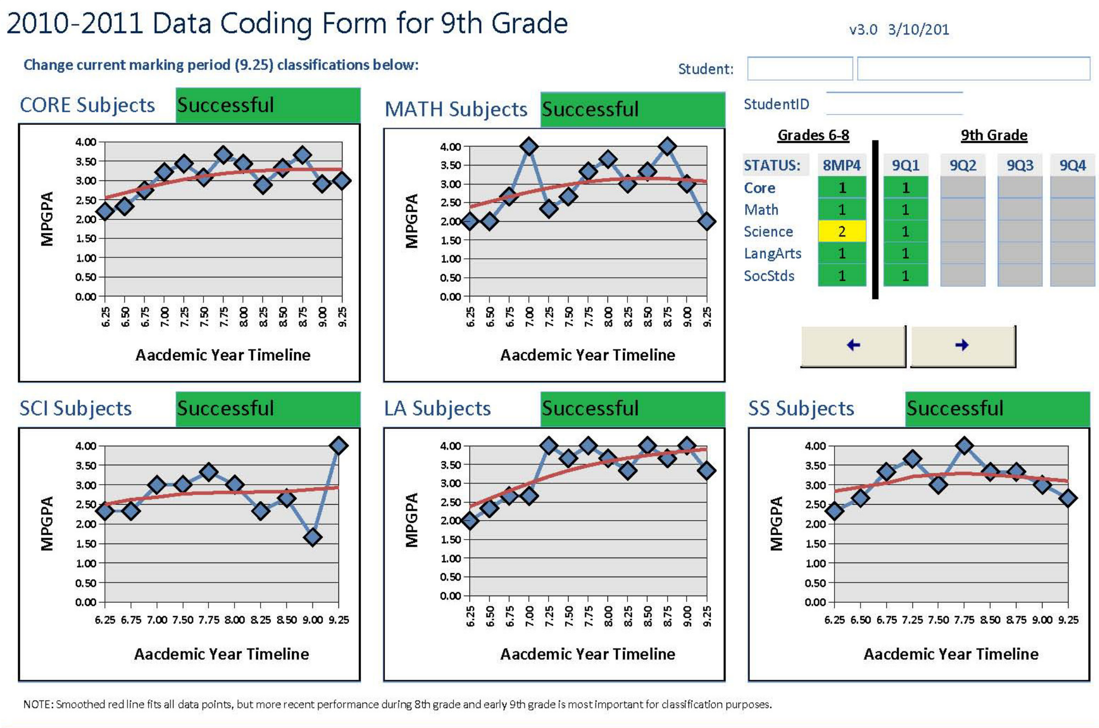


Figure 6 cont.: Historical and Diagnostic Progress Dashboards

C)



D)

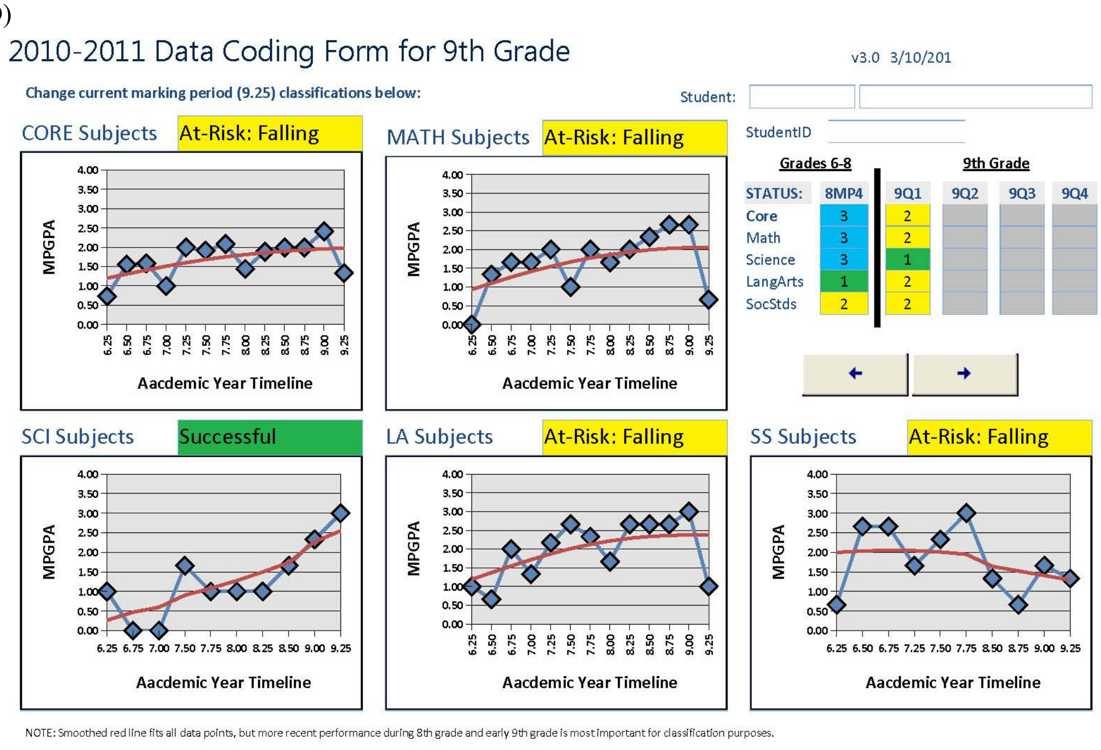
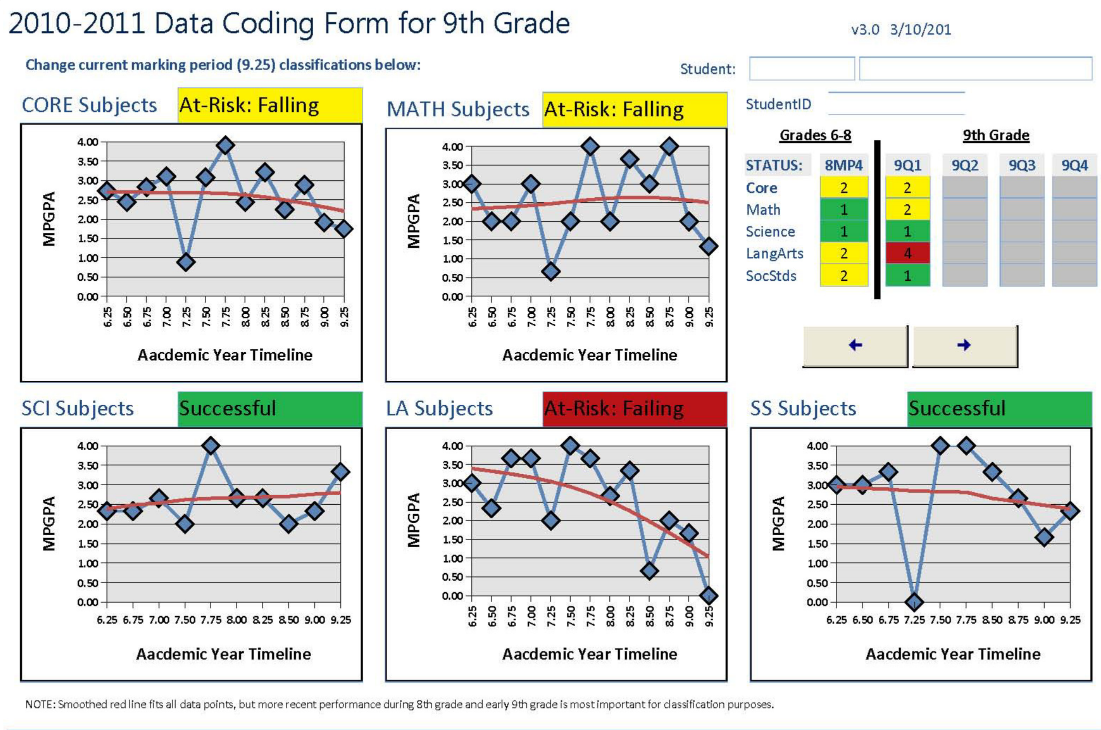
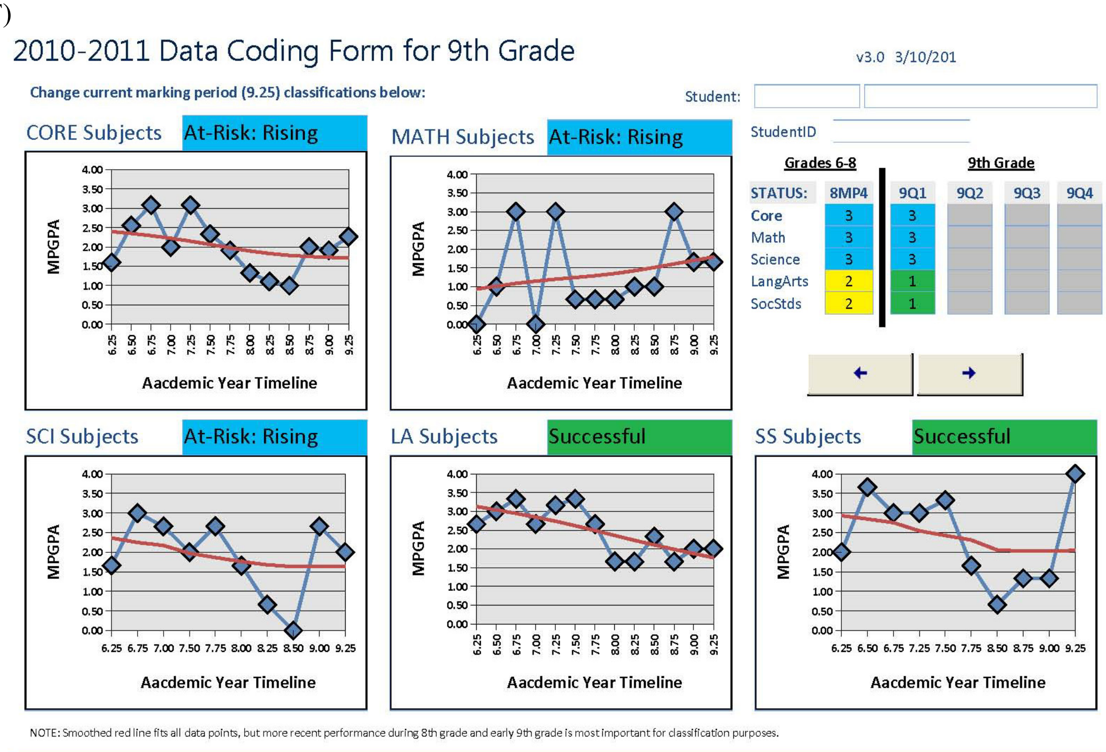


Figure 6 cont.: Historical and Diagnostic Progress Dashboards

E)



F)



References:

Armstrong, J., Anthes, K. (2001). How data can help: Putting information to work to raise student achievement, American School Board Journal, 188 (11), 38-41.

Bernhardt, V. (2009). Data driven decision making takes a big-picture view of the needs of teachers and students, Journal of Staff Development,. 30 (1), 24-27.

Cromey, A., van der Ploey, A., Masini, B. (2000). The call for data-driven decision making in the Midwest's schools: NCREL response.

Jackson, A., Davis, G. (2000). Turning points 2000: Educating adolescents in the 21st century. New York: Teachers College Press.

Kretovic, J. (PD) (9/15/2005 – 9/14/2011) MERC GEAR UP learning centers II project, U.S. Department of Education: Gaining Early Awareness & Readiness for Undergraduate Programs. CFDA # 84.334A, Project #P334A050257.

Lacefield, W.E., Zeller, P., Van Kannel-Ray, N. (May 1, 2010) Graduation Coaching in High-Need Urban High Schools. A paper presentation during the 2010 Annual Meeting of the American Educational Research Association, Denver, CO. ERIC ED509289

Swanson, C.B., (2003) Keeping count and losing count: Calculating graduation rates for all students under NCLB accountability. The Urban Institute Education Policy Center, 2100 M Street NW, Washington DC 20037.

Swanson, C.B. (2008) Cities in crisis: A special analytic report on high school graduation. Editorial Projects in Education Inc. Research Center. 6935 Arlington Road, Suite 100, Bethesda, MD 20814-5233

Van Kannel-Ray, N. (co-PD) & Carpenter, S. (co-PD) (2010-2011) Graduation coaching in rural school settings project, North Carolina: Research Triangle Institute, Project # 0209234.002.016-7.

Van Kannel-Ray, N., Lacefield, W.E., Zeller, P. (2008) Academic case managers: Evaluating a middle school intervention for children at-risk. Journal of Multi-Disciplinary Evaluation, 5 (10), 21-29.

Van Kannel-Ray, N., Zeller, P., Lacefield, W.E. (2009) Academic case management: Promising interventions for closing achievement gaps in multicultural urban settings. ERS Spectrum, 27 (3), 1.