

The Early Warning Indicator Index: Identifying High School Students at Risk of Not Graduating on Time

What is the Early Warning Indicator Index?

The Early Warning Indicator Index (EWII) assigns students who were enrolled in eighth grade at the end of the 2009-2010 school year one of five risk levels (*Very High Risk*, *High Risk*, *Borderline*, *At Risk*, or *Low Risk*) based on their grade 8 MCAS results, grade 8 attendance rate, age as of September 1, 2010, and number of in- and out-of-school suspensions in the 2009-2010 school year. The EWII can be used by the district as a starting point for identifying and supporting students who may be at risk of not graduating on time from high school.

Detailed rosters for these students—the majority of whom are likely to be high school freshmen in the district in the 2010-2011 school year—are provided to the district in a Microsoft Excel workbook via its Accountability 2010 Drop Box in the Massachusetts Department of Elementary and Secondary Education's (ESE's) online Security Portal. A tool that automatically calculates risk levels for students who are not listed on the roster—the "EWII Risk Calculator"—is also provided.

Research is clear that a variety of factors contribute to a student's likelihood of not graduating on time. Because the EWII risk levels are calculated using statewide data, the district should also use local data to identify and support students, including, but not limited to, information on course grades, course completion rates, and freshman year attendance rates.

How Were the Risk Levels Determined?

ESE used data from the 2008 and 2009 high school graduation cohorts to evaluate the connection between a student's middle school performance and whether they graduated on time. After considering a variety of models, ESE identified one consisting of five factors that, taken together, predicted with a substantial degree of reliability whether a student graduated on time:

1. The student's eighth grade MCAS mathematics score;
2. The student's eighth grade MCAS English language arts (ELA) score;
3. The student's attendance rate in eighth grade;
4. The number of times the student was suspended (either in or out of school) in eighth grade; and
5. The student's age following the end of eighth grade.

While each of the indicators is predictive of on-time graduation for students statewide, ESE found that these degree of predictability of individual indicators differed somewhat for students residing in urban communities versus students who reside in suburban or rural communities. As a result, risk levels were established separately for these populations in proportion to their actual on-time graduation rates. For example, the calculations used to assign urban students a risk level of *Very High Risk* were based on an analysis of students whose collective on-time graduation rate was 19.3 percent. Suburban and rural students assigned this risk level had a collective on-time graduation rate of 27.8 percent. (See Table 6 at the end of this document for a list of the Commonwealth's 24 urban districts.)

What Information is Included on My District's 2010-2011 Roster?

Each roster contains identifying information, MCAS results, attendance data, and suspension data for each individual student who was enrolled in eighth grade in the district as of the June 2010 Student Information Management System (SIMS) collection. The rosters do not contain information for high school freshman who may have transferred into the district after June 2010.

Specifically, risk levels assigned in the rosters are based on the five data elements below:

1. Spring 2010 grade 8 MCAS mathematics scores;
2. Spring 2010 grade 8 English language arts (ELA) scores;
3. 2009-10 grade 8 attendance rate;
4. Number of 2009-10 grade 8 suspensions (in or out of school);
5. Age as of September 1, 2010.

In addition to these five indicators, the rosters include other descriptive information drawn from 2009-2010 SIMS and MCAS data, including: the student's name, gender, race/ethnicity, the school they were enrolled in as of June 2010, their town of residency as of June 2010, the student's mobility rate, the student's growth percentiles in mathematics and ELA, if available, and whether they received free or reduced price lunch, special education services, and/or were identified as limited English proficient (LEP) in 2009-2010.

Please refer to Table 5 at the end of this document for the complete list of variables included on the roster and a detailed description of each.

Which Districts Were Assigned Rosters?

Rosters are provided for all districts that served eighth grade students in 2009-2010. Regional vocational schools, charter high schools, and other single school districts that begin at grade 9 have not been provided rosters. (See "Why Don't the Rosters Include Grade 9 School Assignments" below.)

Why Do the Rosters not Include Grade 9 School Assignments?

EWII rosters released in early November 2010 do not include current grade 9 school assignments because ESE does not know which schools students are enrolled in until October SIMS data have been submitted by all districts and fully processed. We recommend that districts enter grade 9 school assignment information to the rosters based on local records.

Were Students with Missing Information Assigned Risk Levels?

Yes. All students in the rosters have an age, calculated based on the student's date of birth, and grade 8 suspension data (the default being no suspensions). However MCAS or attendance data were incomplete for some students. In these cases, ESE generated imputed (estimated) MCAS scores and attendance rates for students who were missing these data. The imputed values allowed ESE to assign risk levels to every student on the roster, even if they were missing some data.

To provide additional context, each student record, whether based on actual or imputed values, also contains a "probability" variable (represented as a number between 0 and 1) which gives the student's predicted probability of graduating on time based on their eighth grade data. The higher the value, the more likely the student will graduate on time; the lower the value, the more likely the student will not graduate on time.

Taken together, these data give districts information they need to begin a more thorough inquiry into the interventions, supports, and unique circumstances of each student. For reference, the roster contains three "flag" variables indicating if a student lacked MCAS or attendance data and as a result has imputed values for one or more of these variables.

Do the Student Groups at each Risk Level Share Common Characteristics?

Generally speaking, students assigned lower risk levels had higher MCAS scores, higher attendance rates, lower incidents of suspension, and are closer to age 14 than students assigned to higher risk levels, who tended to have inverse characteristics.

Tables 1 and 2 display the characteristics of the 21,439 eighth grade students from urban communities and the 48,999 students from suburban and rural communities who did not have missing data and were assigned a risk level for 2010-2011. For example, the 9,512 students from urban communities assigned to the Very Low Risk category had, on average, an ELA scaled score of 252, a mathematics scaled score of 249, an attendance rate of 97 percent, less than one suspension, and were about 14 and a half years of age. Conversely, the 1,243 students from urban communities assigned to the Very High Risk category had, on average, an ELA scaled score of 223, a mathematics scaled score of 214, an attendance rate of 78 percent, more than 2 suspensions, and were about 15 and a quarter years of age. It is important to keep in mind that these figures are only averages; for example, there are students assigned very low risk levels with poor MCAS scores and vice-versa.

How Accurate is the EWII in Predicting On-Time Graduation?

As stated previously, ESE used actual data from the 2008 and 2009 graduating cohorts to evaluate the connection between each student's middle school performance and whether they graduated on time. In turn, this information was then used to estimate the likelihood that a student enrolled in eighth grade in the 2009-2010 school year will graduate on time four years later in the spring/summer of 2014.

Tables 3 and 4 display the actual on-time graduation rates for all of the students in the 2008 and 2009 graduating cohorts whose middle school data from 2003-2004 and the 2004-2005 school years, respectively, were used in the model.

Can I Assign a Risk Level to a Student who is New to the District?

Yes, ESE has developed the "EWII Calculator" in Microsoft Excel for this purpose. As long as the student was enrolled in a Massachusetts school in 2009-2010 and data are available for that student from that year, districts can use the spreadsheet to plug in information for all of the eighth grade data elements (i.e., MCAS ELA, MCAS mathematics, attendance, suspensions, or age). The calculator will automatically generate a risk level for the student using the same algorithm used to generate risk levels for the student on the district's roster.

There are separate calculators for students whose town of residence is urban and for those whose town of residence is suburban or rural. Within those, there are different calculators for students missing specific data elements. Because the rosters were generated by district—as opposed to town of residence—rosters for some districts contain data for both urban and non-urban students (e.g., an "urban" student living in Boston may attend school in a suburban community). It is therefore important to first identify whether a student's town of residence is urban or non-urban and then choose the appropriate calculator.

Who Can Access the Rosters?

Because the rosters contain student-level data, only district staff assigned the authority to view the data may access them. District personnel with the "Accountability – District" security role may access these data. This security role is assigned by the district's Directory Administrator. For a list of directory administrators by district, please visit <http://www.doe.mass.edu/infoservices/data/diradmin/list.aspx>.

The rosters can be accessed via the Accountability 2010 Data Drop Box in Drop Box Central in the ESE Security Portal.

Who May I Contact if I have Questions or Need Assistance?

General questions about the EWII, its use in identifying and supporting students, and ongoing networking activities and grant opportunities related to high school graduation initiatives may be directed to the Office of College and Career Readiness via email to achievement@doe.mass.edu.

Technical questions about the EWII may be directed to the Office of School Improvement Grant Programs via email to ayp@doe.mass.edu.

**Table 1: Characteristics of Student Groups by EWII Risk Level
(Students from Suburban and Rural Communities with Complete Data)**

N = 48,999 Students		Range	Average (Unrounded)	Std. Deviation
Very Low Risk N = 40,319 Students	Grade 8 ELA MCAS Scaled Score	210-280	253.41	9.44
	Grade 8 Math MCAS Scaled Score	202-280	249.27	16.40
	Grade 8 Attendance Rate	74-100	96.50	2.80
	Number of Suspensions	0-9	.03	.22
	Age as of September 1, 2010	12-17	14.52	.36
Low Risk N = 4,521 Students	Grade 8 ELA MCAS Scaled Score	206-276	236.98	10.63
	Grade 8 Math MCAS Scaled Score	204-274	222.91	9.83
	Grade 8 Attendance Rate	71-100	93.35	4.31
	Number of Suspensions	0-15	.25	.84
	Age as of September 1, 2010	13-17	14.70	.46
Borderline N = 3,040 Students	Grade 8 ELA MCAS Scaled Score	204-270	231.55	12.25
	Grade 8 Math MCAS Scaled Score	202-270	219.42	8.74
	Grade 8 Attendance Rate	55-100	90.54	5.48
	Number of Suspensions	0-30	.88	1.77
	Age as of September 1, 2010	13-17	14.81	.52
High Risk N = 625 Students	Grade 8 ELA MCAS Scaled Score	204-270	228.81	11.81
	Grade 8 Math MCAS Scaled Score	202-260	217.63	7.55
	Grade 8 Attendance Rate	56-99	85.95	6.53
	Number of Suspensions	0-23	2.69	3.25
	Age as of September 1, 2010	14-17	14.91	.58
Very High Risk N = 494 Students	Grade 8 ELA MCAS Scaled Score	204-262	226.07	12.73
	Grade 8 Math MCAS Scaled Score	204-258	216.38	7.80
	Grade 8 Attendance Rate	1-93	73.62	12.43
	Number of Suspensions	0-37	3.94	5.09
	Age as of September 1, 2010	13-17	15.05	.60

**Table 2: Characteristics of Student Groups by EWII Risk Level
(Students from Urban Communities with Complete Data)**

N = 21,439 Students		Range	Average (Unrounded)	Std. Deviation
Very Low Risk N = 9,512 Students	Grade 8 ELA MCAS Scaled Score	218-280	251.68	8.25
	Grade 8 Math MCAS Scaled Score	212-280	249.21	14.47
	Grade 8 Attendance Rate	81-100	96.88	2.63
	Number of Suspensions	0-12	.05	.35
	Age as of September 1, 2010	12.3-17.1	14.49	.42
Low Risk N = 4,218 Students	Grade 8 ELA MCAS Scaled Score	208-270	241.01	8.21
	Grade 8 Math MCAS Scaled Score	204-280	226.78	10.84
	Grade 8 Attendance Rate	72-100	95.02	3.62
	Number of Suspensions	0-14	.19	.70
	Age as of September 1, 2010	12.1-17.5	14.66	.52
Borderline N = 5,049 Students	Grade 8 ELA MCAS Scaled Score	204-266	230.69	11.50
	Grade 8 Math MCAS Scaled Score	202-266	219.24	8.21
	Grade 8 Attendance Rate	67-100	93.03	4.54
	Number of Suspensions	0-20	.46	1.16
	Age as of September 1, 2010	13.2-17.3	14.87	.62
High Risk N = 1,417 Students	Grade 8 ELA MCAS Scaled Score	202-262	225.82	11.65
	Grade 8 Math MCAS Scaled Score	202-258	216.16	6.91
	Grade 8 Attendance Rate	64-100	87.80	5.54
	Number of Suspensions	0-27	1.23	1.98
	Age as of September 1, 2010	11.1-18.4	15.08	.69
Very High Risk N = 1,243 Students	Grade 8 ELA MCAS Scaled Score	204-256	222.95	11.36
	Grade 8 Math MCAS Scaled Score	202-258	214.47	6.43
	Grade 8 Attendance Rate	9-97	77.69	10.67
	Number of Suspensions	0-30	2.38	2.88
	Age as of September 1, 2010	13.6-18	15.28	.69

Table 3: Comparison of Actual On-Time Graduation Rates for Students in 2008 and 2009 Graduating Cohorts and EWII Risk Levels (Students from Suburban and Rural Communities)

		Very Low Risk	Low Risk	Borderline	High Risk	Very High Risk	Total
Did Not Graduate On Time	Number of Students	2,166	1,876	2,503	737	687	7,969
	Percentage	3.3%	14.6%	33.4%	58.5%	72.2%	
Graduated On Time	Number of Students	63,056	10,957	4,992	522	264	79,791
	Percentage	96.7%	85.4%	66.6%	41.5%	27.8%	
Total	Number of Students	65,222	12,833	7,495	1,259	951	87,760
	Percentage	100%	100%	100%	100%	100%	

Table 4: Comparison of Actual On-Time Graduation Rates for Students in 2008 and 2009 Graduating Cohorts and EWII Risk Levels (Students from Urban Communities)

		Very Low Risk	Low Risk	Borderline	High Risk	Very High Risk	Total
Did Not Graduate On Time	Number of Students	552	1,218	4,152	1,945	2,099	9,966
	Percentage	5.2%	13.7%	33.3%	60.6%	80.7%	
Graduated On Time	Number of Students	10,118	7,658	8,301	1,265	502	27,844
	Percentage	94.8%	86.3%	66.7%	39.4%	19.3%	
Total	Number of Students	10,670	8,876	12,453	3,210	2,601	37,810
	Percentage	100%	100%	100%	100%	100%	

Table 5: 2010-2011 EWII Roster File Layout

Variable Name		Type	Width	Decimals	Definition
1	District_Code	String	4	0	4-digit district organization code based on June SIMS
2	SASID	String	10	0	state-assigned student identifier
3	First_Name	String	25	0	first name
4	Middle_Name	String	25	0	middle name
5	Last_Name	String	25	0	last name
6	School	String	8	0	8-digit code of enrolled school as of June 2010
7	Grade	String	2	0	enrolled grade as of June 2010
8	Truancy	Numeric	1	0	number of school days student was recorded as truant
9	Gender	String	1	0	gender
10	District_Name	String	30	0	name of enrolled district where student was assessed in the spring of 2010
11	School_Name	String	30	0	name of enrolled school based on June SIMS
12	Low_Income	Numeric	1	0	free/reduced price lunch status as of June 2010 (1 = y, 0 = n)
13	Limited_English_Proficient	Numeric	1	0	limited or formerly limited English proficient status as of June 2010 (1 = y, 0 = n)
14	Special_Education	Numeric	1	0	special education status as of June 2010 (1 = y, 0 = n)
15	MCAS_Alt_ELA	String	1	0	took ELA alternate assessment in eighth grade (1 = y)
16	Actual_Gr8_ELA_Scaled_Score	Numeric	3	0	Actual eighth grade ELA scaled score
17	Gr8_ELA_Performance	String	2	0	eighth grade ELA performance level: A = Advanced P = Proficient NI = Needs Improvement F = Failing W = Warning
18	Gr8_ELA_Growth_Percentile	Numeric	3	0	eighth grade ELA student growth percentile
19	MCAS_Alt_MTH	String	1	0	took math alternate assessment in eighth grade (1 = y)
20	Actual_Gr8_Math_Scaled_Score	Numeric	3	0	Actual eighth grade math scaled score

Table 5: 2010-2011 EWII Roster File Layout

Variable Name		Type	Width	Decimals	Definition
21	Gr8_Math_Performance	String	2	0	eighth grade math performance level: A = Advanced P = Proficient NI = Needs Improvement F = Failing W = Warning
22	Gr8_MTH_Growth_Percentile	Numeric	3	0	eighth grade math student growth percentile
23	Race	String	1	0	race/ethnicity as of June 2010: A = Asian & Pacific Islander B = African American/Black H = Hispanic N = Native American W = White
24	Date_of_Birth	String	10	0	date of birth
25	Age	Numeric	4	2	Age as of 9/1/10
26	Residency_Code	String	8	0	8-digit code of town of residency as of June 2010
27	Town_of_Residence	String	30	0	name of town of residency as of June 2010
28	Urban	Numeric	1	0	whether town of residency is urban (1 = y, 0 = n)
29	Gr8_Attendance	Numeric	3	0	actual or imputed eighth grade attendance rate
30	In_School_Suspension	Numeric	1	0	number of incidents coded as in-school suspension as per district definition
31	Out_of_School_Suspension	Numeric	1	0	number of incidents coded as out-of-school suspension as per district definition
32	Sum_School_Suspension	Numeric	3	0	sum of above two numbers
33	Gr8_ELA_Scaled_Score	Numeric	3	0	actual or imputed ELA scaled score
34	Gr8_Math_Scaled_Score	Numeric	3	0	actual or imputed math scaled score
35	Probability	Numeric	4	3	predicted probability of graduating on time based on their eighth grade data; the higher the value, the more likely the student will graduate on time; the lower the value, the more likely the student will not graduate on time (number between .0 and 1)
36	Risk_Level	String	25	0	risk level
37	Transfer_in	Numeric	1	0	student transferred into a school after the beginning of the school year (1 = y, 0 = n)

Table 5: 2010-2011 EWII Roster File Layout

Variable Name		Type	Width	Decimals	Definition
38	Transfer_out	Numeric	1	0	student transferred out of a school after the beginning of the school year (1 = y, 0 = n)
39	Stable	Numeric	1	0	student remained in the same school for the entire school year (1 = y, 0 = n)
40	Missing_ELA	Numeric	1	0	flag denoting whether student is missing eighth grade ELA data (1 = the data were missing and estimated data imputed as a substitute for the missing data; 0 = actual data displayed)
41	Missing_MTH	Numeric	1	0	flag denoting whether student is missing eighth grade math data (1 = the data were missing and estimated data imputed as a substitute for the missing data; 0 = actual data displayed)
42	Missing_Attendance	Numeric	1	0	flag denoting whether student is missing eighth grade attendance data (1 = the data were missing and estimated data imputed as a substitute for the missing data; 0 = actual data displayed)

Table 6: List of 24 Urban Districts

District Code	District Name
0035	Boston
0044	Brockton
0049	Cambridge
0057	Chelsea
0061	Chicopee
0093	Everett
0095	Fall River
0097	Fitchburg
0100	Framingham
0128	Haverhill
0137	Holyoke
0149	Lawrence
0153	Leominster
0160	Lowell
0163	Lynn
0165	Malden
0201	New Bedford
0236	Pittsfield
0243	Quincy
0248	Revere
0274	Somerville
0281	Springfield
0293	Taunton
0348	Worcester