Measuring Growth Toward College Readiness:

Using MAP (Measures of Academic Progress)

Scores to Predict Success on the ACT Test Benchmark Scores

Robert A. Theaker

Northwest Evaluation Association, Portland, Oregon

Clay S. Johnson

University of Arkansas

Poster session presented at the 2012 meeting of

The American Educational Research Association, Vancouver, BC

May 2013: Revised to remove grade 3 fall and spring term along with grade 4 fall term data due to small sample size. Minor copy edits were also completed for this revision.

MEASURING GROWTH TOWARD COLLEGE READINESS

Abstract

This study expands the authors' continuing line of research toward accurately estimating college readiness benchmarks through elementary and middle grades in mathematics and reading. Benchmark scores established by ACT® (22 in mathematics and 21 in reading) were used as score criteria. The MAP® assessments were used in order to connect third through 11th grade student growth with college readiness. A large, multistate sample allowed for the extension of the authors previous analyses. The use of 11th grade MAP and ACT score correlations to backward map benchmark scores resulted in predictive accuracy beginning in grade 4 spring. This study provides the current findings of a work in progress, whose final goal is a model for predicting college readiness.

Objectives

As college readiness is soon to become the accepted measure of accountability in K-12 education, reliable interim measures of progress toward this goal are essential for the success of students and the effectiveness of teachers. This study continues the authors' line of research to link ACT scores associated with college success with Measures of Academic Progress (MAP) scores from NWEA, which can be used to track student growth from kindergarten through grade 12.

The goals of this study were threefold. First, previously calculated benchmarks of college readiness from past analyses (Theaker & Mellema, 2010) were re-investigated with a recent sample and were validated using receiver operating characteristic (ROC) curve analysis. Second, grade-by-grade benchmarks were established for predicting a student's likely growth toward levels of ACT performance desired for admission using ACT's college readiness benchmarks. Third, an initial work with grade by grade benchmarks for different institution types.

Significance

The current conversation in Washington, DC regarding reauthorization of the federal Elementary and Secondary Education Act focuses heavily on college readiness for all students. As this represents an important shift from the NCLB mandates toward proficiency for all, our assessments for measuring student progress must soon shift correspondingly.

This study – and the line of research it continues – builds an essential link between two highly reliable, well validated, current measures of student progress in order to offer a practical and accessible means of early identification of growth toward college readiness. Our establishment of

growth benchmark MAP scores through the elementary grades, and our ability to predict future ACT performance offers a unique way for teachers to appropriately target early interventions needed for students' continued academic success.

Theoretical Framework

The term *college readiness* is as inconsistently defined as it is widely used. For our current purposes, ACT scores were used for measuring students' preparedness for college, as research from ACT has repeatedly shown to be appropriate (ACT, 2010). Progress toward reaching that goal, beginning in the eighth grade, was also tracked through their EXPLORE, PLAN, ACT assessment system (ACT, 2009; Allen, Bassiri, & Noble, 2009). While others have scrutinized the relationship between ACT scores and students' demonstrated college success, as in a recent study from the National Bureau of Economic Research (Bettinger, Evans, & Pope, 2011), the near-ubiquity of the ACT exams in college admissions, course placement, and scholarship decisions seems to warrant their consideration in our study as a pragmatic measure of college readiness.

Likewise, MAP assessments have been designed and validated for measuring student growth in mathematics and reading (NWEA, 2008). These growth measures have been validated extensively to link students' early progress toward proficiency on exams for state accountability (NWEA, 2009), but their utility in measuring progress in preparation for college is yet to be established on a large scale. Our access to the Growth Research Database at NWEA, which includes students who have an 11th grade MAP score and an ACT score in near succession, gives us a unique means for linking students' progress from third grade through high school toward an ACT benchmark predictive of college success.

MEASURING GROWTH TOWARD COLLEGE READINESS

To date, no national exam has been used on a large scale to backward map predicted college success through the elementary grades. This study is unique in making predictions of college readiness much earlier in a student's career and therefore allowing earlier intervention toward keeping students on track toward college success.

Table 1 shows ACT's published benchmarks based on empirical data of students' college performance (ACT, 2009). The given scores for mathematics and reading provide our reference points for linking the ACT system with MAP scale scores. Based on the student data available for this study, we will focus on the subjects of mathematics and reading.

Table 1

College Readiness Benchmarks for the ACT Assessments

Subtest	ACT
Mathematics	22
Reading	21
English	18
Science	24
Composite	21

Note. Values were calculated by ACT from empirical data to identify the scores of students with a 50% likelihood of achieving at least a *B* in a related freshman-level course. From "College Readiness Standards for EXPLORE®, PLAN®, and the ACT®," p. 3. Copyright 2011 by ACT, Inc.

Data Sources

Data included in our analyses were contained in two groups: a small sample of previously analyzed data and a larger dataset for validating and extending our previous work. The initial data came from a subset of students in our larger study for which we had access to MAP scores (Theaker &

Mellema, 2010). This included approximately 350 students from eight charter schools within a single management system. These data included test scores for grades 3 through 8 for the Explore® exam only. The larger dataset included in the current study represents 201,168 test records for reading and math from 10 school districts in three states. This study used the fall and spring testing terms for reading or math. A student's grade 11 ACT score was matched with a MAP score within grades 4 through 11, which was not available for our past analysis. In all, the sample contained over 201,000 matched pairs of scores from 28,000 student-score combinations from 10 school districts in three states.

Methods

First, our data were matched to include valid tests of students in grades 3 through 11 to link individual ACT scale scores and MAP assessment RIT scores for a sample of students who had completed both exams in the same subject. ACT scores were provided by NWEA partnering school districts and individually linked to those students' MAP scores from each testing season, either fall or spring. All valid matched data from the resulting sample (i.e., data with valid scores and linking IDs) were included in the analyses; no attempt was made to rebalance the sample in order to simulate a state- or nationally-representative population.

As was reported in *Using RIT Scores to Predict College Readiness* (NWEA, 2012), NWEA conducted an alignment of the MAP® assessment RIT scales with college readiness benchmarks for EXPLORE, PLAN, and ACT exams. "Visual examinations of scatter plots of the data revealed curvilinear relationships between the MAP scale scores and the ACT scale scores. Consequently, a series of

curvilinear (quadratic) regression models were fitted to the data, using MAP RIT scores as the single predictor of performance on each of the college readiness tests. MAP assessments in reading and language usage were both fit to predictive models of performance on college readiness tests of English and Reading" (NWEA, 2012, p. 2).

MAP mathematics was used to predict mathematics college readiness and MAP reading was used to predict reading college readiness. In all, 18 predictive models were fitted. This ordinary least squares regression was used to determine the strength of the correlation between fourth through 11th grade MAP benchmarks and actual ACT performance in both fall and spring test administrations. Table 3 reports Pearson's r coefficients of correlation at each grade-level benchmark. Correlations ranged from $r = .71$ to $r = .91$, indicating between 50% and 83% of variation in ACT scores predicted by MAP performance. Due to our large sample size, all correlations were highly statistically significant.

Table 3

Correlations Between MAP Benchmarks and 11th Grade ACT Performance

	Mathematics				Reading			
	Fall		Spring		Fall		Spring	
Grade	*n*	*r*	*n*	*r*	*n*	*r*	*n*	*r*
4			1,046	0.75			1,057	0.72
5	1,534	0.78	1,766	0.77	1,528	0.72	1,770	0.74
6	2,046	0.79	2,434	0.81	2,082	0.75	2,425	0.77
7	4,081	0.84	4,446	0.84	4,328	0.78	4,332	0.79
8	5,779	0.86	6,048	0.85	5,980	0.78	6,021	0.77
9	5,990	0.89	7,536	0.87	6,019	0.80	7,422	0.78
10	5,727	0.91	6,502	0.88	5,917	0.79	7,012	0.78
11	2,362	0.89	2,468	0.86	2,386	0.77	2,359	0.77

Note. All coefficients are significant at $p < .001$.

The quadratic regression model was as follows:

$$ACT = \alpha + \mathrm{X}\,\beta + \mathrm{X}^2\,\beta + \mathrm{e}$$

where X = RIT score

Second, the Pearson's chi-squared test was used to evaluate whether the model fit of the data was appropriate. We found the fit to be significant at $p < .001$ for each grade, subject, and season combination for reading and math. Next we ran a set of cross-tabs to determine the proportions of correct predictions for each grade, subject, and season. Mathematics scores were found to have best prediction, with the highest accuracy ranging from 79% to 91%. Reading was approximately the same with a range between 77% and 84%. As expected, the correct prediction percentage increased for students' MAP testing seasons closer to the ACT test season in 11th grade. The use of MAP scores as a

predictor of ACT benchmark scores was found to have a low false-positive prediction rate of 4% or less.

The key fit statistics for the quadratic regression model used for this analysis are provided below in Table 4.

Table 4

Quadratic Regression Model Fit Statistics for Mathematics and Reading Scores

	Math scores		Reading scores	
Grade level	$\chi^2(1)$	Correct prediction	$\chi^2(1)$	Correct prediction
		Fall testing season		
5	415.65	81%	381.90	77%
6	586.34	80%	559.79	77%
7	1501.78	81%	1462.75	79%
8	2385.81	83%	2014.48	79%
9	2987.66	86%	2324.21	81%
10	2982.74	87%	2214.22	81%
11	1133.59	91%	699.07	82%
		Spring testing season		
4	211.03	81%	262.17	79%
5	442.02	81%	483.66	79%
6	804.17	82%	749.96	80%
7	1782.40	82%	1577.65	80%
8	2488.58	83%	2114.01	80%
9	3268.74	83%	2545.29	79%
10	2891.81	85%	2498.10	80%
11	936.39	90%	799.50	84%

Note. All Pearson's chi-squared tests were significant at $p < .001$.

Finally, ROC curve analyses were performed to provide an examination of the predictive accuracy of the benchmarks from our regression model. A ROC curve simply compares the precision and sensitivity of our predictions. For our purposes, we were able to analyze the AUC (area under the curve) precision with which our grade-level benchmark scores were associated with actual student performance in relation to ACT college readiness benchmarks. Findings from the ROC curve analyses are given in Tables 5 and 6.

Table 5

ROC Curve Statistics for Mathematics Score Predicting ACT Benchmark Reached

			Asymptotic 95% confidence interval	
Grade level	AUC	*SE*	Lower bound	Upper bound
		Fall testing season		
5	0.831	0.010	0.811	0.851
6	0.855	0.008	0.839	0.870
7	0.863	0.005	0.852	0.873
8	0.867	0.004	0.858	0.875
9	0.878	0.004	0.869	0.886
10	0.875	0.004	0.867	0.884
11	0.873	0.008	0.858	0.888
		Spring testing season		
4	0.807	0.008	0.792	0.823
5	0.793	0.006	0.781	0.805
6	0.789	0.005	0.779	0.799
7	0.802	0.004	0.794	0.809
8	0.799	0.003	0.792	0.805
9	0.806	0.003	0.800	0.812
10	0.806	0.003	0.799	0.812
11	0.802	0.006	0.790	0.814

Note. ROC = receiver operating characteristic. AUC = area under curve. All statistics were significant at $p < .001$.

Table 6

ROC Curve Statistics for Reading Score Predicting ACT Benchmark Reached

			Asymptotic 95% confidence interval	
Grade level	AUC	*SE*	Lower bound	Upper bound
		Fall testing season		
5	0.866	0.009	0.848	0.884
6	0.867	0.008	0.852	0.882
7	0.893	0.005	0.884	0.902
8	0.912	0.004	0.905	0.919
9	0.929	0.003	0.923	0.935
10	0.938	0.003	0.933	0.944
11	0.945	0.005	0.935	0.955
		Spring testing season		
4	0.823	0.008	0.806	0.839
5	0.813	0.006	0.801	0.825
6	0.805	0.005	0.795	0.815
7	0.814	0.004	0.807	0.822
8	0.820	0.003	0.813	0.826
9	0.825	0.003	0.820	0.831
10	0.826	0.003	0.819	0.832
11	0.812	0.007	0.799	0.826

Note. ROC = receiver operating characteristic. AUC = area under curve. All statistics were significant at $p < .001$.

Results

As shown above, MAP tests were found to be highly correlated to the ACT. The model resulted in the following grade, subject, and season benchmarks given in Table 7. It is important to note that benchmarks for the third grade fall season start at a RIT score of 204 (NWEA 83rd percentile) for mathematics and 199 (NWEA 73rd percentile) for reading.

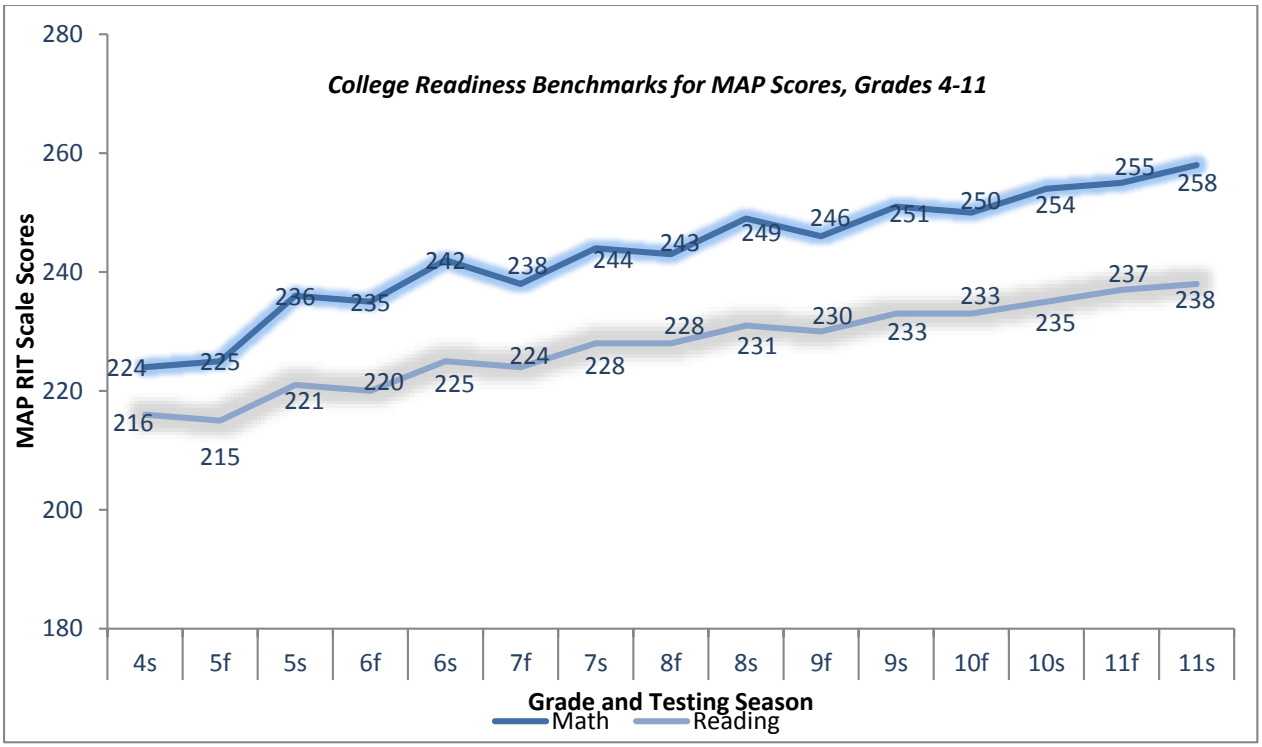
MEASURING GROWTH TOWARD COLLEGE READINESS

Table 7

Predicted College Readiness Benchmarks as MAP Score and Percentile Ranking by Grade

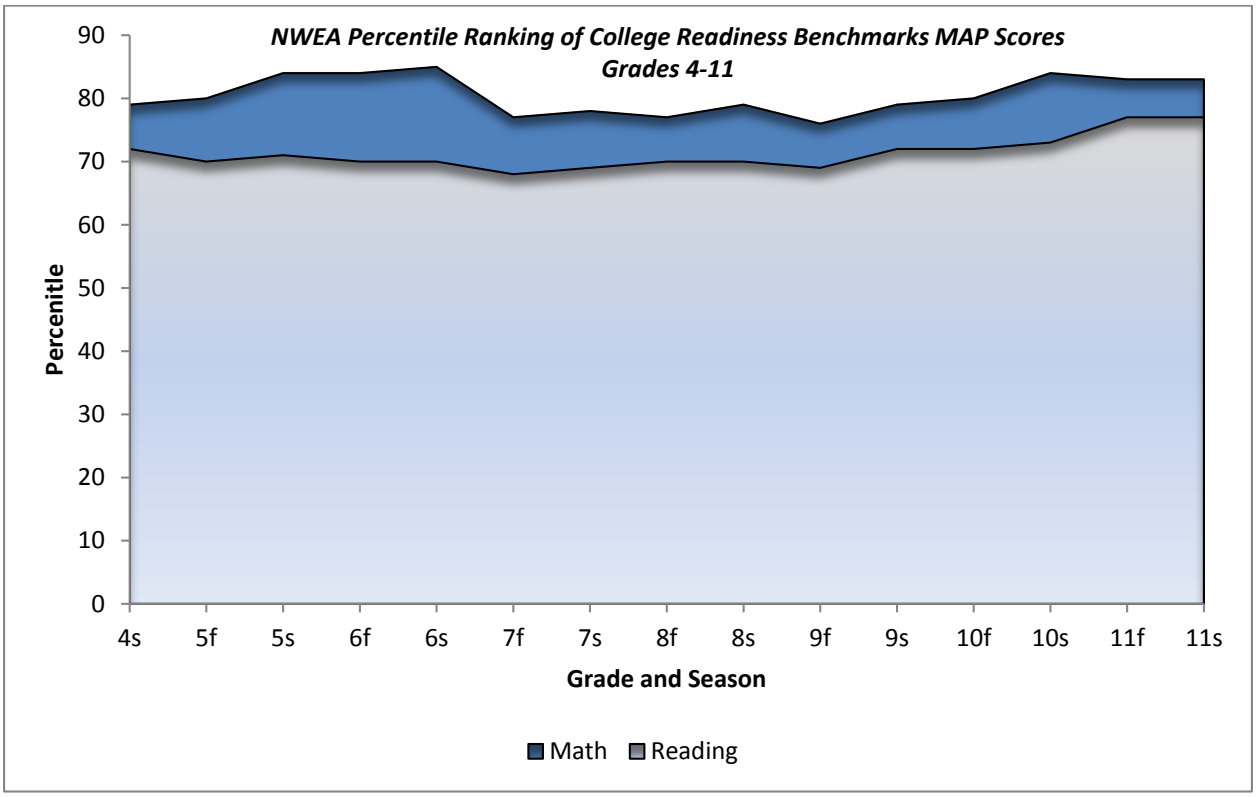
	Mathematics				Reading			
	Fall		Spring		Fall		Spring	
Grade	RIT	Percentile	RIT	Percentile	RIT	Percentile	RIT	Percentile
4			224	79			215	72
5	225	80	236	84	214	70	220	71
6	235	84	242	85	220	70	224	70
7	238	77	244	78	223	68	227	69
8	243	77	249	79	227	70	230	70
9	246	76	251	79	229	69	232	72
10	250	80	254	82	232	72	234	73
11	255	83	258	83	236	77	237	77

Figure 1. MAP Scale Scores for ACT® College Readiness Benchmarks.



MEASURING GROWTH TOWARD COLLEGE READINESS

Figure 2. MAP Scale Scores as Percentile Ranks for ACT® College Readiness Benchmarks.



Finally, we extended our predicted scores beyond a single set of college readiness benchmarks by performing similar analyses toward benchmarks for various types of goal institutions students might choose. Since it is difficult to justify the application of the same set of college-readiness goals for a student attending an open-enrollment university as for one applying in a top ranked university, sets of grade-by-grade benchmarks allow students and teachers to make an early comparison of expected growth toward a variety of higher education goals. Table 8 gives a set of suggested ACT goal scores for Ivy League, top public universities, state universities, and open-enrollment institutions. Figures 3 and 4 illustrate the MAP benchmarks across grades associated with institution type.

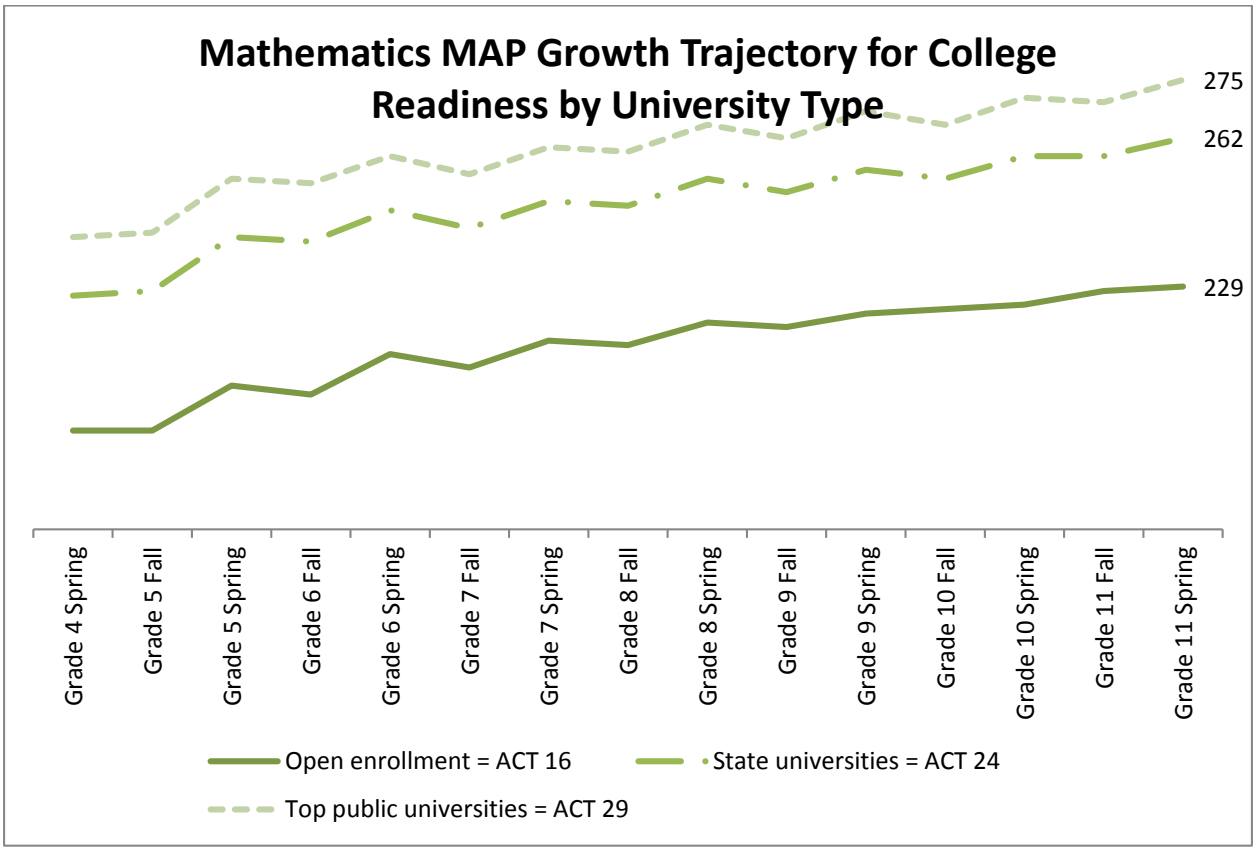
Table 8

Typical ACT Benchmark Scores by Institution Type

Type of institution	ACT score
Top ranked	32*
Top public universities	29
State universities	24
Open-enrollment	16

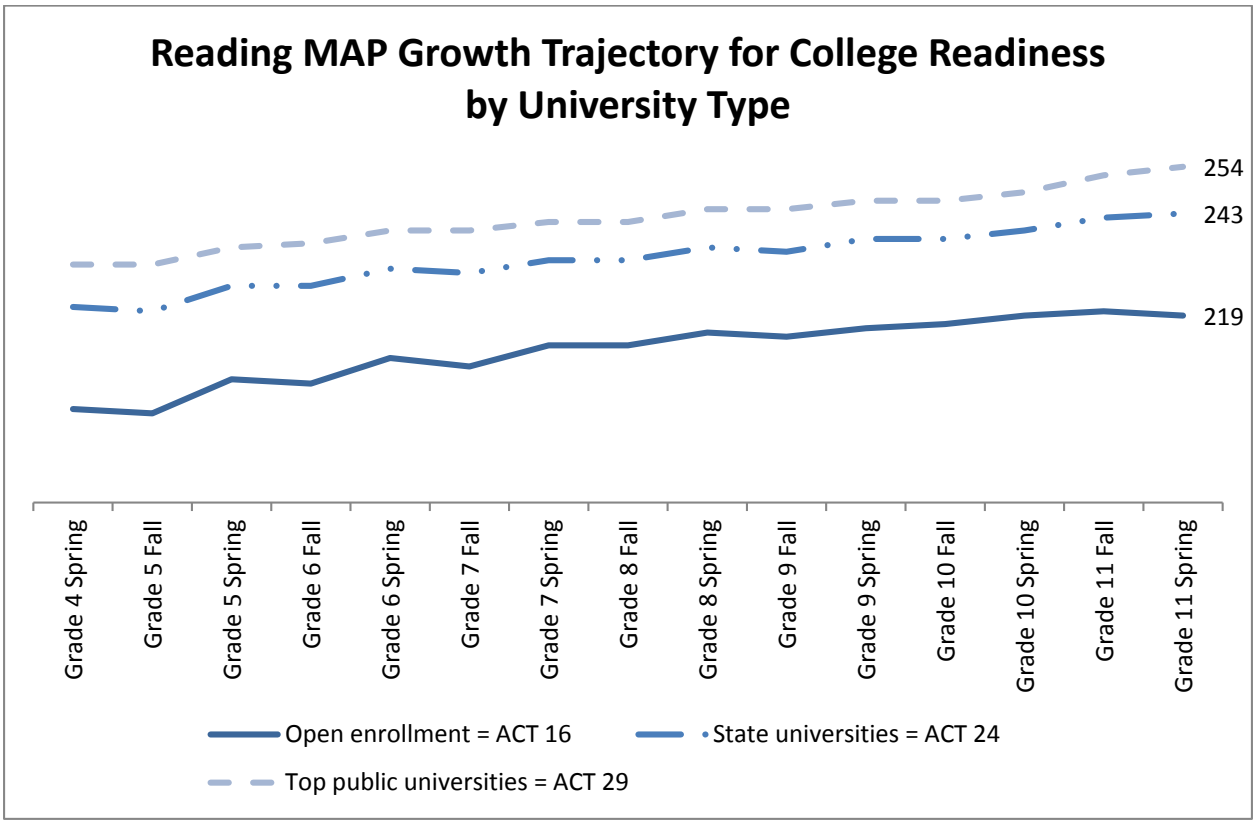
Notes: The scores are average ACT score of incoming freshmen.
**Not illustrated due to small sample size*

Figure 3. Mathematics MAP Growth Trajectory for College Readiness by University Type.



MEASURING GROWTH TOWARD COLLEGE READINESS

Figure 4. Reading MAP Growth Trajectory for College Readiness by University Type.



References

ACT. (2009). *How much growth toward college readiness is reasonable to expect in high school? Issues in college readiness.* Iowa City, IA: ACT, Inc. Retrieved from ERIC database. (ED504656)

ACT. (2010). *Mind the gaps: How college readiness narrows achievement gaps in college success.* Iowa City, IA: ACT, Inc. Retrieved from ERIC database. (ED511829)

Allen, J., & Sconing, J. (2005). *Using ACT assessment scores to set benchmarks for college readiness* (Research Report No. 2005-3). Iowa City, IA: ACT, Inc. Retrieved from ERIC database. (ED489766)

Bettinger, E. P., Evans, B. J., & Pope, D. G. (2011). *Improving college performance and retention the easy way: Unpacking the ACT exam* (NBER Working Paper No. 17119). Retrieved from http://www.nber.org/papers/w17119

Kingsbury Center at Northwest Evaluation Association. (2010). *South Carolina linking study* (Technical Report). Lake Oswego, OR: NWEA. Retrieved from http://www.nwea.org/sites/www.nwea.org/files/resources/SC%20Linking%20Study%20August2010_Revised_Final.pdf

Northwest Evaluation Association. (2008). *RIT scale norms for use with Measures of Academic Progress* (Technical Manual). Portland, OR: NWEA.

Northwest Evaluation Association. (2009). *Technical manual for Measures of Academic Progress and Measures of Academic Progress for Primary Grades.* Portland, OR: NWEA.

Northwest Evaluation Association. (2012). *Using RIT Scores to Predict College Readiness*. Portland, OR: NWEA.

Theaker, R., & Mellema, P. (2010). *Derivation of college-readiness benchmarks* (Technical Report).