Making Connections

April 2016

Can scores on an interim high school reading assessment accurately predict low performance on college readiness exams?

Sharon Koon
Yaacov Petscher
Florida State University

Key findings

This study provides preliminary evidence that scores on an interim reading assessment in grade 9, the Florida Assessments for Instruction in Reading–Florida Standards, can be used to identify students who may score below the college readiness benchmark on the Preliminary SAT/National Merit Scholarship Qualifying Test or ACT Plan in grade 10. The percentage of students not meeting the college readiness benchmark who were correctly identified by the Florida Assessments for Instruction in Reading–Florida Standards as at risk ranged from 81 percent to 96 percent. Classification accuracy depended on the assessment and subject area being predicted. Using scores on an existing interim reading assessment in an early warning system could enable districts to identify at-risk students without additional testing burden, time away from instruction, or cost.

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences
U.S. Department of Education

REL SOUTHEAST
Regional Educational Laboratory
At Florida State University

U.S. Department of Education
John B. King, Jr., *Secretary*

Institute of Education Sciences
Ruth Neild, *Deputy Director for Policy and Research*
Delegated Duties of the Director

National Center for Education Evaluation and Regional Assistance
Joy Lesnick, *Acting Commissioner*
Amy Johnson, *Action Editor*
Sandra Garcia, *Project Officer*

REL 2016–124

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

April 2016

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0011 by Regional Educational Laboratory Southeast administered by the Florida Center for Reading Research, Florida State University. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Koon, S., & Petscher, Y. (2016). *Can scores on an interim high school reading assessment accurately predict low performance on college readiness exams?* (REL 2016–124). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. Retrieved from http://ies.ed.gov/ncee/edlabs.

This report is available on the Regional Educational Laboratory website at http://ies.ed.gov/ncee/edlabs.

Summary

During the 2013/14 school year two Florida school districts sought to develop an early warning system to identify students at risk of low performance on college readiness measures in grade 11 or 12 (such as the SAT or ACT) in order to support them with remedial coursework prior to high school graduation. The districts partnered with Regional Educational Laboratory Southeast to study the extent to which scores on an interim reading assessment in grade 9, the Florida Assessments for Instruction in Reading—Florida Standards (FAIR-FS), could identify students who may score below the college readiness benchmark on the Preliminary SAT/National Merit Scholarship Qualifying Test (PSAT/NMSQT) or ACT Plan in grade 10. Almost all Florida students in grade 10 take one of these assessments, allowing for a more robust sample of students than is available for counterpart assessments in grade 11 or 12 (the SAT and ACT). To conduct the analysis, this study used scores in word recognition, vocabulary knowledge, reading comprehension, and syntactic knowledge from the FAIR-FS; scores in critical reading, math, and writing from the PSAT/NMSQT; and scores in English, reading, science, and math from the ACT Plan.

Classification and regression tree (CART) analyses, which provide an easy-to-interpret "tree" format, were used to classify students as at risk or not at risk of low performance on the PSAT/NMSQT and ACT Plan college readiness measures based on FAIR-FS scores (Berk, 2008; Breiman, Friedman, Olshen, & Stone, 1984; Lewis, 2000). Sensitivity, which is a measure of classification accuracy, was used to judge the accuracy of the prediction models. In this study, sensitivity provides the proportion (or percentage) of students identified as at risk by their FAIR-FS scores among all students who did not meet the college readiness benchmark of interest.

Key findings for the sample include:

- FAIR-FS reading comprehension scores predict PSAT/NMSQT critical reading performance with 89 percent sensitivity.
- FAIR-FS syntactic knowledge scores predict PSAT/NMSQT math performance with 81 percent sensitivity and writing performance with 84 percent sensitivity.
- FAIR-FS syntactic knowledge scores predict ACT Plan English performance with 81 percent sensitivity, reading performance with 84 percent sensitivity, and science performance with 96 percent sensitivity.
- FAIR-FS reading comprehension scores predict ACT Plan math performance with 83 percent sensitivity.

Contents

Why this study?

Identifying students who may not meet college readiness benchmarks prior to high school graduation is critical to addressing their academic deficiencies. Potentially useful identification tools are interim assessments, defined by the Council of Chief State School Officers (2008, pp. 2–3) as:

> …assessments administered multiple times during a school year, usually outside of instruction, to evaluate students' knowledge and skills relative to a specific set of academic goals in order to inform policymaker or educator decisions at the student, classroom, school, or district level. The specific interim assessment designs are driven by the purposes and intended uses, but the results of any interim assessment must be reported in a manner allowing aggregation across students, occasions, or concepts.

Identifying students who may not meet college readiness benchmarks prior to high school graduation is critical to addressing their academic deficiencies

Interim assessments may predict student college readiness in reading. As noted by Porter and Polikoff (2012, p. 411), "readiness predictors based on specific skills and student performance would undoubtedly be more useful for planning policies to improve instruction and student performance" than simply knowing what percentage of students met the benchmark. If specific skills can be identified as those most likely to indicate college readiness, early identification of deficiencies in those skills could enable teachers and schools to provide targeted intervention. However, research in this area is limited.

Interim reading assessments may also predict student readiness in other subject areas. There is a high correlation between performance in reading and performance in math (Dorans, 2000; Thurber, Shinn, & Smolkowski, 2002) and science (Dorans, 2000). For example, Dorans found a .75 correlation between ACT science performance and each of the other SAT and ACT subject area scores. Additionally, literacy skills provide a critical foundation for students' overall academic success (Whitehurst & Lonigan, 2001), possibly because students who read well read more, thus acquiring more knowledge in various academic domains (Cunningham & Stanovich, 1998).

During the 2013/14 school year two Florida school districts (referred to as districts 1 and 2) sought to develop an early warning system to identify students at risk of low performance on college readiness measures in grade 11 or 12 (such as the SAT or ACT) in order to support them with remedial coursework prior to high school graduation. The districts partnered with Regional Educational Laboratory Southeast to study the extent to which scores on an interim reading assessment in grade 9, the Florida Assessments for Instruction in Reading–Florida Standards (FAIR-FS), could identify students who may score below the college readiness benchmark on the Preliminary SAT/National Merit Scholarship Qualifying Test (PSAT/NMSQT) or the ACT Plan in grade 10. Almost all Florida students in grade 10 take one of these assessments, allowing for a more robust sample of students than is available for their counterpart assessments in grade 11 or 12 (the SAT and ACT).

What the study examined

Two research questions guided the study:

- How do scores on the FAIR-FS in grade 9 predict performance on the PSAT/NMSQT in grade 10?
- How do scores on the FAIR-FS in grade 9 predict performance on the ACT Plan in grade 10?

FAIR-FS data are from five high schools in district 1 and two high schools in district 2; they include scores in word recognition, vocabulary knowledge, reading comprehension, and syntactic knowledge. PSAT/NMSQT data are from district 1, and ACT Plan data are from district 2. FAIR-FS is described in box 1, and the PSAT/NMSQT and ACT Plan are described in box 2.

How scores on the FAIR-FS in grade 9 predict performance on the PSAT/NMSQT and ACT Plan in grade 10 was examined using classification and regression tree analyses

How scores on the FAIR-FS in grade 9 predict performance on the PSAT/NMSQT and ACT Plan in grade 10 was examined using classification and regression tree (CART) analyses. CART is predictive modeling that presents the results in an easy-to-interpret tree format that classifies students as at risk or not at risk of a future outcome based on a set of if-then statements (Berk, 2008; Breiman et al., 1984; Lewis, 2000). Previous studies

Box 1. About the Florida Assessments for Instruction in Reading—Florida Standards

After grade 3, students' reading comprehension is determined by their ability to decode words and their oral language skills (Storch & Whitehurst, 2002). Florida Assessments for Instruction in Reading—Florida Standards (FAIR-FS) is a multivariate measure of four important component skills of reading: word recognition, vocabulary knowledge, reading comprehension, and syntactic knowledge (Foorman, Petscher, & Schatschneider, 2015). Performance in each skill is reported using a developmental scale, with scores ranging from 200 to 800, with a mean of 500 and a standard deviation of 100.

The word recognition test, referred to as a task, asks students to identify the correctly written form of a word pronounced by the computer. The vocabulary knowledge task measures recognition of morphological patterns in words, which can generalize to recognition of new words (rather than measuring isolated vocabulary words), by having students complete a sentence with one of three morphologically related words (for example, "In some states you can get a driver's [permission, permissive, permit] when you are fourteen years old"). The task was designed to inform instruction on the metalinguistic aspects of vocabulary essential to building vocabulary depth rather than instruction that favors memorization of a large corpus of isolated words in an attempt to increase vocabulary breadth.

Scores on the word recognition and vocabulary knowledge tasks determine which of several complex text passages are used to assess reading ability in the reading comprehension task. For that task, students silently read up to three passages and answer seven to nine multiple-choice questions written to the Language Arts Florida Standards.

The syntactic knowledge task measures understanding of the grammatical relationships among words and sentences. In this task, students listen to a sentence being read and select from a dropdown menu which of three connectives, pronoun references, or verb tenses best completes a sentence (for example, "Pizza is one of my very favorite foods, [although, as, when] we only get to eat it on special occasions").

Box 2. About the Preliminary SAT/National Merit Scholarship Qualifying Test and the ACT Plan

To help increase college readiness, Florida law requires that all grade 10 students (unless exempted by parent request) take either the Preliminary SAT/National Merit Scholarship Qualifying Test (PSAT/NMSQT) or the ACT Plan, which serve as preliminary measures of student performance on the SAT or ACT. The PSAT/NMSQT scale of 20–80 is comparable to the SAT scale of 200–800 (College Board, 2013), and an ACT Plan composite score (on a scale of 1–32) can be used to estimate an ACT composite score range based on the applicable testing season and student's grade level. Correlations between the ACT Plan and the ACT range from .60 to .72 for English, reading, science, and math (ACT, 2013).

Grade 10 students who achieve a PSAT/NMSQT score of 42 in critical reading, 44 in math, and 42 in writing meet college and career readiness benchmarks, which are associated with a 65 percent probability of obtaining a freshman year grade point average of a B– or higher (College Board, 2011, 2013). Students who achieve an ACT Plan score of 15 in English, 18 in reading, 20 in science, and 19 in math are considered ready for college; the ACT Plan scores are associated with a 50 percent chance of earning a B or higher or a 75 percent chance of earning a C or higher in the corresponding first-year English composition, introductory social science, biology, or college algebra course (ACT, 2013, 2014).

Florida student performance on the PSAT/NMSQT and ACT Plan lags the nation (ACT, 2014; College Board, 2014). Approximately 32 percent of Florida grade 10 students who took the PSAT/NMSQT in 2013/14 met all subject area benchmarks, compared with 41 percent nationwide, and roughly 18 percent of Florida grade 10 students who took the ACT Plan met the college readiness benchmarks in all subject areas, compared with 20 percent nationwide.

have found CART results to be consistent with those from logistic regression (Koon, Petscher, & Foorman, 2014) and easier for a nonstatistical audience to understand because of CART's graphic format.

Dominance analysis (Azen & Budescu, 2003) was used to determine the relative importance of each FAIR-FS task in predicting PSAT/NMSQT and ACT Plan performance and to validate the CART results. The dominance analyses served as an informative complement to the CART analyses. Details on the data sources and analyses are in appendix A.

Traditional measures of classification accuracy, including sensitivity and specificity, are provided for all prediction models (box 3). However, sensitivity values were used to judge the performance of the prediction models due to the importance of identifying students who are at risk. Researchers have proposed different threshold values for sensitivity; many look for levels of at least .80 (Piasta, Petscher, & Justice, 2012), with some recommending at least .90 (Compton, Fuchs, Fuchs, & Bryant, 2006; Jenkins, Hudson, & Johnson, 2007). Based on these guidelines, a sensitivity value of .80 or higher was used in this study.

Box 3. Measures of classification accuracy

Traditional measures of classification accuracy can be derived from a 2×2 classification table that provides counts of individuals in four categories (Schatschneider, Petscher, & Williams, 2008). In this study students are categorized based on their score on an interim assessment, the Florida Assessments for Instruction in Reading—Florida Standards, and an outcome assessment, the Preliminary SAT/National Merit Scholarship Qualifying Test or the ACT Plan (see table).

Sample 2×2 classification table

	Outcome assessment (Preliminary SAT/National Merit Scholarship Qualifying Test or ACT Plan)	
Interim assessment (Florida Assessments for Instruction in Reading–Florida Standards)	Fail	Pass
At risk	A: True positive	B: False positive
Not at risk	C: False negative	D: True negative

Source: Authors' illustration.

In addition to the overall classification accuracy rate (proportion of students who were correctly identified as meeting or not meeting the college readiness benchmark on the outcome assessment), four standard measures of classification accuracy were applied to determine how accurately the analysis identifies students at risk. The first measure, sensitivity, is the proportion of students identified as at risk on the interim assessment among all students who fail the outcome assessment—the number of true positives—divided by the sum of true positives and false negatives (A/[A+C]). The second measure, specificity, is the proportion of students identified as not at risk among all students who pass the outcome assessment—or the number of true negatives divided by the sum of true negatives and false positives (D/[D+B]). The third measure, positive predictive power, is the proportion of students who fail the outcome assessment among all students who are identified as at risk on the interim assessment—or the number of true positives divided by the sum of true positives and false positives (A/[A+B]). The fourth measure, negative predictive power, is the proportion of students who pass the outcome assessment among all students who are identified as not at risk on the interim assessment—or the number of true negatives divided by the sum of false negatives and true negatives (D/[C+D]).

What the study found

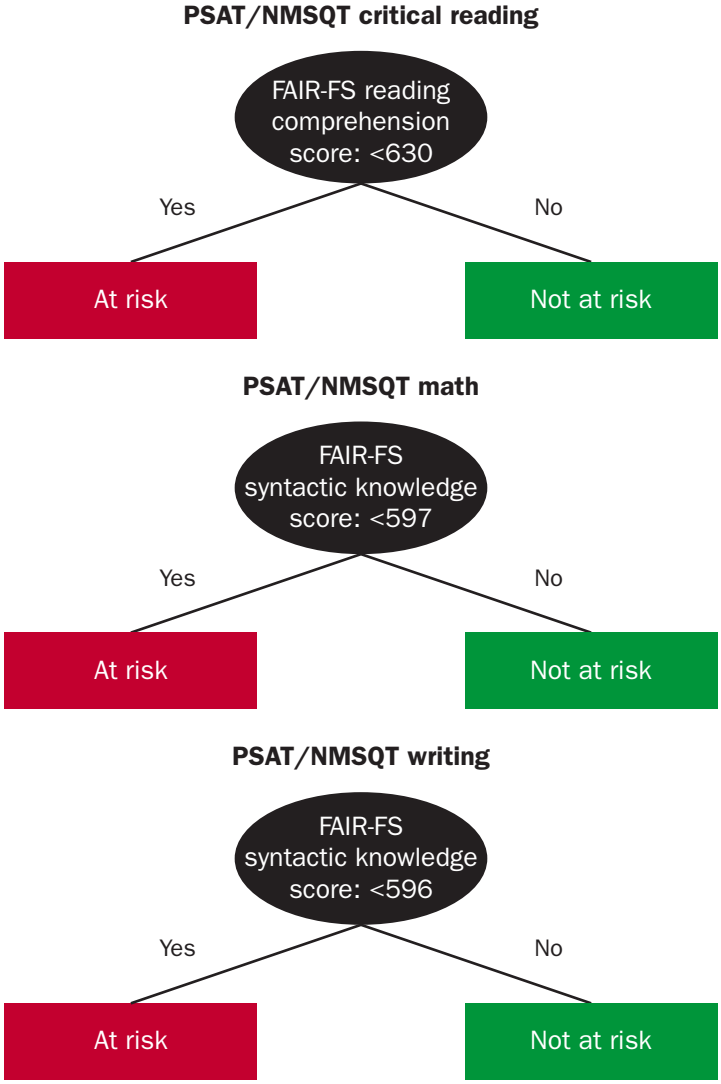
The findings indicate that FAIR-FS scores can predict performance on college readiness assessments with acceptable sensitivity.

FAIR-FS reading comprehension scores predict PSAT/NMSQT critical reading performance with 89 percent sensitivity, and FAIR-FS syntactic knowledge scores predict PSAT/NMSQT math performance with 81 percent sensitivity and writing performance with 84 percent sensitivity

The CART analyses identified only one decision rule in each subject area for classifying students as at risk or not at risk of not reaching the college readiness benchmark on the PSAT/NMSQT (figure 1):

- Critical reading: students who score less than 630 in reading comprehension on the FAIR-FS are identified as at risk.

Figure 1. Classification and regression tree model decision rules for classifying Florida district 1 students as at risk or not at risk of not reaching the college readiness benchmark on the PSAT/NMSQT based on FAIR-FS score



The CART analyses identified only one decision rule in each subject area for classifying students as at risk or not at risk of not reaching the college readiness benchmark on the PSAT/NMSQT

PSAT/NMSQT is Preliminary SAT/National Merit Scholarship Qualifying Test. FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards.

Note: FAIR-FS scores range from 200 to 800, with a mean of 500 and a standard deviation of 100.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 1 (see appendix A).

- Math: students who score less than 597 in syntactic knowledge on the FAIR-FS are identified as at risk.
- Writing: students who score less than 596 in syntactic knowledge on the FAIR-FS are identified as at risk.

The overall classification accuracy rate of the decision rules was 75–83 percent when tested with a validation sample of students (table 1). The sensitivity rate (81–89 percent) was above the minimum standard of 80 percent. The specificity rate was much lower (66–80 percent), which impacted the overall classification accuracy rate. This study considered it more important to judge the prediction models based on sensitivity rates because

Table 1. Preliminary SAT/National Merit Scholarship Qualifying Test classification and regression tree analysis results for Florida district 1 students (percent)

PSAT/NMSQT subject area	Sensitivity rate	Specificity rate	Positive predictive power rate	Negative predictive power rate	Overall classification accuracy rate
Critical reading	89	69	79	83	80
Math	81	66	79	70	75
Writing	84	80	89	72	83

PSAT/NMSQT is Preliminary SAT/National Merit Scholarship Qualifying Test.

Note: *n* = 210.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 1 (see appendix A).

the intention is to identify students at risk of low performance on the PSAT/NMSQT. The positive predictive power rate generally exceeded the negative predictive power rate.

A supplementary dominance analysis triangulating the CART results supported the decision rules (see appendix A). The dominance analysis found that FAIR-FS reading comprehension score is the dominant predictor of PSAT/NMSQT reading performance and that FAIR-FS syntactic knowledge and reading comprehension scores are dominant predictors of PSAT/NMSQT math and writing performance (table 2).

As with the PSAT/NMSQT results, the CART analyses identified only one decision rule in each subject area for classifying students as at risk or not at risk of not reaching the college readiness benchmark on the ACT Plan

FAIR-FS syntactic knowledge scores predict ACT Plan English performance with 81 percent sensitivity, reading performance with 84 percent sensitivity, and science performance with 96 percent sensitivity, and FAIR-FS reading comprehension scores predict ACT Plan math performance with 83 percent sensitivity

As with the PSAT/NMSQT results, the CART analyses identified only one decision rule in each subject area for classifying students as at risk or not at risk of not reaching the college readiness benchmark on the ACT Plan (figure 2):

- English: students who score less than 592 in syntactic knowledge on the FAIR-FS are identified as at risk.
- Reading: students who score less than 620 in syntactic knowledge on the FAIR-FS are identified as at risk.
- Science: students who score less than 680 in syntactic knowledge on the FAIR-FS are identified as at risk.
- Math: students who score less than 710 in reading comprehension on the FAIR-FS are identified as at risk.

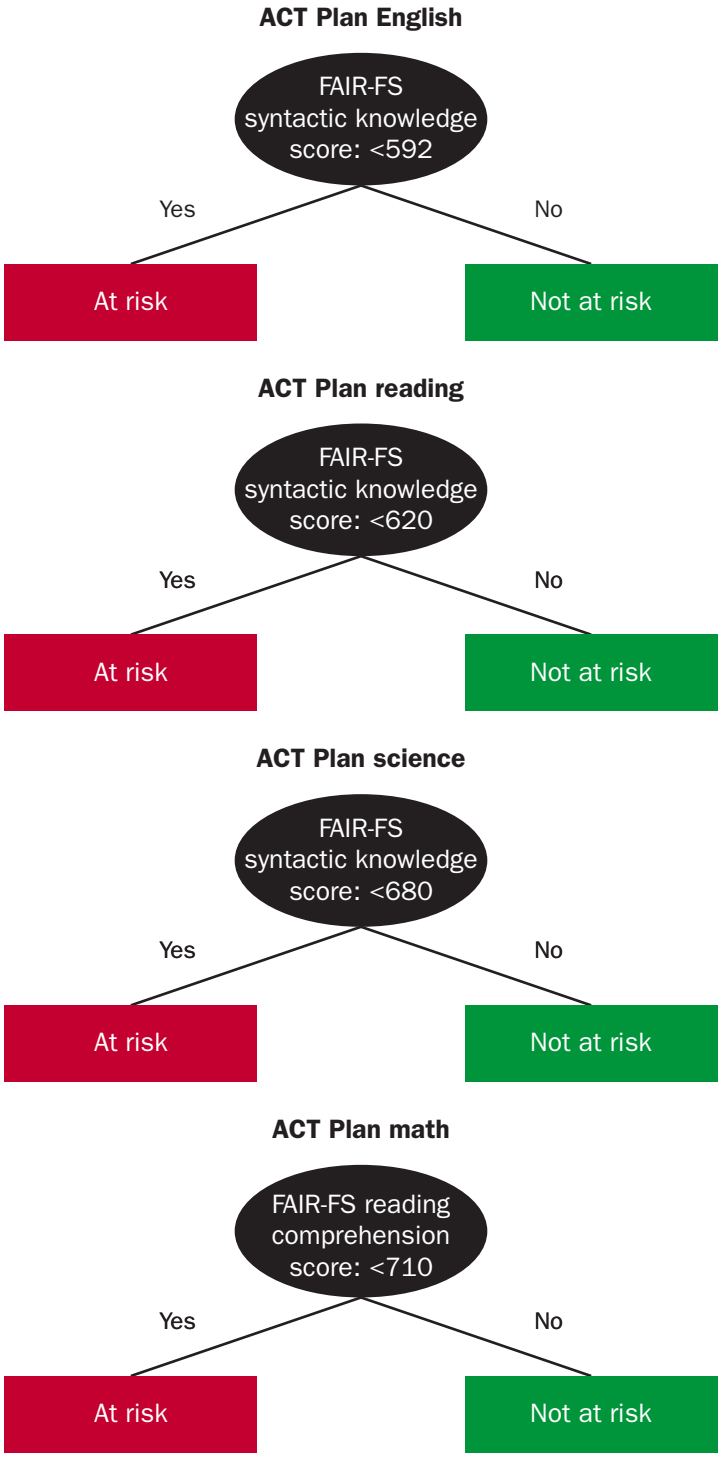
Table 2. Correspondence of predictor importance from dominance analysis and classification and regression tree analysis for Florida district 1 students based on Preliminary SAT/National Merit Scholarship Qualifying Test results

PSAT/NMSQT subject area	Primary FAIR FS predictor in CART analysis	Dominant FAIR FS predictor in dominance analysis
Critical reading	Reading comprehension	Reading comprehension
Math	Syntactic knowledge	Reading comprehension Syntactic knowledge
Writing	Syntactic knowledge	Reading comprehension Syntactic knowledge

PSAT/NMSQT is Preliminary SAT/National Merit Scholarship Qualifying Test. FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards. CART is classification and regression tree.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 1 (see appendix A).

Figure 2. Classification and regression tree model decision rules for classifying Florida district 2 students as at risk or not at risk of not reaching the college readiness benchmark on the ACT Plan based on FAIR-FS score



FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards.

Note: FAIR-FS scores range from 200 to 800, with a mean of 500 and a standard deviation of 100.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 2 (see appendix A).

The overall classification accuracy rate of the decision rules was 71–81 percent when tested with a validation sample of students (table 3). The sensitivity rate (81–96 percent) was above the minimum standard of 80 percent, while the specificity rate was much lower (41–76 percent). The positive predictive power rate ranged from 59 percent to 89 percent, and the negative predictive power rate ranged from 44 percent to 85 percent.

A supplementary dominance analysis triangulating the CART results supported some of the decision rules. The dominance analysis found that FAIR-FS syntactic knowledge score is the dominant predictor of ACT Plan reading and math performance (table 4). This finding is consistent with the CART results for ACT Plan reading but seemingly inconsistent with the CART results for ACT Plan math (in which the decision rule is based on reading comprehension score). However, the use of reading comprehension as a splitting variable likely results from the high correlation between FAIR-FS reading comprehension and FAIR-FS syntactic knowledge in this sample of students (r = .80).

A supplementary dominance analysis triangulating the CART results found that FAIR-FS syntactic knowledge score is the dominant predictor of ACT Plan reading and math performance

Syntactic knowledge score is the first surrogate variable in the CART results for predicting ACT Plan math performance. Surrogate variables are identified as alternatives to the specified variable (Therneau & Atkinson, 2015). The alternatives can be used when an individual's score on the specified variable is missing. For example, the syntactic knowledge score could form the basis for a decision if a student is missing the reading comprehension score. The alternatives are likely less efficient than the original scores because the most optimal rule is the one specified in the decision tree. In this study, using syntactic

Table 3. ACT Plan classification and regression tree analysis results for Florida district 2 students (percent)

ACT Plan subject area	Sensitivity rate	Specificity rate	Positive predictive power rate	Negative predictive power rate	Overall classification accuracy rate
English	81	65	59	85	71
Reading	84	76	89	67	81
Science	96	48	79	85	80
Math	83	41	81	44	73

Note: n = 70.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 2 (see appendix A).

Table 4. Correspondence of predictor importance from dominance analysis and classification and regression tree analysis for Florida district 2 students based on ACT Plan results

ACT Plan subject area	Primary FAIR FS predictor in CART analysis	Dominant FAIR FS predictors in dominance analysis
English	Syntactic knowledge	None
Reading	Syntactic knowledge	Syntactic knowledge
Science	Syntactic knowledge	None
Math	Reading comprehension	Syntactic knowledge

FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards. CART is classification and regression tree.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 2 (see appendix A).

knowledge score as an alternative to reading comprehension score results in the same risk classification for 93 percent of the sample. The dominance analysis did not identify any completely dominant predictors for ACT Plan English or science performance.

Implications of the study findings

This study provides preliminary evidence that FAIR-FS scores in grade 9 could be used to create an early warning system to identify students whose PSAT/NMSQT or ACT Plan performance in grade 10 may be below the college readiness benchmark. Using FAIR-FS scores in an early warning system could enable districts to identify at-risk grade 9 students without an additional testing burden, time away from instruction, or additional cost.

Limitations of the study

This study provides preliminary evidence that FAIR-FS scores in grade 9 could be used to create an early warning system to identify students whose PSAT/NMSQT or ACT Plan performance in grade 10 may be below the college readiness benchmark

This study has two important limitations. First, the analyses are based on a single year's results, which does not allow for confirming the stability of the results or performing a longitudinal analysis. This limitation is especially relevant to FAIR-FS scores, which were drawn from the assessment's field-test sample. Second, the findings for each college readiness assessment are based on a single school district: district 1 for the PSAT/NMSQT and district 2 for the ACT Plan. The analyses could be replicated statewide to verify the stability of the models and the generalizability of the results to the larger Florida student population.

Both the College Board and ACT have implemented new assessments, replacing the versions of the PSAT/NMSQT and ACT Plan assessments used in this study (ACT, 2015; College Board, 2015). New analyses would be required to determine whether the study results remain valid with the new assessments.

PSAT/NMSQT and ACT Plan performance serves as a proxy measure of college readiness due to the complexity of identifying all the factors that determine whether a student is truly ready for success in college. District decisionmakers should carefully consider other issues, such as the implications of over- and under-identification, student access to the core curriculum and other typical activities at the school, and the school process for determining when a student may successfully exit an intervention.

Appendix A. Data sources and methodology

This appendix provides detailed information on the study's data sources and methodology.

Data sources

Florida Assessments for Instruction in Reading—Florida Standards (FAIR-FS) data were obtained from an archive of data on 1,853 grade 9 students in five high schools in district 1 and 480 grade 9 students in two high schools in district 2. The same student sample was previously used in an experimental linking study to estimate item parameters and a vertical scale for the FAIR-FS reading component skills. That study was part of Florida State University's subaward from the Educational Testing Service's assessment grant in the Institute of Education Sciences National Center for Educational Research's Reading for Understanding initiative (Sabatini, PI; Award number R305F100005, https://ies.ed.gov/ncer/projects/grant.asp?ProgID=62&grantid=987). The linking study conducted purposeful sampling of schools so that students within the sampled schools in each district would correspond to (or be similar to) the district's student demographic profile.

The FAIR-FS was administered in December 2012 and January 2013 in district 1 and district 2 as part of the linking study sample (which also served as the FAIR-FS field test). Although all students statewide take the Preliminary SAT/National Merit Scholarship Qualifying Test (PSAT/NMSQT) or ACT Plan in grade 10, the availability of FAIR-FS scores was limited to districts 1 and 2 in 2012/13, narrowing the study to those districts.

PSAT/NMSQT data were obtained from district 1, where the test was administered to all students in grades 8–10 in mid-October 2013. ACT Plan data were obtained from district 2, where the test was administered at the end of October 2013. District 2 is one of a few districts in Florida that administers the ACT Plan to all students in grade 10 to meet the statutory requirement to administer either the PSAT/NMSQT or ACT Plan.

Preliminary analyses

The initial district 1 dataset included 1,853 students. Some 250 students were dropped because of missing PSAT/NMSQT scores, and 7 students were dropped because of missing data on all FAIR-FS tasks. Multiple imputation with SAS 9.4 software created a dataset with complete cases for all predictor variables. This approach was justified because the FAIR-FS missing data were considered missing completely at random—the result of the planned missing data design of the linking study. The mean imputed value (after 20,000 imputations) was used for missing values because there is no accepted procedure for analyzing and summarizing classification trees generated from multiple imputed files. A stratified random sample of 1,100 students was selected from the dataset to approximate the racial/ethnic minority distribution of district 1 grade 9 students. An analysis of univariate and multivariate outliers resulted in one student being dropped. The final dataset included 1,099 students.

The same approach was used with the district 2 dataset. Of the initial 480 students, 1 was dropped because of missing ACT Plan scores. Again, 20,000 imputations generated mean imputed values for missing data. A stratified random sample of 300 students was selected, and 4 of them were dropped after being identified as outliers. The final data set included 296 students.

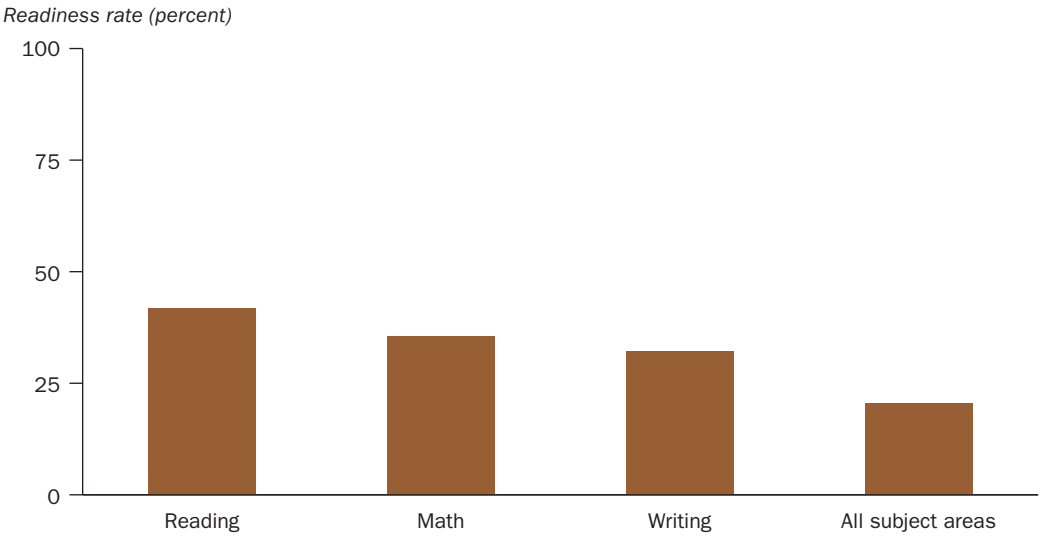
Classification and regression tree analyses

FAIR-FS task scores were used in a series of classification and regression tree (CART) models to predict performance on the PSAT/NMSQT and ACT Plan. The CART model classifies individuals into mutually exclusive subgroups using a nonparametric approach that results in a classification tree (Breiman et al., 1984). It does this by identifying the best predictors and predictor levels that most efficiently split the sample into the most homogeneous subgroups of individuals who are identified as at risk or not at risk based on their observed scores. A variable may appear in the CART model multiple times because the search for the single variable that will result in the best split to the data includes all variables at each split (Therneau & Atkinson, 2015).

The PSAT/NMSQT and ACT Plan scores were dummy-coded to indicate whether a student met the college readiness benchmark for each subject area. Scores at or above the benchmark were coded 1 for not at risk and scores below the benchmark were coded 0 for at risk. Approximately 21 percent of the district 1 sample and 9 percent of the district 2 sample met the college readiness benchmark in all subject areas (figures A1 and A2).

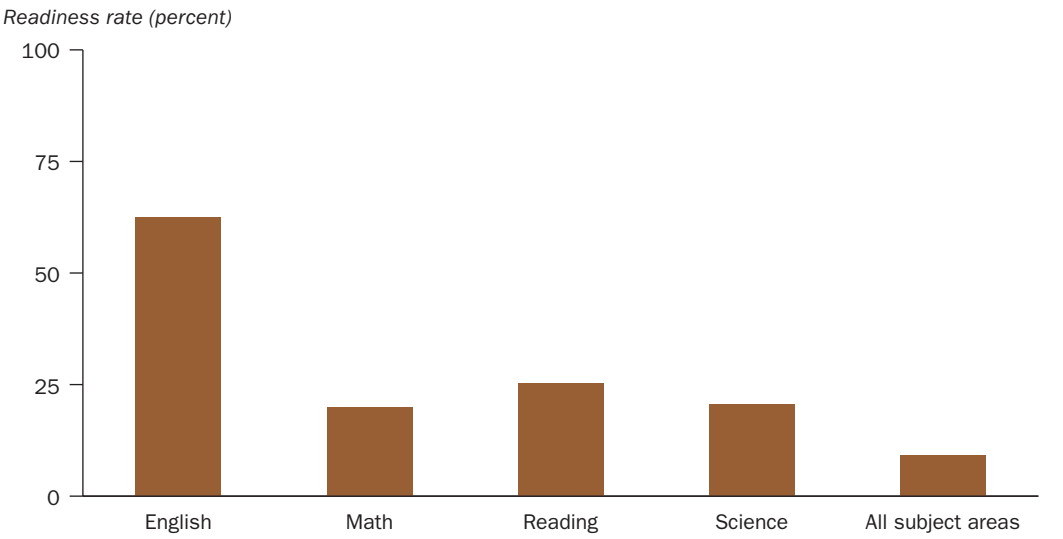
The final datasets for each grade were then split into two. The calibration dataset (used to build the CART models) consisted of a random sample of 80 percent of the students in each grade. The validation dataset (used to test the CART models) consisted of the remaining 20 percent. CART analyses were run using the Recursive Partitioning and Regression Trees package (rpart; R 3.0.1 package).

Figure A1. College readiness of Florida students in the district 1 sample, overall and by subject area, 2013/14 (*n* = 1,099)



Source: Authors' analysis of data from district 1.

Figure A2. College readiness of Florida students in the district 2 sample, overall and by subject area, 2013/14 (*n* = 296)



Source: Authors' analysis of data from district 2.

The initial model for each outcome of interest was fit using all FAIR-FS scores as predictors, represented generally by

Outcome ~ Word recognition score + Vocabulary knowledge score +
Reading comprehension score + Syntactic knowledge score

in addition to specifying a minimum split size of three students and tenfold cross-validation for evaluating the quality of the prediction tree and determining the appropriate minimum complexity parameter for pruning the tree (Breiman et al., 1984).

As is the case with other statistical methods, the principle of parsimony is applicable to CART models. This principle suggests that the best model is often the simplest model that fits the data. In a CART model this principle is applied by pruning the classification tree using model specifications (such as the minimum split size included in the initial model) so that the resulting tree is not overfit to the data. As such, a minimum reduction in the cross-validation relative error (that is, a minimum complexity parameter) was added to the model specifications in a revised model. The minimum complexity parameter specifies the minimum decrease in the overall lack of fit that must result from an additional split. The value selected for the minimum complexity parameter was the value resulting in the fewest number of splits with a cross-validation relative error less than one standard error above the minimum cross-validation relative error (Therneau, Atkinson, & Ripley, 2014). Plots of the cross-validation relative error against minimum complexity parameter values were consulted for this decision. For all models, adding the recommended minimum complexity parameter resulted in a pruned tree with one split. The pruned trees with the final classification rules are shown in figures 1 and 2 in the main report.

The classification rules were applied to the validation dataset to predict group membership and to derive the classification tables used to calculate sensitivity, specificity, and the other measures of classification accuracy (see box 3 in the main report).

Dominance analyses

Dominance analysis (Azen & Budescu, 2003) can be used to identify the most useful predictor in a set of predictors in either an exploratory or confirmatory approach. This study used dominance analysis to determine the relative importance of each FAIR-FS task in predicting scores on the PSAT/NMSQT and ACT Plan and to judge the validity of the CART results. The dominance analyses provide an informative complement to the CART analyses that present a nonparametric approach to predicting risk classification.

Dominance analysis provides complete, conditional, and general measures of importance for each predictor in a regression framework. These measures of importance are an advantage over multiple regression results, which do not provide accepted measures of relative importance when correlated predictors are compared because of the inability of multiple regression analysis to appropriately partition variance to the various predictors. Because dominance analysis belongs to the regression framework, it assumes that the relationships between the predictors and the outcome are linear, the errors are normally and independently distributed, and the error variance is constant. In addition, issues of multicollinearity and influential observations can impact the results. The dominance probability macro (Azen & Budescu, 2003) was used to execute dominance analysis in SAS 9.4 software.

Dominance analysis uses the change in model fit (that is, *R*-squared) to define a predictor's contribution and compares predictor contributions across all possible subset models. Specifically, the technique examines all combinations of predictors to determine an order of importance measured by a variable's reduction of error in predicting the criterion (Azen & Budescu, 2003). The number of subset models is equal to $2^p - 1$, where p is the number of predictors.

Scale scores on the college readiness assessments served as the outcome measures, and FAIR-FS scores served as predictors, represented by

$$Y_i = B_0 + B_1(\textit{Word recognition})_i + B_2(\textit{Vocabulary knowledge})_i + B_3(\textit{Reading comprehension})_i + B_4(\textit{Syntactic knowledge})_i$$

where Y_i is the scale score for student i on the outcome assessment (that is, the PSAT/NMSQT or ACT Plan) and is a function of the intercept (B_0), slopes (B_1, B_2, B_3, B_4), and residual (e_i).

Seven dominance analyses were conducted to examine each PSAT/NMQST and ACT Plan outcome separately. Correlations between the predictor and outcome variables show strong, positive bivariate relationships (tables A1 and A2) supporting the use of all variables in each model.

Table A1. Correlations, means, and standard deviations for Preliminary SAT/ National Merit Scholarship Qualifying Test and Florida Assessments for Instruction in Reading—Florida Standards for Florida district 1 students, 2013/14 (*n* = 1,099)

Assessment and subject area	PSAT/NMSQT			FAIR FS			
	Reading	Math	Writing	Reading comprehension	Syntactic knowledge	Vocabulary knowledge	Word recognition
PSAT/NMSQT							
Reading	1.00						
Math	.68	1.00					
Writing	.74	.66	1.00				
FAIR-FS							
Reading comprehension	.69	.55	.64	1.00			
Syntactic knowledge	.65	.56	.66	.67	1.00		
Vocabulary knowledge	.60	.46	.56	.57	.65	1.00	
Word recognition	.61	.47	.60	.61	.63	.60	1.00
Mean	40.06	40.55	37.91	591.32	558.36	560.09	543.56
Standard deviation	9.18	8.89	8.98	122.96	93.23	66.70	93.00

PSAT/NMSQT is Preliminary SAT/National Merit Scholarship Qualifying Test. FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards.

Note: All correlations are significant at the .001 level (two-tailed).

Source: Authors' analysis of data from the Florida Center for Reading Research and district 1.

Table A2. Correlations, means, and standard deviations for ACT Plan and Florida Assessments for Instruction in Reading—Florida Standards for Florida district 2 students, 2013/14 (*n* = 296)

Assessment and subject area	ACT Plan				FAIR FS			
	English	Reading	Science	Math	Reading comprehension	Syntactic knowledge	Vocabulary knowledge	Word recognition
ACT Plan								
English	1.00							
Reading	.66	1.00						
Science	.68	.66	1.00					
Math	.63	.57	.69	1.00				
FAIR–FS								
Reading comprehension	.64	.61	.59	.53	1.00			
Syntactic knowledge	.67	.65	.60	.58	.80	1.00		
Vocabulary knowledge	.66	.57	.58	.53	.73	.67	1.00	
Word recognition	.62	.59	.54	.46	.74	.67	.69	1.00
Mean	15.60	15.90	17.20	16.36	609.21	576.60	587.45	565.81
Standard deviation	3.74	3.93	3.45	3.50	118.61	95.87	60.92	85.01

FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards.

Note: All correlations are significant at the .001 level (two-tailed).

Source: Authors' analysis of data from the Florida Center for Reading Research and district 2.

PSAT/NMSQT dominance analyses. The results of the dominance analyses for PSAT/NMSQT performance are reported in tables A3–A5. The tables provide the *R*-squared value for each subset model considered and the unique variance contribution added to that subset model by the variables not included in the estimation of *R*-squared. Standardized regression coefficients are provided for each predictor.

Examination of the variance components for PSAT/NMSQT reading performance shows that the contribution of FAIR-FS reading comprehension scores is greater than that of all other predictors (table A3). On average, FAIR-FS reading comprehension scores uniquely explain 20 percent of the variance in PSAT/NMSQT reading performance across all subset models. In predicting PSAT/NMSQT math performance, FAIR-FS reading comprehension and syntactic knowledge scores contribute comparable amounts of unique variance (table A4). On average, FAIR-FS reading comprehension scores uniquely explain 13 percent of the variance in PSAT/NMSQT math performance across all subset models compared

Table A3. Dominance analysis results for Preliminary SAT/National Merit Scholarship Qualifying Test reading performance for Florida district 1 students, 2013/14 (*n* = 1,099)

		Additional contribution of FAIR FS score in:			
Subset model	*R* squared	Reading comprehension	Syntactic knowledge	Vocabulary knowledge	Word recognition
Null and *k*=0 average	na	0.48	0.43	0.36	0.37
Reading comprehension	0.48	na	0.06	0.06	0.05
Syntactic knowledge	0.43	0.12	na	0.06	0.06
Vocabulary knowledge	0.36	0.18	0.12	na	0.09
Word recognition	0.37	0.16	0.12	0.09	na
k=1 average	na	0.15	0.10	0.07	0.07
Reading comprehension and syntactic knowledge	0.54	na	na	0.02	0.02
Reading comprehension and vocabulary knowledge	0.54	na	0.03	na	0.02
Reading comprehension and word recognition	0.53	na	0.03	0.03	na
Syntactic knowledge and vocabulary knowledge	0.48	0.09	na	na	0.03
Syntactic knowledge and word recognition	0.49	0.08	na	0.03	na
Vocabulary knowledge and word recognition	0.46	0.11	0.06	na	na
k=2 average	na	0.09	0.04	0.03	0.03
Reading comprehension, syntactic knowledge, and vocabulary knowledge	0.57	na	na	na	0.01
Reading comprehension, syntactic knowledge, and word recognition	0.56	na	na	0.02	na
Reading comprehension, vocabulary knowledge, and word recognition	0.56	na	0.02	na	na
Syntactic knowledge, vocabulary knowledge, and word recognition	0.51	0.07	na	na	na
k=3 average	na	0.07	0.02	0.02	0.01
Reading comprehension, syntactic knowledge, vocabulary knowledge, and word recognition	0.58	na	na	na	na
Overall average	na	0.20	0.14	0.12	0.12
Standardized regression coefficient	na	0.37	0.20	0.17	0.15

na is not applicable. PSAT/NMSQT is Preliminary SAT/National Merit Scholarship Qualifying Test. FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards.

Note: *k* is the number of variables in the subset model.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 1.

Table A4. Dominance analysis results for Preliminary SAT/National Merit Scholarship Qualifying Test math performance for Florida district 1 students, 2013/14 (*n* = 1,099)

Subset model	*R* squared	Additional contribution of FAIR FS score in:			
		Reading comprehension	Syntactic knowledge	Vocabulary knowledge	Word recognition
Null and *k*=0 average	na	0.31	0.31	0.21	0.22
Reading comprehension	0.31	na	0.06	0.03	0.03
Syntactic knowledge	0.31	0.06	na	0.02	0.02
Vocabulary knowledge	0.21	0.13	0.12	na	0.06
Word recognition	0.22	0.12	0.12	0.05	na
k=1 average	na	0.10	0.10	0.03	0.03
Reading comprehension and syntactic knowledge	0.37	na	na	0.00	0.01
Reading comprehension and vocabulary knowledge	0.34	na	0.04	na	0.01
Reading comprehension and word recognition	0.33	na	0.04	0.01	na
Syntactic knowledge and vocabulary knowledge	0.33	0.05	na	na	0.01
Syntactic knowledge and word recognition	0.33	0.04	na	0.01	na
Vocabulary knowledge and word recognition	0.27	0.08	0.07	na	na
k=2 average	na	0.06	0.05	0.01	0.01
Reading comprehension, syntactic knowledge, and vocabulary knowledge	0.37	na	na	na	0.00
Reading comprehension, syntactic knowledge, and word recognition	0.38	na	na	0.00	na
Reading comprehension, vocabulary knowledge, and word recognition	0.35	na	0.03	na	na
Syntactic knowledge, vocabulary knowledge, and word recognition	0.34	0.04	na	na	na
k=3 average	na	0.04	0.03	0.00	0.00
Reading comprehension, syntactic knowledge, vocabulary knowledge, and word recognition	0.38	na	na	na	na
Overall average	na	0.13	0.12	0.06	0.07
Standardized regression coefficient	na	0.28	0.27	0.07	0.08

na is not applicable. PSAT/NMSQT is Preliminary SAT/National Merit Scholarship Qualifying Test. FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards.

Note: *k* is the number of variables in the subset model.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 1.

with 12 percent explained by FAIR-FS syntactic knowledge scores. FAIR-FS reading comprehension and syntactic knowledge scores each contribute on average 16 percent of the unique variance to explaining PSAT/NMSQT writing performance (table A5).

Table A5. Dominance analysis results for Preliminary SAT/National Merit Scholarship Qualifying Test writing performance for Florida district 1 students (*n* = 1,099), 2013/14

Subset model	*R* squared	Additional contribution of FAIR FS score in:			
		Reading comprehension	Syntactic knowledge	Vocabulary knowledge	Word recognition
Null and *k*=0 average	na	0.41	0.43	0.31	0.36
Reading comprehension	0.41	na	0.09	0.05	0.07
Syntactic knowledge	0.43	0.07	na	0.03	0.06
Vocabulary knowledge	0.31	0.16	0.15	na	0.11
Word recognition	0.36	0.12	0.13	0.06	na
k=1 average	na	0.12	0.12	0.05	0.08
Reading comprehension and syntactic knowledge	0.51	na	na	0.01	0.03
Reading comprehension and vocabulary knowledge	0.47	na	0.05	na	0.04
Reading comprehension and word recognition	0.48	na	0.05	0.02	na
Syntactic knowledge and vocabulary knowledge	0.46	0.06	na	na	0.04
Syntactic knowledge and word recognition	0.49	0.04	na	0.01	na
Vocabulary knowledge and word recognition	0.42	0.08	0.08	na	na
k=2 average	na	0.06	0.06	0.01	0.03
Reading comprehension, syntactic knowledge, and vocabulary knowledge	0.52	na	na	na	0.02
Reading comprehension, syntactic knowledge, and word recognition	0.53	na	na	0.00	na
Reading comprehension, vocabulary knowledge, and word recognition	0.50	na	0.03	na	na
Syntactic knowledge, vocabulary knowledge, and word recognition	0.50	0.04	na	na	na
k=3 average	na	0.04	0.03	0.00	0.02
Reading comprehension, syntactic knowledge, vocabulary knowledge, and word recognition	0.54	na	na	na	na
Overall average	na	0.16	0.16	0.09	0.12
Standardized regression coefficient	na	0.28	0.28	0.10	0.20

na is not applicable. PSAT/NMSQT is Preliminary SAT/National Merit Scholarship Qualifying Test. FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards.

Note: *k* is the number of variables in the subset model.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 1.

ACT Plan dominance analyses. Examination of the variance components for ACT Plan English performance shows that there are no completely dominant FAIR-FS predictors (table A6). FAIR-FS syntactic knowledge scores contribute, on average, 16 percent of the unique variance in explaining ACT Plan reading performance (table A7). Examination of the unique variance components for ACT Plan science performance shows that there are no completely dominant FAIR-FS predictors (table A8). FAIR-FS syntactic knowledge scores contribute, on average, 13 percent of the unique variance in explaining ACT Plan math performance (table A9).

Table A6. Dominance analysis results for ACT Plan English performance for Florida district 2 students, 2013/14 (*n* = 296)

Subset model	*R* squared	Additional contribution of FAIR FS score in:			
		Reading comprehension	Syntactic knowledge	Vocabulary knowledge	Word recognition
Null and *k*=0 average	na	0.42	0.45	0.43	0.39
Reading comprehension	0.42	na	0.07	0.08	0.05
Syntactic knowledge	0.45	0.03	na	0.08	0.06
Vocabulary knowledge	0.43	0.06	0.10	na	0.06
Word recognition	0.39	0.07	0.12	0.10	na
k=1 average	na	0.06	0.10	0.08	0.05
Reading comprehension and syntactic knowledge	0.48	na	na	0.05	0.03
Reading comprehension and vocabulary knowledge	0.49	na	0.04	na	0.02
Reading comprehension and word recognition	0.46	na	0.05	0.05	na
Syntactic knowledge and vocabulary knowledge	0.53	0.00	na	na	0.02
Syntactic knowledge and word recognition	0.51	0.01	na	0.04	na
Vocabulary knowledge and word recognition	0.49	0.02	0.06	na	na
k=2 average	na	0.01	0.05	0.04	0.02
Reading comprehension, syntactic knowledge, and vocabulary knowledge	0.53	na	na	na	0.01
Reading comprehension, syntactic knowledge, and word recognition	0.52	na	na	0.03	na
Reading comprehension, vocabulary knowledge, and word recognition	0.51	na	0.04	na	na
Syntactic knowledge, vocabulary knowledge, and word recognition	0.55	0.00	na	na	na
k=3 average	na	0.00	0.04	0.03	0.01
Reading comprehension, syntactic knowledge, vocabulary knowledge, and word recognition	0.55	na	na	na	na
Overall average	na	0.12	0.16	0.15	0.12
Standardized regression coefficient	na	0.04	0.33	0.28	0.18

na is not applicable. PSAT/NMSQT is Preliminary SAT/National Merit Scholarship Qualifying Test. FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards.

Note: *k* is the number of variables in the subset model.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 2.

Table A7. Dominance analysis results for ACT Plan reading performance for Florida district 2 students, 2013/14 (*n* = 296)

Subset model	*R* squared	Additional contribution of FAIR FS score in:			
		Reading comprehension	Syntactic knowledge	Vocabulary knowledge	Word recognition
Null and *k*=0 average	na	0.37	0.42	0.33	0.35
Reading comprehension	0.37	na	0.07	0.04	0.05
Syntactic knowledge	0.42	0.02	na	0.03	0.05
Vocabulary knowledge	0.33	0.08	0.13	na	0.08
Word recognition	0.35	0.06	0.12	0.05	na
k=1 average	na	0.06	0.11	0.04	0.06
Reading comprehension and syntactic knowledge	0.44	na	na	0.02	0.03
Reading comprehension and vocabulary knowledge	0.41	na	0.06	na	0.03
Reading comprehension and word recognition	0.42	na	0.05	0.02	na
Syntactic knowledge and vocabulary knowledge	0.46	0.01	na	na	0.02
Syntactic knowledge and word recognition	0.47	0.00	na	0.01	na
Vocabulary knowledge and word recognition	0.40	0.03	0.08	na	na
k=2 average	na	0.01	0.06	0.01	0.02
Reading comprehension, syntactic knowledge, and vocabulary knowledge	0.46	na	na	na	0.02
Reading comprehension, syntactic knowledge, and word recognition	0.47	na	na	0.01	na
Reading comprehension, vocabulary knowledge, and word recognition	0.43	na	0.05	na	na
Syntactic knowledge, vocabulary knowledge, and word recognition	0.48	0.00	na	na	na
k=3 average	na	0.00	0.05	0.01	0.02
Reading comprehension, syntactic knowledge, vocabulary knowledge, and word recognition	0.48	na	na	na	na
Overall average	na	0.11	0.16	0.10	0.11
Standardized regression coefficient	na	0.06	0.37	0.14	0.21

na is not applicable. PSAT/NMSQT is Preliminary SAT/National Merit Scholarship Qualifying Test. FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards.

Note: *k* is the number of variables in the subset model.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 2.

Table A8. Dominance analysis results for ACT Plan science performance for Florida district 2 students, 2013/14 (*n* = 296)

Subset model	*R* squared	Additional contribution of FAIR FS score in:			
		Reading comprehension	Syntactic knowledge	Vocabulary knowledge	Word recognition
Null and *k*=0 average	na	0.35	0.35	0.33	0.29
Reading comprehension	0.35	na	0.04	0.05	0.02
Syntactic knowledge	0.35	0.04	na	0.06	0.04
Vocabulary knowledge	0.33	0.06	0.08	na	0.04
Word recognition	0.29	0.08	0.10	0.08	na
k=1 average	na	0.06	0.07	0.06	0.03
Reading comprehension and syntactic knowledge	0.39	na	na	0.03	0.01
Reading comprehension and vocabulary knowledge	0.40	na	0.03	na	0.01
Reading comprehension and word recognition	0.37	na	0.03	0.03	na
Syntactic knowledge and vocabulary knowledge	0.41	0.01	na	na	0.01
Syntactic knowledge and word recognition	0.39	0.01	na	0.03	na
Vocabulary knowledge and word recognition	0.37	0.03	0.05	na	na
k=2 average	na	0.02	0.04	0.03	0.01
Reading comprehension, syntactic knowledge, and vocabulary knowledge	0.42	na	na	na	0.00
Reading comprehension, syntactic knowledge, and word recognition	0.41	na	na	0.02	na
Reading comprehension, vocabulary knowledge, and word recognition	0.40	na	0.02	na	na
Syntactic knowledge, vocabulary knowledge, and word recognition	0.42	0.01	na	na	na
k=3 average	na	0.01	0.02	0.02	0.00
Reading comprehension, syntactic knowledge, vocabulary knowledge, and word recognition	0.43	na	na	na	na
Overall average	na	0.11	0.12	0.11	0.08
Standardized regression coefficient	na	0.14	0.26	0.23	0.10

na is not applicable. PSAT/NMSQT is Preliminary SAT/National Merit Scholarship Qualifying Test. FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards.

Note: *k* is the number of variables in the subset model.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 2.

Table A9. Dominance analysis results for ACT Plan math performance for Florida district 2 students, 2013/14 (*n* = 296)

Subset model	*R* squared	Additional contribution of FAIR FS score in:			
		Reading comprehension	Syntactic knowledge	Vocabulary knowledge	Word recognition
Null and *k*=0 average	na	0.28	0.33	0.28	0.21
Reading comprehension	0.28	na	0.06	0.04	0.01
Syntactic knowledge	0.33	0.01	na	0.04	0.01
Vocabulary knowledge	0.28	0.05	0.09	na	0.02
Word recognition	0.21	0.08	0.13	0.09	na
k=1 average	na	0.05	0.09	0.05	0.01
Reading comprehension and syntactic knowledge	0.35	na	na	0.02	0.00
Reading comprehension and vocabulary knowledge	0.33	na	0.04	na	0.00
Reading comprehension and word recognition	0.29	na	0.06	0.03	na
Syntactic knowledge and vocabulary knowledge	0.37	0.00	na	na	0.00
Syntactic knowledge and word recognition	0.34	0.01	na	0.03	na
Vocabulary knowledge and word recognition	0.29	0.03	0.07	na	na
k=2 average	na	0.01	0.06	0.03	0.00
Reading comprehension, syntactic knowledge, and vocabulary knowledge	0.37	na	na	na	0.00
Reading comprehension, syntactic knowledge, and word recognition	0.35	na	na	0.02	na
Reading comprehension, vocabulary knowledge, and word recognition	0.33	na	0.04	na	na
Syntactic knowledge, vocabulary knowledge, and word recognition	0.37	0.00	na	na	na
k=3 average	na	0.00	0.04	0.02	0.00
Reading comprehension, syntactic knowledge, vocabulary knowledge, and word recognition	0.37	na	na	na	na
Overall average	na	0.09	0.13	0.10	0.06
Standardized regression coefficient	na	0.08	0.36	0.23	0.00

na is not applicable. PSAT/NMSQT is Preliminary SAT/National Merit Scholarship Qualifying Test. FAIR-FS is Florida Assessments for Instruction in Reading–Florida Standards.

Note: *k* is the number of variables in the subset model.

Source: Authors' analysis of data from the Florida Center for Reading Research and district 2.

References

ACT. (2013). *PLAN technical manual*. Iowa City, IA: Author. Retrieved January 14, 2014, from http://www.act.org/plan/pdf/PlanTechnicalManual.pdf.

ACT. (2014). *The condition of college and career readiness 2014: Florida*. Iowa City, IA: Author. Retrieved March 9, 2015, from http://www.act.org/newsroom/data/2014/states/pdf/Florida.pdf.

ACT. (2015). *ACT Plan overview*. Retrieved December 1, 2015, from http://www.act.org/products/k-12-act-plan/.

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods*, 8(2), 129–148.

Berk, R. A. (2008). *Statistical learning from a regression perspective*. New York, NY: Springer.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth International Group.

College Board. (2011). *SAT benchmarks: Development of a college readiness benchmark and its relationship to secondary and postsecondary school performance* (Research Report 2011–5). New York, NY: Author. Retrieved July 9, 2015, from http://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-2011–5-sat-college-readiness-benchmark-secondary-performance.pdf.

College Board. (2013). *Understanding 2013 PSAT/NMSQT scores*. New York, NY: Author. Retrieved January 14, 2014, from http://professionals.collegeboard.com/profdownload/understanding-psat-nmsqt-scores.pdf.

College Board. (2014). *2013–2014 college-bound high school sophomores: Summary report [Florida]*. New York, NY: Author. Retrieved March 9, 2015, from http://media.collegeboard.com/digitalServices/pdf/research/2014/FL_14_05_03_01.pdf.

College Board. (2015). *Redesigned SAT*. New York, NY: Author. Retrieved April 9, 2015, from https://www.collegeboard.org/delivering-opportunity/sat/redesign.

Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology*, 98(2), 394–409. http://eric.ed.gov/?id=EJ742190

Council of Chief State School Officers. (2008). *Interim assessment practices and avenues for state involvement* (TILSA SCASS Report). Washington, DC: Author. Retrieved July 9, 2015, from http://www.ccsso.org/Documents/2008/Interim_Assessment_Practices_and_2008.pdf.

Cunningham, A. E., & Stanovich, K. E. (1998). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(1), 934–945. http://eric.ed.gov/?id=EJ561721

Dorans, N. J. (2000). *Distinctions among classes of linkages* (Research Note 11). New York, NY: The College Board. Retrieved July 9, 2015, from http://research.collegeboard.org/sites/default/files/publications/2012/7/researchnote-2000–11-distinctions-classes-linkages.pdf.

Foorman, B. R., Petscher, Y., & Schatschneider, C. (2015). *Florida assessments for instruction in reading, aligned to the language arts Florida standards (FAIR-FS): Grades 3 through 12: Technical manual*. Tallahassee, FL: Florida Center for Reading Research. Retrieved March 30, 2015, from http://www.fcrr.org/_/images/FAIR_tech_manuals/FAIR-FS_3-12_Tech_Manual_v1_2.pdf.

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for service delivery in an RTI framework: Candidate measures. *School Psychology Review*, *36*(4), 582–601.

Koon, S., Petscher, Y., & Foorman, B. R. (2014). *Using evidence-based decision trees instead of formulas to identify at-risk readers* (REL 2014–036). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. http://eric.ed.gov/?id=ED545225

Lewis, R. J. (2000). An introduction to classification and regression tree (CART) analysis. In *Annual Meeting of the Society for Academic Emergency Medicine in San Francisco, California* (pp. 1–14). Des Plaines, IL: Society for Academic Emergency Medicine. Retrieved July 9, 2015, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.95.4103&rep=rep1&type=pdf.

Piasta, S. B., Petscher, Y., & Justice, L. M. (2012). How many letters should preschoolers in public programs know? The diagnostic efficiency of various preschool letter-naming benchmarks for predicting first-grade literacy achievement. *Journal of Educational Psychology*, *104*(4), 945–958.

Porter, A. C., & Polikoff, M. S. (2012). Measuring academic readiness for college. *Educational Policy*, *26*(3), 394–417. http://eric.ed.gov/?id=EJ970620

Schatschneider, C., Petscher, Y., & Williams, K. M. (2008). How to evaluate a screening process: The vocabulary of screening and what educators need to know. In L. Justice & C. Vukelic (Eds.), *Every moment counts: Achieving excellence in preschool language and literacy instruction* (pp. 304–317). New York, NY: Guilford Press.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*(6), 934–947. http://eric.ed.gov/?id=EJ663347

Therneau, T. M., Atkinson, B., & Ripley, B. (2014). '*rpart': Recursive Partitioning.* R package version 4.1–9. Retrieved March 13, 2015, from http://cran.r-project.org/web/packages/rpart/rpart.pdf.

Therneau, T. M., & Atkinson, E.J. (2015). *An introduction to recursive partitioning using the 'rpart' routines*. Retrieved March 13, 2015, from https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf.

Thurber, R. S., Shinn, M. R., & Smolkowski, K. (2002). What is measured in mathematics tests? Construct validity of curriculum-based mathematics measures. *School Psychology Review, 31*(4), 498–513. http://eric.ed.gov/?q=EJ667623&id=EJ667623

Whitehurst, G. J., & Lonigan, C. J. (2001). Emergent literacy: Development from prereaders to readers. In S. B Neuman & D. K. Dickensen (Eds.), *Handbook of early literacy research* (pp. 11–29). New York, NY: Guilford Press.

The Regional Educational Laboratory Program produces 7 types of reports

Making Connections
Studies of correlational relationships

Making an Impact
Studies of cause and effect

What's Happening
Descriptions of policies, programs, implementation status, or data trends

What's Known
Summaries of previous research

Stated Briefly
Summaries of research findings for specific audiences

Applied Research Methods
Research methods for educational settings

Tools
Help for planning, gathering, analyzing, or reporting data or research