

APPENDIX E:
NEW DEVELOPED CUT SCORES

Establishing MME and MEAP Cut Scores Consistent with College and Career Readiness

A study conducted by the Michigan Department of
Education (MDE) and ACT, Inc.

December, 2011

Executive Summary

The purpose of this study was to identify three new sets of cut scores on the Michigan Educational Assessment Program (MEAP) and the Michigan Merit Examination (MME), classifying students into four performance levels: Advanced, Proficient, Partially Proficient, and Not Proficient, where Proficient is defined as being on track to succeed in a postsecondary educational experience. Three types of linkages were needed to identify the cut scores. The first was to link 11th grade MME scores to freshman college grades to identify cut scores on the MME. The second was to link MEAP scores to MME scores to identify cut scores on one or more grades of the MEAP. The third was to link MEAP scores in other grades to the MEAP scores that linked to the MME scores to identify cut scores on the remaining grades of the MEAP. The new cut scores appear below for mathematics, reading, science, and social studies.

New MEAP and MME Mathematics Cut Scores

Assessment	Grade	Partially Proficient	Proficient	Advanced
MME	11	1093	1116	1138
MEAP	8	809	830	865
MEAP	7	714	731	776
MEAP	6	614	629	675
MEAP	5	516	531	584
MEAP	4	423	434	470
MEAP	3	322	336	371

New MEAP and MME Reading Cut Scores

Assessment	Grade	Partially Proficient	Proficient	Advanced
MME	11	1081	1108	1141
MEAP	8	796	818	853
MEAP	7	698	721	760
MEAP	6	602	619	653
MEAP	5	501	521	565
MEAP	4	395	419	478
MEAP	3	301	324	364

New MEAP and MME Science Cut Scores

Assessment	Grade	Partially Proficient	Proficient	Advanced
MME	11	1106	1126	1144
MEAP	8	826	845	863
MEAP	5	526	553	567

New MEAP and MME Social Studies Cut Scores

Assessment	Grade	Partially Proficient	Proficient	Advanced
MME	11	1097	1129	1158
MEAP	9	899	928	960
MEAP	6	593	625	649

To provide some information regarding the impact of implementing the new cut scores, the new cut scores were applied to the MME and MEAP data from the 2010-2011 school year, comparing the percentage of students in each category (Advanced, Proficient, Partially Proficient, Not Proficient) to the percentages obtained using the previous cut scores.

Establishing MME and MEAP Cut Scores Consistent with College and Career Readiness

Introduction

The purpose of this study was to identify three new sets of cut scores on the Michigan Educational Assessment Program (MEAP) and the Michigan Merit Examination (MME), classifying students into four performance levels: Advanced, Proficient, Partially Proficient, and Not Proficient, where Proficient is defined as being on track to succeed in a postsecondary educational experience. Three types of linkages were needed to identify the cut scores. The first was to link 11th grade MME scores to freshman college grades to identify cut scores on the MME. The second was to link MEAP scores to MME scores to identify cut scores on one or more grades of the MEAP. The third was to link MEAP scores in other grades to the MEAP scores that linked to the MME scores to identify cut scores on the remaining grades of the MEAP. The work was conducted by the Michigan Department of Education (MDE) and ACT, Inc.

The impact of the new cut scores was studied by applying the new cut scores to existing MEAP and MME 2010-2011 data, and comparing these results to the operational results based on the previous cut scores. The new cut scores were higher than the previous cut scores, resulting in lower percentages of students attaining “Proficient” or “Advanced” classification.

New cut scores were set in mathematics, reading, science, and social studies. Writing was not included in the study because (1) the MEAP writing test was new in Fall 2011 and thus does not have the data necessary to map cut scores on the MEAP back from cut scores on the MME, (2) the MME writing cut score is already similar to the ACT writing college readiness benchmark, and (3) the MEAP writing cut scores were already set to be consistent with the MME writing cut scores.

Establishing MME Cut Scores Consistent with College and Career Readiness.

The first step in setting new cut scores was to set college readiness cutoff scores on the 11th grade MME. This was accomplished by relating course grades from first-year college students enrolled in Michigan public postsecondary institutions (two- and four-year) to MME scores.

Determining which College Courses to Include

All Michigan postsecondary institutions were asked to provide a list of first-year credit-bearing courses that they felt would be appropriate. Ten institutions responded and their course selections were included. For institutions that did not provide lists, courses were selected by ACT from college course catalogs. The final list was reviewed and approved by Michigan Department of Education (MDE) staff with some changes (additions and deletions). Each course was assigned to a subject area (mathematics, reading, science, or social studies). Some courses were used for both reading and social studies; a list of the type of courses included in the study appears in Table 1. Using the final list, grades for courses were pulled by ACT from the Center for Educational Performance and Information (CEPI) grade file provided by MDE. The final file included 13 four-year and 26 two-year public institutions (see Table 2).

Table 1.

College Courses Used for the Analysis of each MME Content Area

MME content area	College courses types used
Mathematics	College Algebra.
Reading	Courses identified by 4-year universities. Reading-heavy courses such as entry-level literature, history, philosophy, or psychology for 2-year universities.
Science	Courses identified by 4-year universities. Entry level biology, chemistry, physics, or geology for 2-year universities.
Social Studies	Courses identified by 4-year universities. Entry level history, geography, or economics for 2-year universities.

Table 2.

Institution List

Institution names	
Alpena Community College	Oakland University
Kellogg Community College	Muskegon Community College
Delta College	North Central Michigan College
Central Michigan University	Northern Michigan University
Mid Michigan Community College	Monroe County Community College
Lake Michigan College	Northwestern Michigan College
Eastern Michigan University	Oakland Community College
Ferris State University	St. Clair County Community College
Mott Community College	Macomb Community College
Glen Oaks Community College	Southwestern Michigan College
Gogebic Community College	Saginaw Valley State University
Grand Rapids Community College	The University of Michigan - Ann Arbor
Grand Valley State University	University of Michigan – Flint
Henry Ford Community College	Wayne State University
Jackson Community College	Western Michigan University
Kalamazoo Valley Community College	Washtenaw Community College
Kirtland Community College	Schoolcraft College
Lansing Community College	University of Michigan Dearborn
Montcalm Community College	West Shore Community College
Michigan State University	

Obtaining College Grade Data

College course grade data were obtained from the CEPI, who provided a data file of grades for over 16 million records for students who enrolled between 1955 and 2011. Grade definitions in

the CEPI file varied by institution. These were recoded by ACT using the Postsecondary Electronic Standards Council (PESC) college transcript implementation guide provided by CEPI. Grades were put on a numeric scale of 0-4. Grades of AU, AWF, DR, R, FA, FR, T, TR, and X were not used in any of the analyses. When a grade of C was defined as a successful outcome, grades of P and CR were set to 2.0 and grades of IN, N, NC, NE, NS, W, WF, WP, WX, and U were set to 0.0. None of these grades were used when success was defined as a grade of B or higher or an A or higher. A list of courses by subject area can be found in Table 3. Because the hierarchical nature of the math subject area tends to create differences in course type difficulty, this subject area was limited to one course type - College Algebra. This is typically the first credit-bearing math course taken by a large number of college students.

Data Matching and Cleaning

Students with first college enrollment dates of Fall 2009 and Fall 2010 were used in the study. These were the first cohorts that had both 11th grade MME scores and college course grades. After matching and cleaning, the final sample size was 104,691 records. Because the data file was structured with one course type per record, there were multiple records for most students. Table 3 provides the number of student grades used for each individual course. Table 4 lists the descriptive statistics for the four MME subject areas being studied.

Table 3.

Student Grades Used by Course

MME test	Course type	Number of students
Math	College Algebra	6,567
Reading	Literature	456
	American History	1,731
	Other History	3,010
	Psychology	16,231
	Sociology	8,236
	Political Science	6,114
	Philosophy	1,869
	Other	2,517
Science	Biology/Life Sciences	8,355
	General Chemistry	5,807
	Physics	535
	Other	1,483
Social Studies	American History	1,734
	Other History	3,006
	Psychology	16,230
	Sociology	8,231
	Geography	612
	Political Science	6,108
	Economics	3,498
	Other	2,361

Table 4.

Descriptive Statistics for Final Data Set

MME subject area	Sample size	Percent B or higher	Course grade		MME score	
			Mean	SD	Mean	SD
Math	6,286	47.0	2.49	1.18	1112.18	13.17
Reading	37,952	54.9	2.64	1.23	1117.16	24.60
Science	15,608	50.0	2.54	1.19	1123.45	23.54
Social Studies	39,721	54.4	2.63	1.22	1135.37	26.33

Methodologies Considered in Setting Cut Scores

Three different methodologies were used in identifying the MME cut scores: Logistic Regression (LR), Signal Detection Theory (SDT), and Conditional Means (CM).

The LR model used in this study takes the form

$$P(\text{success}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

where

success is defined as at or above the grade corresponding to college readiness (e.g., C or higher), as proficiency on the MME, or as proficiency on the MEAP;

$P(\text{success})$ is the probability of success;

e is the base of the natural logarithm;

β_0 is the intercept of the logistic regression;

β_1 is the slope of the logistic regression; and

x is the MME or MEAP score being used to predict success.

The cut score identified with the LR model is the score on the MME that gives a 50% probability of success. This score gives the highest accuracy rate by maximizing the percentage of students who:

- were successful in the course (e.g., received a B or better) AND were considered proficient on the MME (at or above the cutoff score), or
- were not successful in the course (e.g., received a score of B- or lower) AND were considered not proficient on the MME (below the cutoff score).

The SDT model used in this study also maximizes the rates of consistent classification. The observed percentage of students who were classified correctly was calculated using the following student subsets for the entire MME score range:

$A_{00}(X)$ = the number of students who score below X on the MME, and get a grade of below B in the college class (were not successful).

$A_{01}(X)$ = the number of students who score below X on the MME, and get a grade of B or better in the college class (were successful).

$A_{10}(X)$ = the number of students who score at or above X on the MME, and get a grade of below B in the college class (were not successful).

$A_{11}(X)$ = the number of students who score at or above X on the MME, and get a grade of B or better in the college class (were successful).

The cut score is the score X that maximizes $A_{00}(X) + A_{11}(X)$.

Note that under mild monotonicity assumptions, this method is equivalent to choosing the score point such that the conditional probability of exceeding the cut score equals .5. To the extent that the assumption holds, LR and SDT should give similar solutions. Finally, the SDT observed classification rates were smoothed to avoid any effects of small sample sizes at some score points on the results.

The CM method calculates the mean course grade for each score on the MME scale. The cutoff score is the score at which the average course grade was closest to the success criterion (e.g., 3.0 = B or higher).

Three success criteria were used as measures of college readiness in the analyses: a grade of A or higher, B or higher, and C or higher. Separate analyses were performed for two-year and four-year institutions.

Results

Preliminary results revealed that the A or higher and C or higher criteria were not feasible as measures of college readiness. For many institutions, the A or higher cutoff scores were too high to back map to MEAP scores and the C or higher cutoffs were below chances values. In addition, the results from CM were not considered stable because of small sample sizes at some score points.

Initially, separate analyses were run for 2- and 4-year institutions. Although the results provided different cut scores, they were within measurement error of each other. Therefore, all institutions were combined to establish the Proficient cut scores. The cut score between Not Proficient and Partially Proficient and the cut score between Proficient and Advanced were identified using the 33% and 67% chance of success levels.

The results from LR and SDT were very similar. SDT was the methodology preferred by MDE and its results were used for the MME cut scores. The final cut scores were based on a success criterion of B or higher, and two-year and four-year institutions were combined.

For each subject area, SDT was used to generate a distribution of consistency classifications across MME test score by institution. The median consistency at each score was calculated

across institutions and a logistic regression function was fit to this distribution to smooth the results. The MME scores with the highest median consistency were selected as the 11th grade MME college readiness cutoff scores and are shown in Table 5.

Table 5.

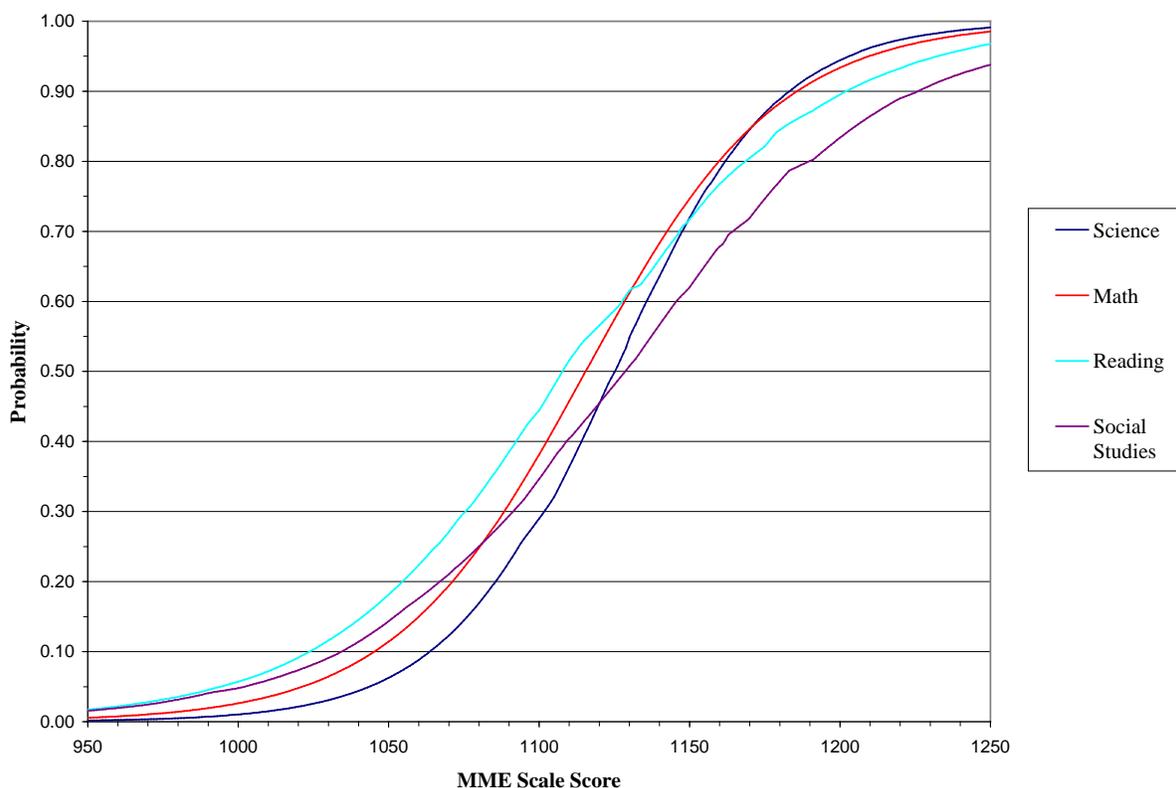
*MME College Readiness Cutoff Scores**

Subject area	Classification consistency	Partially proficient	Proficient	Advanced
Math	65%	1093	1116	1138
Reading	63%	1081	1108	1141
Science	67%	1106	1126	1144
Social Studies	63%	1097	1129	1158

* Grades of P, CR, S, IN, N, NC, NE, NS, W, WF, WP, WX, and U were not included.

A score that gives the highest classification consistency also has a probability of success of 0.50 – meaning that students with this score have a 50% chance of receiving a B or higher course grade in the subject area. Partially Proficient and Advanced cutoff scores were selected as the scores at which students had a 33% and 67% chance of success. Figure 1 shows the smoothed probability curves for the four MME subject areas.

Figure 1. *Probability of a B or Higher in College Courses by MME Subject Area Test*



Classification consistencies for Partially Proficient and Advanced cutoff scores could not be calculated and would not be meaningful because they were based on the Proficient cutoff scores probability curves. In other words, the definition of Partially Proficient success and Advanced success were not based on specific course outcomes.

Establishing MEAP Cut Scores Consistent with College and Career Readiness

Once the MME 11th grade cut scores were established, the MEAP to MME linkage was created for mathematics, reading, science, and social studies. As stated above, the MEAP writing cut scores were already set to be consistent with the MME writing cut scores.

In practice, setting the MEAP cut scores informed the final determination of MME cut scores, by providing reasonableness on the results. For example, a combination of criteria, data, and methodology that provided reasonable looking cut scores for MME in grade 11 might back map to an extremely small percentage of students attaining a proficient rating in a lower grade, such as less than 1%, which lacks credibility.

Data Cleaning and Matching

Six years of MEAP data (Fall 2005 to Fall 2010) and four years of MME data (Spring 2008 to Spring 2011) were used in this project. Each MEAP year of data contained students from grade 3 through grade 9 and each MME year of data contained students from either grade 11 or 12.

The matching of students across grades was done to provide an empirical basis for determining the cut scores. Prior to matching, each year of MEAP and MME data were cleaned based on the following rules.

Cleaning rules for MEAP data:

- remove records with invalid UICs: length of the UIC code is not equal to 10 and UIC is not equal to “0000000000”;
- remove students from private schools;
- remove records with invalid scores on all subjects;
- remove records with test type = “FI”;
- remove records with duplicate UICs within and across grades in the same dataset.

Cleaning rules for MME data:

- remove records with invalid UICs: length of the UIC code is not equal to 10 and UIC is not equal to “0000000000”;
- remove records with invalid scores on all subjects;
- remove records with duplicate UICs within and across grades in the same dataset.

Based on all available data after cleaning, 11 cohorts were identified, each spanning at least two grades. Table 6 presents all the cohorts, the grades they cover, and the school year (e.g., 05-06) for each included grade. Each cohort goes back to a minimum of grade 3 (since grade 3 is the lowest grade in which students were tested on MEAP) and to the 2005-06 (05-06) school year (since the MEAP test was new in the 2005-2006 school year).

Students from the same cohort were matched by UIC across different years of data. For example, Cohort 2 data was created by first matching grade 3 students in the 2008 Fall MEAP data to grade 4 students in the 2009 Fall MEAP data, then to grade 5 students in the 2010 Fall MEAP data, by UIC. For MME data, only grade 11 students were selected and matched. Table 6 shows

two sample sizes for the same cohort. N_Total represents the number of cases providing at least one year of data and N_Full_Matched presents the number of cases providing all years of data.

Table 6.

Cohort with at Least Two Years of Data

Cohort	Grade										Sample Size	
	3	4	5	6	7	8	9	10 ^a	11	12	N_Total ^b	N_Full_Matched ^c
1	09-10	10-11	-	-	-	-	-	-	-	-	118,468	107,844
2	08-09	09-10	10-11	-	-	-	-	-	-	-	123,896	105,019
3	07-08	08-09	09-10	10-11	-	-	-	-	-	-	128,557	101,277
4	06-07	07-08	08-09	09-10	10-11	-	-	-	-	-	135,268	98,131
5	05-06	06-07	07-08	08-09	09-10	10-11	-	-	-	-	140,729	94,044
6	-	05-06	06-07	07-08	08-09	09-10	10-11	-	-	-	149,950	93,447
7	-	-	05-06	06-07	07-08	08-09	09-10	10-11	-	-	151,047	98,155
8	-	-	-	05-06	06-07	07-08	08-09	09-10	10-11	-	161,869	88,526
9	-	-	-	-	05-06	06-07	07-08	08-09	09-10	10-11	163,246	93,004
10	-	-	-	-	-	05-06	06-07	07-08	08-09	09-10	160,813	97,172
11	-	-	-	-	-	-	05-06	06-07	07-08	08-09	158,255	104,352

^a No testing in this grade.

^b Students having at least one year of data.

^c Students having all years of data.

Methodologies Considered

Three methodologies were examined in determining the MEAP cut scores: LR, SDT, and Equipercentile Cohort Matching (ECM).

After identifying the cut score for proficiency on the MME, the cut scores were then mapped backward onto the MEAP to achieve the same type of results (meaning that the known outcome was then proficiency on the MME and the unknown outcome was proficiency on the MEAP).

Because both LR and SDT are subject to regression effects, it was important to address these effects by having the minimum number of links in defining each grade level's cut score. When linking each grade to the grade just previous to it, there would be seven links for the third grade cut score as shown here:

1. Linking grade 11 MME to college grades.
2. Linking grade 8 MEAP to grade 11 MME.
3. Linking grade 7 MEAP to grade 8 MEAP.
4. Linking grade 6 MEAP to grade 7 MEAP.
5. Linking grade 5 MEAP to grade 6 MEAP.
6. Linking grade 4 MEAP to grade 5 MEAP.
7. Linking grade 3 MEAP to grade 4 MEAP.

Instead, a different linking scheme was implemented which limited the maximum number of links created to identify any grade level's cut score to three. Table 7 shows the links for each grade and content area to demonstrate that the maximum number of links was three.

Because both LR and SDT are subject to regression around the mean (meaning that they can inflate cut scores if they are above the mean, or deflate them if they are below the mean), the results of the LR and SDT models were carefully inspected to assure that any place in which there was evidence of regression effects, the ECM methodology was used.

Table 7.

Links in Tying Cut Scores on MME and MEAP to College Grades

Cut score		
Content area	Grade	Links created
Mathematics and Reading	3	#1. Grade 11 MME to College Grades #2. Grade 7 MEAP to Grade 11 MME #3. Grade 3 MEAP to Grade 7 MEAP
	4	#1. Grade 11 MME to College Grades #2. Grade 7 MEAP to Grade 11 MME #3. Grade 4 MEAP to Grade 7 MEAP
	5	#1. Grade 11 MME to College Grades #2. Grade 7 MEAP to Grade 11 MME #3. Grade 5 MEAP to Grade 7 MEAP
	6	#1. Grade 11 MME to College Grades #2. Grade 7 MEAP to Grade 11 MME #3. Grade 6 MEAP to Grade 7 MEAP
	7	#1. Grade 11 MME to College Grades #2. Grade 7 MEAP to Grade 11 MME
	8	#1. Grade 11 MME to College Grades #2. Grade 8 MEAP to Grade 11 MME
Science and Social Studies	11	#1. Grade 11 MME to College Grades
	5/6	#1. Grade 11 MME to College Grades #2. Grade 8/9 MEAP to Grade 11 MME #3. Grade 5/6 MEAP to Grade 8/9 MEAP
	8/9	#1. Grade 11 MME to College Grades #2. Grade 8/9 MEAP to Grade 11 MME
	11	#1. Grade 11 MME to College Grades

ECM was also used for the backmapping from MME onto MEAP to check for regression effects. Because ECM is a symmetric methodology, it cannot display any regression effects, and can therefore serve as a check for regression effects in the other two methods. The way ECM was used to backmap cut scores onto MEAP was to:

- Take the cohorts that took both the MME and the highest grade level of the MEAP.
- Identify the percentage of the matched cohorts that were proficient on the MME.
- Identify the score on the MEAP that as the cut score gives the most similar percentage passing the MEAP.

- Take the cohorts that took both the highest grade level of the MEAP and the next grade level down.
- Identify the percentage of the matched cohorts that were proficient on the highest level of the MEAP.
- Identify the score on the next grade level down that as the cut score gives the most similar percentage passing the MEAP.
- Repeat the process with the next grade level down until reaching the lowest grade level of MEAP.

The reasons that three methods were used were the following:

- LR and SDT served as a validation of each other.
- ECM served as a check on regression effects.

The three methodologies have different aims. LR aims to identify the score that gives a fixed probability of success. SDT aims to maximize consistent classifications from one level to the next. ECM aims to identify cut scores across grade levels that are approximately equally rigorous in terms of impact. Although they have different aims, they should give similar results. Therefore, it is important to determine which results to use in what circumstances.

SDT was considered the preferred methodology because its aim was to maximize consistent classification from one level to the next (an inherently desirable outcome in that if a student is classified as Proficient in one grade, they can be reasonably expected to be proficient in the next grade given typical education). Where SDT and LR were affected by regression effects, ECM was preferable in that it would produce non-inflated/deflated cut scores. Therefore, the results were inspected to determine whether SDT and/or LR were affected by regression effects. Where there was no evidence of regression effects, SDT results were used. Where there was evidence of regression effects, ECM results were used. The ECM method was used for the Advanced and Partially Proficient cut scores.

Efforts were made to include at least two cohorts for each backmapping, in order to reduce the sampling variations. Table 7 shows the linkage between grades for these backmappings. For example, for the math test, target scores for Grade 11 were backmapped to those of Grades 7 and 8, and the obtained target scores for Grade 7 served as the baseline for the backmapping to Grades 3-6.

It should be noted that the SDT methodology used for establishing the MEAP cut scores was implemented without smoothing. Note that in order to maximize the number of students who are accurately identified as “on track”, the number of correct classifications is calculated for each score point on the scale of the lower grade. It may be the case that for certain forms of the test, some score points are not possible, and for those score points, the proportion of students correctly classified cannot be calculated. In these cases, it was proposed that the proportions be smoothed and the maximum value taken to be the cut score. In this particular case that did not prove to be necessary, as for each grade, the maximum value for the classification consistency occurred in a range where all score points were achievable.

Results

The final cut scores for MME and MEAP are provided in Tables 8 through 11 for mathematics, reading, science, and social studies, respectively.

Table 8.

Recommended New MEAP and MME Mathematics Cut Scores

Assessment	Grade	Partially Proficient	Proficient	Advanced
MME	11	1093	1116	1138
MEAP	8	809	830	865
MEAP	7	714	731	776
MEAP	6	614	629	675
MEAP	5	516	531	584
MEAP	4	423	434	470
MEAP	3	322	336	371

Table 9.

Recommended New MEAP and MME Reading Cut Scores

Assessment	Grade	Partially Proficient	Proficient	Advanced
MME	11	1081	1108	1141
MEAP	8	796	818	853
MEAP	7	698	721	760
MEAP	6	602	619	653
MEAP	5	501	521	565
MEAP	4	395	419	478
MEAP	3	301	324	364

Table 10.

Recommended New MEAP and MME Science Cut Scores

Assessment	Grade	Partially Proficient	Proficient	Advanced
MME	11	1106	1126	1144
MEAP	8	826	845	863
MEAP	5	526	553	567

Table 11.

Recommended New MEAP and MME Social Studies Cut Scores

Assessment	Grade	Partially Proficient	Proficient	Advanced
MME	11	1097	1129	1158
MEAP	9	899	928	960
MEAP	6	593	625	649

The classification consistency rates are given in Table 12 for the links from MME to college grades, from MEAP to MME, and from one grade to another for MEAP.

Table 12.

Classification Consistency Rates

Content Area	Grade	Classification Consistency		
		Partially Proficient	Proficient	Advanced
Mathematics	11	-	65%	-
	8	83%	86%	95%
	7	81%	84%	95%
	6	82%	83%	96%
	5	81%	84%	95%
	4	80%	82%	94%
	3	77%	80%	95%
Reading	11	-	63%	-
	8	83%	78%	87%
	7	86%	76%	85%
	6	85%	74%	83%
	5	88%	75%	84%
	4	80%	82%	94%
	3	80%	72%	86%
Science	11	-	67%	-
	8	80%	84%	92%
	5	76%	82%	92%
Social Studies	11	-	63%	-
	9	85%	81%	91%
	6	81%	77%	91%

The classification consistency rates presented for grade 11 represents the percentage of students classified as either (1) both receiving a B or better and Proficient or above on MME or (2) both receiving a B- or worse and Partially Proficient or below on MME. It is not possible to create

classification consistency rates for the Partially Proficient and Advanced cuts for grade 11 since the threshold for those two cut scores is not 50%.

The classification consistency rates presented for the Proficient cut in grades 3 through 9 represent the percentage of students who were consistently classified as either Proficient or above or consistently classified as Partially Proficient or below from one grade level to the next grade level up. The classification consistency rates presented for the Partially Proficient cut in grades 3 through 9 represent the percentage of students who were consistently classified as either Partially Proficient or above or consistently classified as Not Proficient from one grade level to the next grade level up. The classification consistency rates presented for the Advanced cut in grades 3 through 9 represent the percentage of students who were consistently classified as either Advanced or consistently classified as Proficient or below from one grade level to the next grade level up.

Table 12 shows that the lowest classification consistency is from MME to college grades. ACT, Inc. indicated that this level of classification consistency is consistent with that obtained in other states for which they have conducted similar analyses. The remaining classification consistency rates indicate a high degree of stability from grade to grade. The difference between MME to college grades and the remainder of the consistency rates is to be expected for two reasons. First, the rates that are based solely on student achievement scores are high because the classifications are made on the most similar constructs: achievement on two standardized tests of the same subjects. These rates should be higher. Second, the rates for grade 11 are based on less similar but still related constructs: achievement on standardized tests versus college grades in related subjects. These rates should be lower.

Impact of the New Cut Scores

To provide insight into the impact of applying the new cut scores operationally, the new cut scores were applied to the 2010-2011 data. The MME Proficient cut score was used to categorize 11th grades as either Proficient or Not Proficient. The students were also categorized as either having met or not met ACT's College and Career Readiness Benchmark. The number of students classified as both Proficient on MME and meeting the benchmark on ACT are shown in Tables 13 through 15, for mathematics, reading, and science, respectively.

Table 13.

Students Classified as Proficient on MME and Meeting ACT's College and Career Readiness Benchmark in Mathematics

		ACT Math Meeting College and Career Readiness Benchmarks	
		Yes	No
MME Math Proficient	Yes	30,534 25.5%	1,241 1.0%
	No	7,004 5.8%	81,041 67.6%

Table 14.

Students classified as Proficient on MME and Meeting ACT's College and Career Readiness Benchmark in Reading

		ACT Reading Meeting College and Career Readiness Benchmarks	
		Yes	No
MME Reading Proficient	Yes	46,268 38.6%	15,637 13.1%
	No	720 0.6%	57,195 47.7%

Table 15.

Students classified as Proficient on MME and Meeting ACT's College and Career Readiness Benchmark in Science

		ACT Science Meeting College and Career Readiness Benchmarks	
		Yes	No
MME Science Proficient	Yes	22,480 18.8%	7,136 6.0%
	No	4,397 3.7%	85,807 71.6%

It can be seen that although the assessments are different (MME versus ACT), students who are college ready (or not college ready) on one tend to be similarly classified on the other. Tables 16 and 17 relate MME proficiency in mathematics and reading, respectively, to WorkKeys National Career Readiness Certificate levels. Although the certificates consist of three WorkKeys assessments, Reading for Information, Applied Mathematics, and Locating Information, it can be seen that students who are proficient on MME in math and reading tend to earn higher certificate levels than those students who are not proficient on these MME subjects.

Table 16.

Distribution of WorkKeys Certificate Levels, by MME Proficiency in Mathematics

	Not proficient on MME	Proficient on MME
	Frequency	Frequency
No certificate	12,812	55
Bronze certificate	22,413	194
Silver certificate	44,115	10,243
Gold certificate	8,697	20,397
Platinum certificate	8	886

Table 17.

Distribution of WorkKeys Certificate Levels, by MME Proficiency in Reading

	Not proficient on MME	Proficient on MME
	Frequency	Frequency
No certificate	12,454	413
Bronze certificate	19,010	3,597
Silver certificate	24,480	29,878
Gold certificate	1,970	27,124
Platinum certificate	1	893

The percentages of students reaching proficiency targets (i.e., scoring in the upper two levels) with the old cut scores and the new recommended cut scores are presented in comparison in Figures 2 through 5, respectively for mathematics, reading, science, and social studies.

Figure 2. *Comparison of 2010-11 Percentages Meeting Mathematics Proficiency Targets Using Old Cut Scores and New Cut Scores*

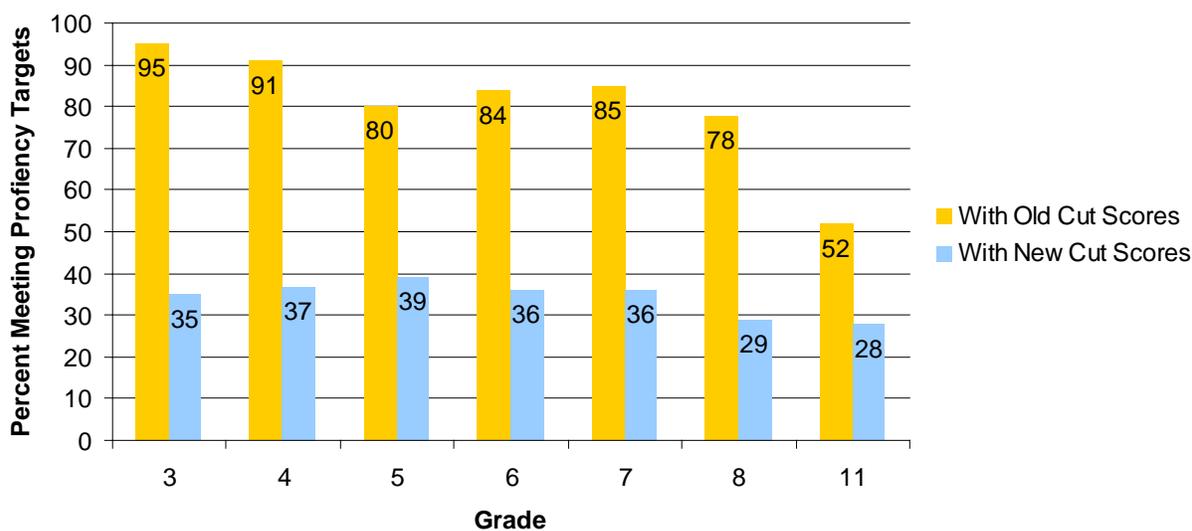


Figure 3. *Comparison of 2010-11 Percentages Meeting Reading Proficiency Targets Using Old Cut Scores and New Cut Scores*

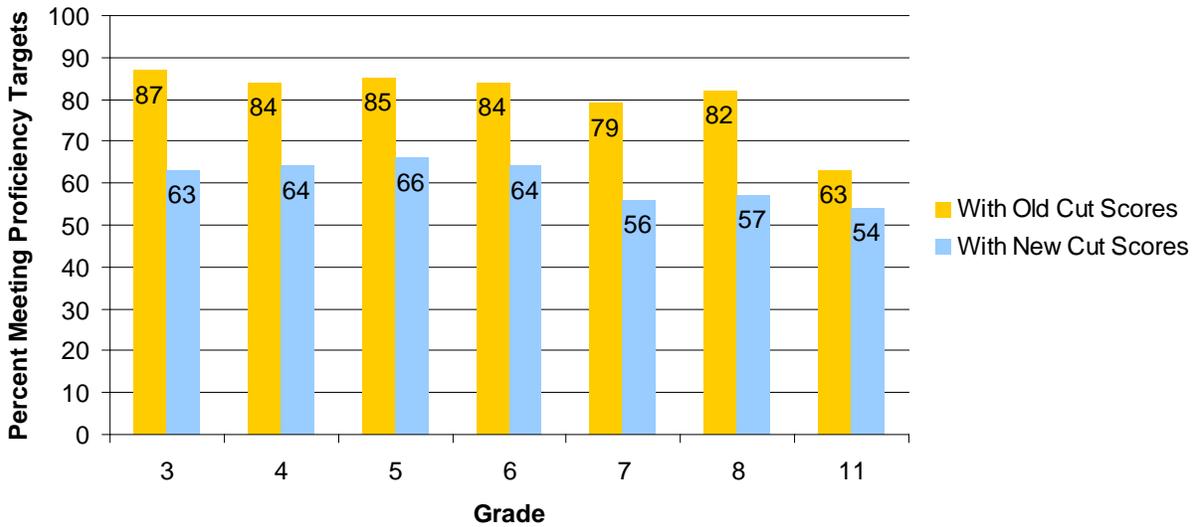


Figure 4. *Comparison of 2010-11 Percentages Meeting Science Proficiency Targets Using Old Cut Scores and New Cut Scores*

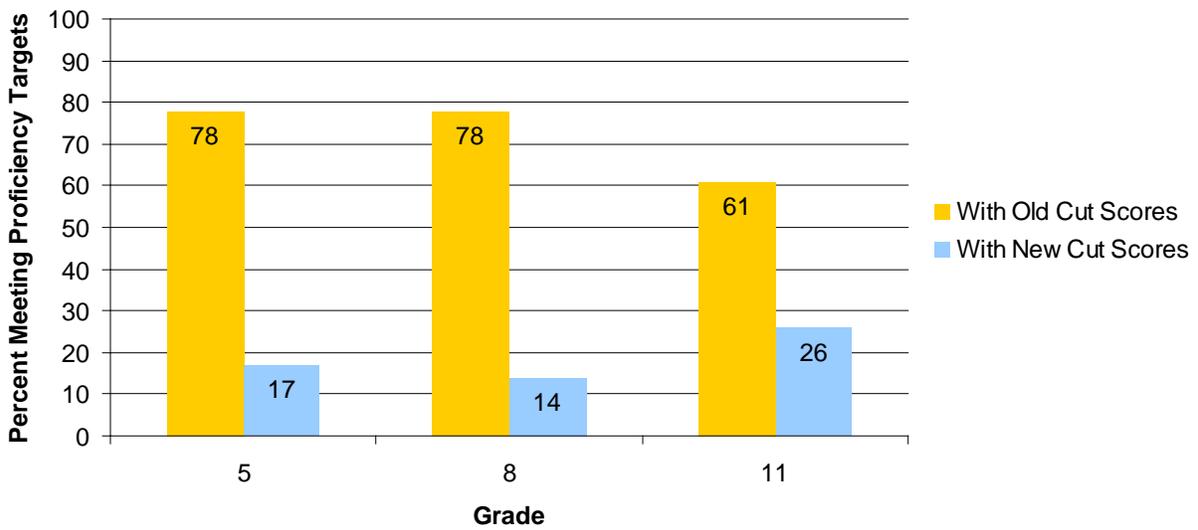
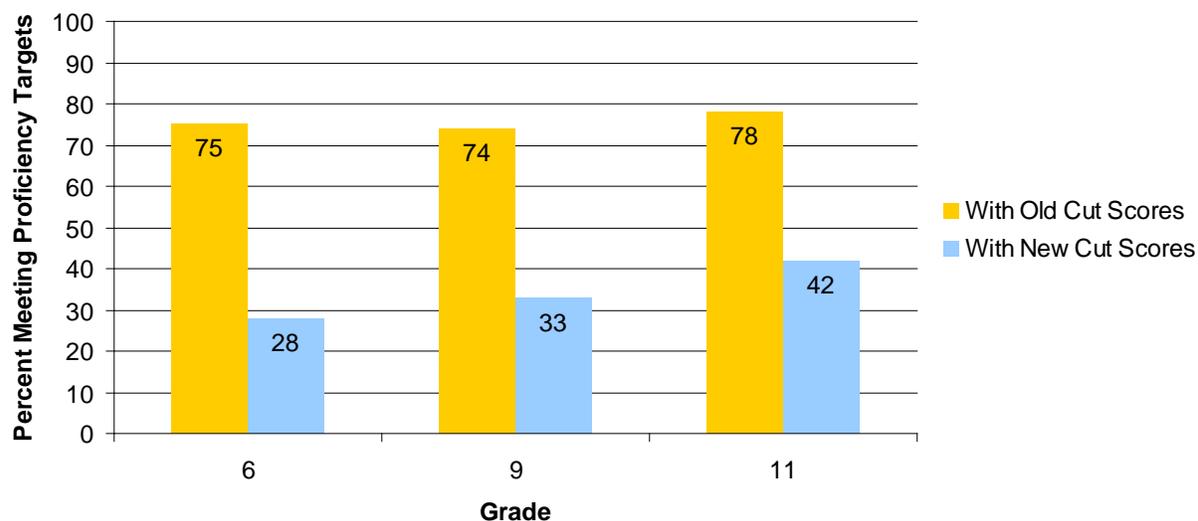


Figure 5. *Comparison of 2010-11 Percentages Meeting Social Studies Proficiency Targets Using Old Cut Scores and New Cut Scores*



The impact the new cut scores would have had had they been in place for the 2010-11 school year is instructive. Had the recommended new cut scores been in place for the 2010-11 school year, the impact would have been as delineated in Figures 6 through 9 for mathematics, reading, science, and social studies respectively.

The same labels as are currently in use remain in use for MEAP and MME: Not Proficient, Partially Proficient, Proficient, and Advanced. The upper two categories are considered “on track” for success in the next level of education (and are thus displayed in green tones), while the lower two categories are considered “off track” (and are thus displayed in red tones).

Figure 6. Retrospective Impact of Cut Scores on 2010-11 MEAP and MME Mathematics

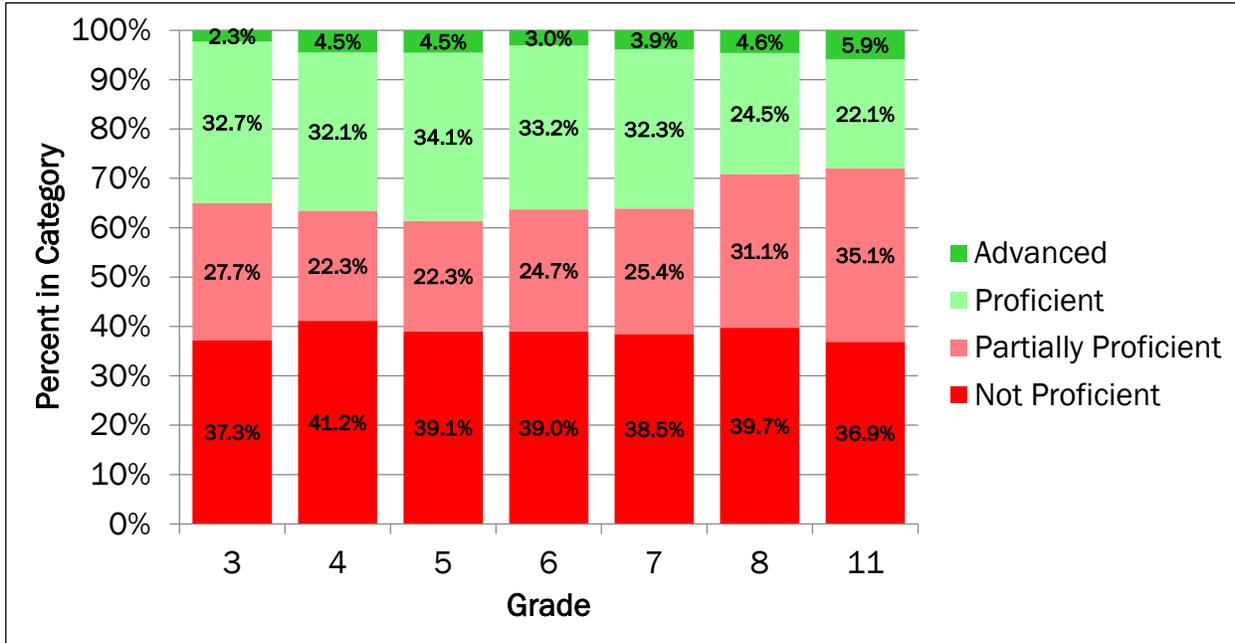


Figure 7. Retrospective Impact of Cut Scores on 2010-11 MEAP and MME Reading

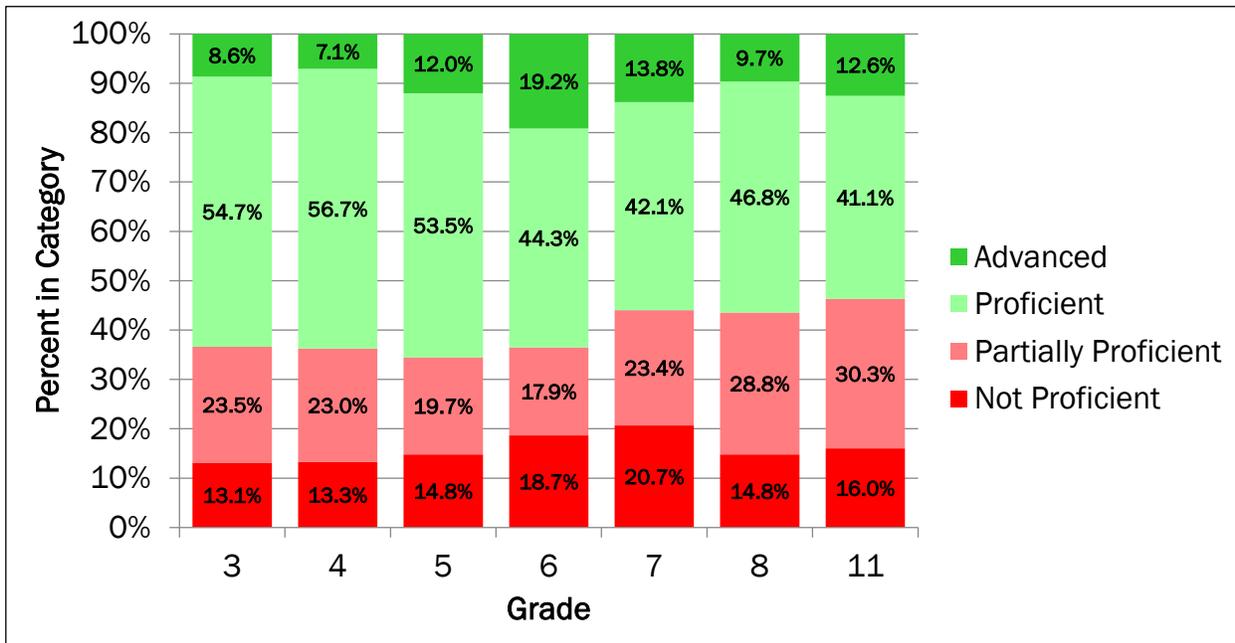
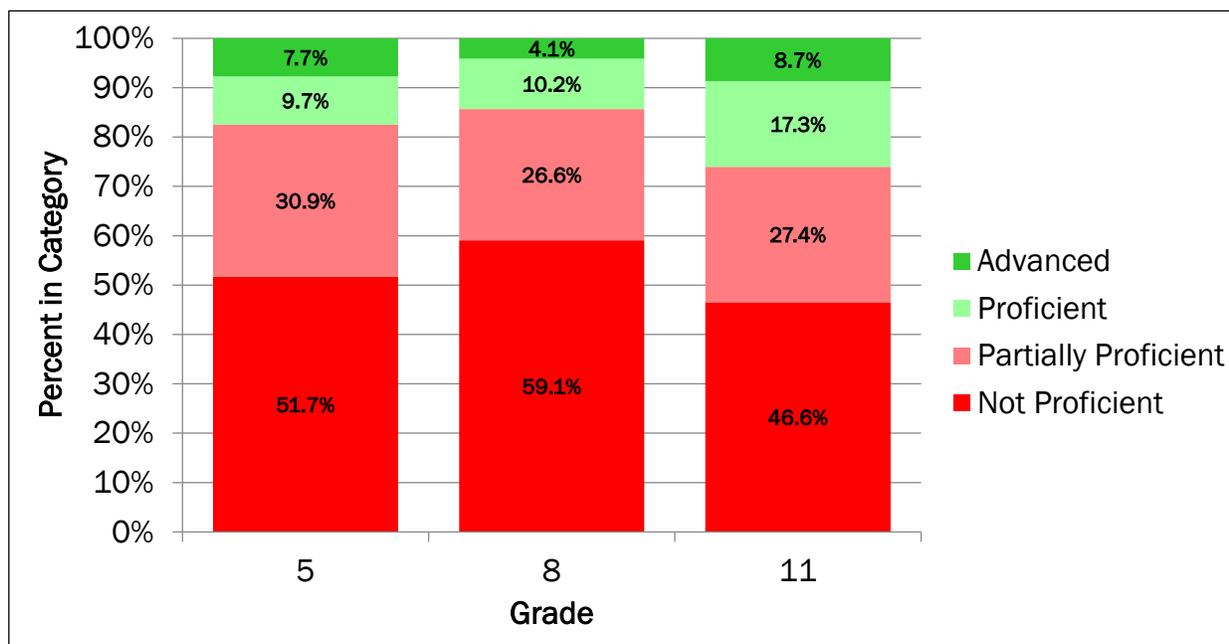
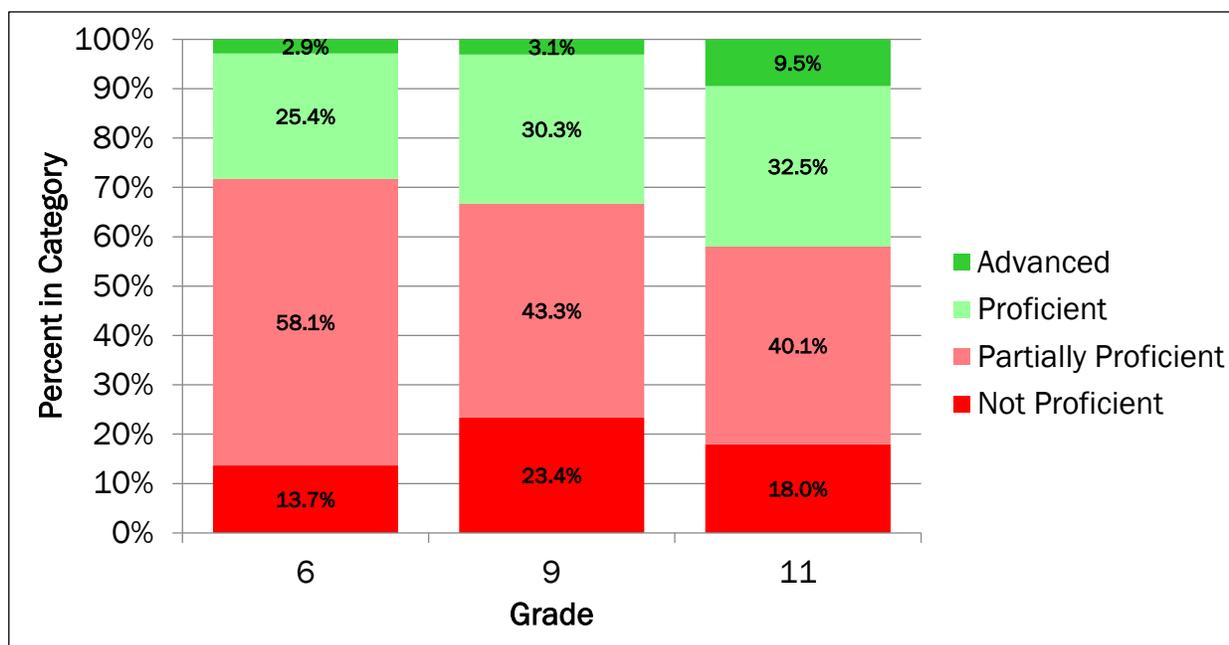


Figure 8. *Retrospective Impact of Cut Scores on 2010-11 MEAP and MME Science*Figure 9. *Retrospective Impact of Cut Scores on 2010-11 MEAP and MME Social Studies*

To demonstrate further the need to set more rigorous cut scores on the MEAP and MME, the approximate percent correct scores needed to pass the MEAP and MME using both the existing

cut scores and the recommended new cut scores are presented below in Figures 10 through 13, respectively, for mathematics, reading, science, and social studies.

Note that the percent correct passing scores are approximate. This is because the difficulty of the MEAP and MME vary slightly from year to year and form to form because different items are presented on the test each year. To compensate for this slight variation from year to year, test equating is conducted, resulting in slightly different percent correct cut scores from year to year and form to form.

Also note that even with the new recommended cut scores, the percent correct scores needed to achieve proficiency range from the high 50s to the low 80s with the majority in the 60s. The differences from grade to grade and subject to subject arise because of differences between the tests from grade to grade and subject to subject.

Figure 10. *Approximate Percent Correct Scores Required to Pass MEAP and MME Mathematics with Old and New Cut Scores*

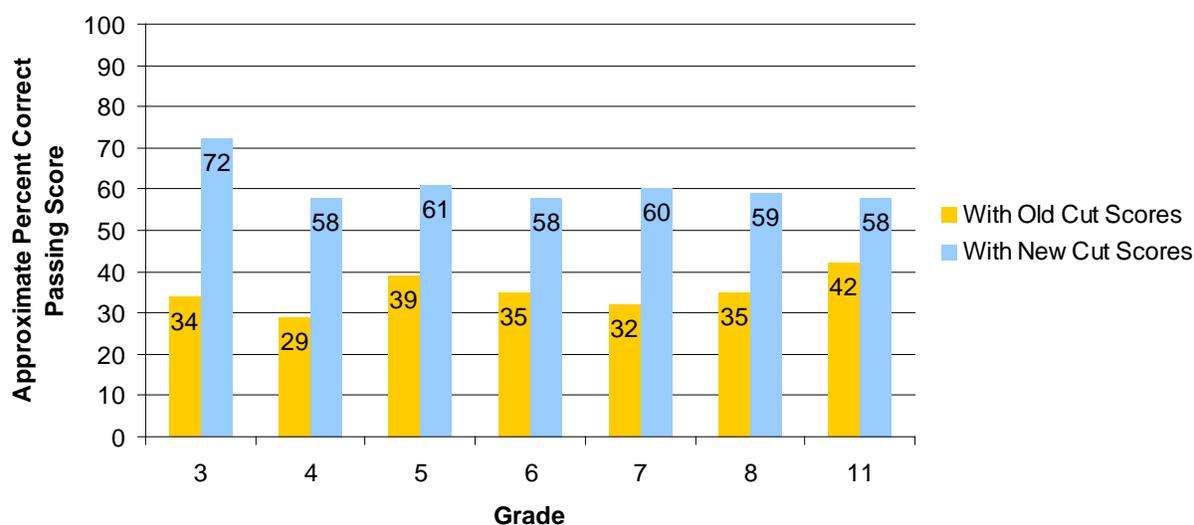


Figure 11. *Approximate Percent Correct Scores Required to Pass MEAP and MME Reading with Old and New Cut Scores*

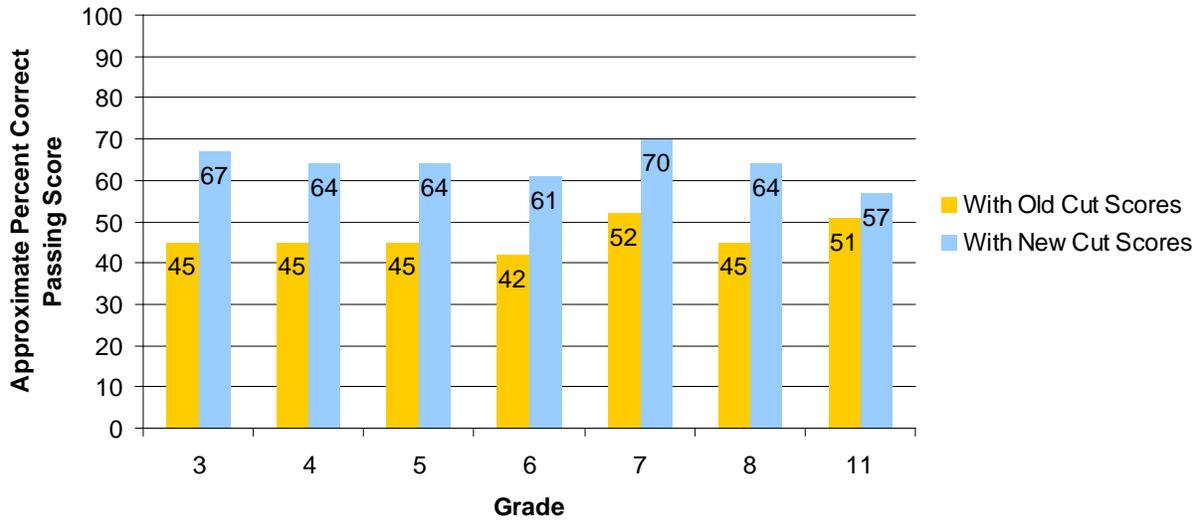


Figure 12. *Approximate Percent Correct Scores Required to Pass MEAP and MME Science with Old and New Cut Scores*

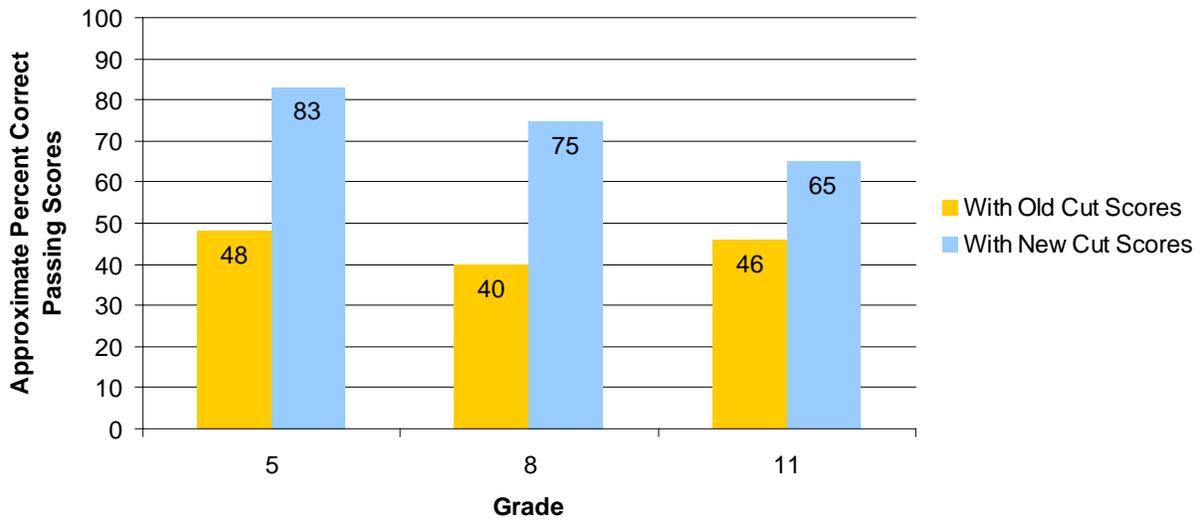
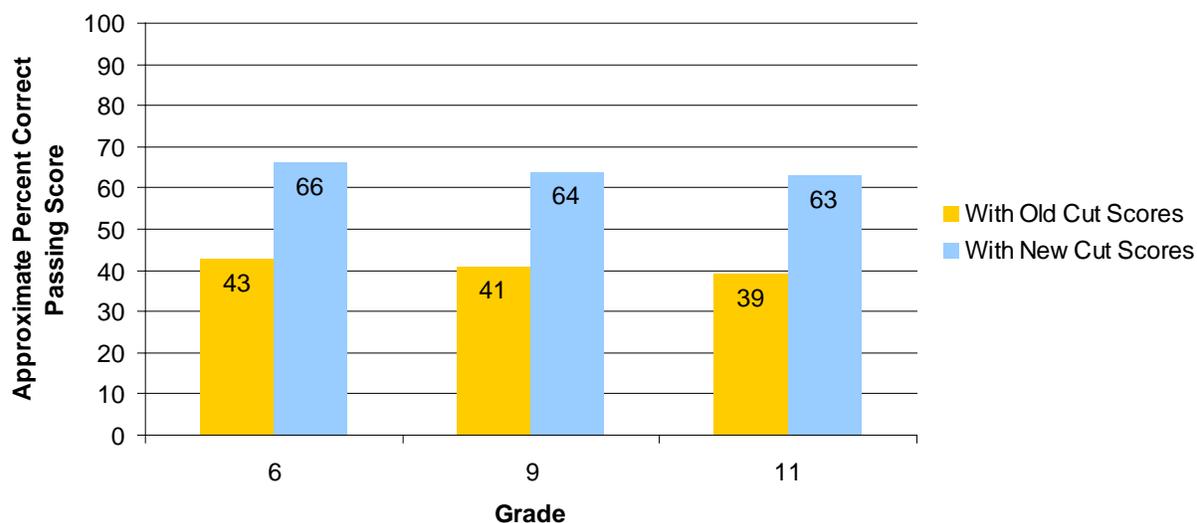


Figure 13. *Approximate Percent Correct Scores Required to Pass MEAP and MME Social Studies with Old and New Cut Score*



Summary

The purpose of this study was to identify new sets of cut scores on the Michigan Educational Assessment Program (MEAP) and the Michigan Merit Examination (MME) in mathematics, reading, science, and social studies, classifying students into four performance levels: Advanced, Proficient, Partially Proficient, and Not Proficient, where Proficient is defined as being on track to succeed in a postsecondary educational experience. The data and methodology used to identify the cut scores was described, as was the impact of applying the new cut scores, in contrast with the previous cut scores, on the 2010-2011 data.