



EDUCATIONAL
RESULTS
PARTNERSHIP



Multiple Measures for Assessment and Placement

White Paper

Presented to the Multiple Measures Work Group of the Common
Assessment Initiative

September 12, 2014

By the Multiple Measures Assessment Project Research Team

Peter Bahr, Ph.D., Associate Professor of Education, Center for the Study of Higher and Postsecondary Education, University of Michigan

Craig Hayward, Ph.D., Director of Research, Planning and Accreditation, Irvine Valley College

John Hetts, Ph.D., Director of Institutional Research, Long Beach City College

Daniel Lamoree, Senior Systems Analyst/Programmer, Mt. San Antonio College

Mallory Newell, Ed.D., Interim Executive Director of Institutional Research and Planning, Foothill-De Anza Community College District

Nathan Pellegrin, M.S., Director of Institutional Research, Peralta Community College District

Ken Sorey, Executive Vice President, Educational Results Partnership (Cal-PASS)

Terrence Willett, M.S., Director of Planning, Research, and Knowledge Systems, Cabrillo College

Scope of Document

This white paper is intended to inform the Multiple Measures work group and the California Common Assessment Initiative (CAI) Steering Committee of the work being conducted by the Multiple Measures Assessment Project (MMAP) research team. We review the context, requirements and practice pertaining to the use of multiple measures in assessing and placing students into California's community college system. High school transcript data are explored in depth and predictive models, derived from logistic regressions and recursive partitioning and regression tree procedures (aka, decision trees), presented and reviewed. It is our hope that this paper will help inform the development of strategies to optimize the use of multiple measures in community college placement and advising. Appendix A includes an overview of the goals of the MMAP project.

Introduction

For over two decades, California's community colleges have been required to assess and place students in the curriculum through means other than a single test score. A 1988 lawsuit, *Romero-Frias et al. v Mertes et al*, initiated system-wide reform requiring the use of "multiple measures" to assess the English and math skills of students. Recent research has found that test scores dominate the placement process across the California system (WestEd, 2014; Willett & Karandjeff, 2014). The influence of multiple measures on student placement is often marginal, however, with the majority of the weight being accorded to the score from the placement test (Hughes & Scott-Clayton, 2011; REL West for WestEd, 2011; WestEd, 2012).

There is growing evidence that multiple measures, such as high school transcripts and non-cognitive variables, can greatly improve the accuracy of the placement process (Belfield & Crosta, 2012; Fuenmayor, Hetts & Rothstein, 2012; Willett, Hayward & Dahlstrom, 2008; Willett & Karandjeff, 2014). At the national level, discourse has centered on the strength of the correlation between students' placement test scores and their likelihood of success in college, and whether additional requirements for remedial education pose undue burdens on students, ultimately hindering their success (Burdman, 2012; Calcagno & Long, 2008; Hayward, 2011; Scott-Clayton & Rodriguez, 2012; Venezia, Bracco & Nodine, 2010; Xu, 2012). Several studies have found that standardized tests are not the strongest predictors of student success, and that high school GPA is a better predictor of college performance (Geiser & Santelices, 2007; Hickson & Dowdy, 2014; Scott-Clayton, 2012).

There may be a shift toward increased reliance on multiple measures, however. Some colleges currently accept completion of coursework at articulated high schools for placement purposes. For example, Long Beach City College now places the majority of recent high school graduates via a placement system that allows students to use the higher of two placements: one based on

their test score and one based solely on their high school transcripts. Since implementing this placement system, the number of students placing into and completing transfer-level course work in English and math has greatly increased (Hetts & Fuenmayor, 2013). In 2012, building on the research at Long Beach City College and the tool LBCC built and made freely available to the California Community Colleges (Fuenmayor et al., 2012), a volunteer effort of local researchers lead by Daniel Lamoree improved the tool and rebranded it as STEPS 2.0 (Willett, 2013). This effort was soon after formally supported by the Research and Planning Group for California Community Colleges (RP Group). The RP Group engaged community colleges in the second iteration of the Student Transcript-Enhanced Placement Study (STEPS 2.0), which built a foundation of evidence concerning the utility of high school transcripts in the college placement process. The work of STEPS resulted in more colleges working to develop equations designed to predict students' success in college-level courses from their high school grades and test scores. Although perfectly predicting students' future success in English and math coursework is not realistic, the results of the work represented a promising addition, and potentially better alternative to test scores (Willett & Karandjeff, 2014).

STEPS found that California Standards Test (CST) scores predicted the level of students' first college English and math course. Since the first course taken in college is the result of the placement process, the conclusion from this finding was stated as "tests predict tests." The corollary of this finding was that student success in college was best predicted from grades received in high school - or "grades predict grades." Other factors were also somewhat predictive. For example, the more high school "A-G" courses for UC/CSU eligibility students completed, the more likely they were to attempt a higher level English course in college.

The MMAP research team has replicated and extended the STEPS work by piloting the development of predictive placement algorithms using all available high school transcript data from Cal-PASS Plus, as well as statewide community college data. These predictive models could play an integral part in a future common assessment system for all CCCs. These models have been developed to create efficient, flexible and powerful rule sets that could be integrated into a dynamic online interface that colleges throughout the system may employ to place and advise students.

Data Sources

At present, the MMAP research team has collected placement data from ACCUPLACER and is awaiting the data from Compass and other sources. In order to make efficient use of project time, Cal-PASS Plus provided a pilot data set containing data on all 112 colleges. The pilot data files (one for English and one for math) include K-12 data and community college MIS data, but no placement data. The files include all of the variables from the STEPS analysis for over 260,000 students. Test results from these initial data files are presented for English and math in

Appendices B and C, respectively. The transcript data used to construct the academic portion of the pilot set of analysis files are generated from high school data submitted to Cal-PASS Plus via a [web interface](#) and the California Community College Chancellor's Office (CCCCO) Management Information System ([MIS](#)) dataset. The structure of the former mirrors the submission requirements for high schools when submitting data to [CALPADS](#) while the latter mirrors the submission requirements for community college when submitting their MIS files. Assessment data are provided for the Early Assessment Program ([EAP](#)), and Standardized Testing and Reporting (STAR) California Standards Test ([CST](#)); these are both submitted by high schools when uploading their high school academic data to Cal-PASS Plus.

File Construction

Two subject-specific flat files have been generated to predict college performance, one for mathematics ($N = 261,994$) and one for English ($N = 256,311$). Each file consists of 62 variables grouped into three categories traversing the academic career of the student as they progress from high school through college: High School, College, and Demographic.

High School

One course is identified as the most appropriate for the student to encapsulate their achievement for mathematics or English during their high school career. In the event a student enrolled in multiple courses for the subject in question, a single course is selected based on a hierarchical sort. This hierarchical sort is defined by six variables; in order of priority, these variables are: (a) highest CBEDS course rank, (b) highest course level, (c) highest A-G status within subject (i.e., primary courses having more priority over elective courses), (d) highest credit units, (e) highest grade the student achieved, (f) and the most recent academic term. Consider a student having enrolled in Algebra I, Geometry, Trigonometry, and Calculus; the aforementioned hierarchical sort would select Calculus as the course that best represents the achievement of the student in mathematics as a result of Calculus having the highest CBEDS course rank relative to the other mathematics courses. However, if a student had enrolled in Algebra I, Geometry, Trigonometry and Math Analysis, the hierarchical sort would rank both Trigonometry and Math Analysis as having the same CBEDS course rank. In such an event, the next variable (i.e., course level) in the hierarchical sort is compared between Trigonometry and Math Analysis. This method of selection continues until a single course is selected.

Two cumulative GPA variables are included: (a) all coursework with the exception of physical education, and (b) all coursework excluding physical education and subject-specific courses (i.e., English or mathematics). For example, the math analysis file has a cumulative GPA variable that excludes both math coursework and physical education coursework. Lastly, variables for EAP and CST results are provided representing the best attempt by the student.

College

Similarly, for the subject in question, a single course is extracted from the college transcript of the student using the same method described above. However, unlike the High School hierarchical sort, when the student has received a grade in two or more applicable sections within the subject, the College hierarchical sort prioritizes the following: (a) least recent term, (b) lowest below college level (i.e., CB21), (c) highest credit units, and (d) highest grade. Essentially, the College hierarchical sort captures the first attempt the student made in the mathematics or English sequence in college.

Demographics

The standardized variables used for MIS reporting were generated for the student, including: gender, ethnicity, disability status, EOPS status, financial aid status, Pell recipient, BOG recipient, and collegiate academic goal. Lastly, a set of term-based variables were calculated. These variables are labeled under the umbrella of *Celerity*. *Celerity* is defined as the number of equivalent full-term college terms having transpired before the student enrolls in a target objective. Consider a student having enrolled in Algebra II, in the Spring 2012 semester during their junior year of high school and then enrolling in college, in Fall 2014, for the first time. The student is then said to have a high school subject to college entry *Celerity* of 2 as two full-term semesters have transpired (i.e., Fall 2013 and Spring 2014). The same student having enrolled in Pre-Algebra, in college, in the Spring 2015 term has a high school subject to college subject *Celerity* of 3.

Non-cognitive Variables

Although MMAP is interested in the potential of ‘affective’ or non-cognitive variables (NCVs), the pilot data-set does not include any NCV data. However, one of the MMAP pilot colleges implemented the School-wide College Student Self-Assessment Survey (SCSSAS) for three years, accumulating thousands of cases. The SCSSAS was administered at the time of assessment, making it a potentially useful source of additional information. The MMAP research team is working to obtain a data set from the college in order to test the utility of these NCVs for predicting future student performance and enhancing student placement algorithms.

Modeling: Description of Methods Tested

Logistic Regression

Logistic regression is a much-used method for analyzing binary outcomes, such as success versus non-success in a college course. The logistic regression model offers a prediction about the likelihood that an individual will experience the outcome of interest, given a set of predictor variables chosen by the researcher. In the example of predicting success in a college course, predictor variables might include high school grade point average (GPA), a student’s most recent score on the math portion of the California Standards Test (CST), and the skill-level of the student’s most recent math course. The logistic regression algorithm attempts to minimize

the residual error between the actual outcomes observed for each student and the outcomes predicted by the regression equation:

$$y = \frac{1}{1 + e^{-(a + b_1x_1 + \dots + b_nx_n)}}$$

In the equation above, y is the outcome variable, x_n are predictor variables, a is the y -intercept, b_n are estimated parameters, and e is the natural or Naperian number. The model output is not terribly intuitive and typically requires someone who is familiar with interpretive techniques to make sense of the results and assess the robustness of the model. Some methods of assessing a logistic model include examining misclassification rates and pseudo R^2 values.

Logistic regression benefits from high quality data and the inclusion of all important variables. It can be sensitive to missing data and the distribution of the underlying data.

The performance of logistic regressions can be evaluated according to the following five dimensions:

1. **Sensitivity** - the percentage of cases that actually experienced the outcome (e.g., "success") that were correctly predicted by the model (i.e., true positives).
2. **Specificity** - the percentage of cases that did not experience the outcome (e.g., "unsuccessful") that were correctly predicted by the model (i.e., true negatives).
3. **Positive predictive value** - the percentage of correctly predicted successful cases relative to the total number of cases predicted as being successful.
4. **Negative predictive value** - the percentage of correctly predicted unsuccessful cases relative to the total number of cases predicted as being unsuccessful.
5. **Misclassification rate** - the number of incorrect predictions divided by the total number of classifications.

Decision Trees

Decision trees are a form of data modeling that results in a set of "if-then" rules rather than a mathematical equation with linear combinations of estimated parameters. For example, let us consider a hypothetical decision tree for determining if a student is likely to be successful in a college introductory statistics class given his/her high school grade point average (GPA), most recent score on the math portion of the California Standards Test (CST), and the skill-level of his/her most recent math course. A resulting tree could be:

If high school GPA is less than or equal to 2.5 and

if CST is less than or equal to 300 and

if last math was intermediate algebra or lower, then **not likely to successful**

if last math was higher than intermediate algebra, then **likely to be successful**

if CST is greater than 300, then **likely to be successful**
 If high school GPA is greater than 2.5 and
 if last math was intermediate algebra or higher, then **likely to be successful**
 if last math was below intermediate algebra and
 if CST was above 300, then **likely to be successful**
 if CST was equal to or below 300 then **not likely to be successful**

These models are intuitive for general audiences and are similar to diagnostic methods used by mechanics and taxonomists. The trees begin with all the data in a single group called a “node.” A computer algorithm then determines how the data will be “split” into two nodes to form the first decision point. In the hypothetical example above, the first decision is whether a student’s high school GPA is above 2.5 or not. Each new node is then evaluated by the algorithm to determine if a node should be split further. This process continues until researcher-defined stopping rules are met. The result is a tree with branches and terminal nodes referred to as “leaves.” In the example above, the leaves are in bold and represent predicted outcomes.

There are a variety of algorithms that can be used to create decision trees, and some can split nodes into more than two resultant nodes. One commonly used decision tree is a classification and regression tree (CART or CRT), which uses binary or two-way splits (Breiman, et al., 1994). The splitting decision tree attempts to create two nodes that most reduce the dissimilarity among the data in each node. In the example, the first split reduced dissimilarities in GPA’s among students in each resulting node. The dissimilarity often is measured using the Gini index:

$$D = 1 - \sum_{i=1}^n p_i^2$$

In the equation above, p is the probability of selecting two individuals with the same characteristics. The algorithm determines which variables at what levels (values of the variables) most reduce dissimilarity, and then continues until a threshold of minimal change is reached. This is a computationally intensive process that is sensitive to the operating rules defined by the researcher. For example, in addition to the stopping rule, the research also can set rules for the minimum number of cases to be included in a resultant node to avoid an overgrown tree with trivial leaves. Decision trees also are sensitive to the variables selected by the research, and a robust analytic procedure will include procedures such as bootstrap aggregating or “bagging,” where a large number of possible trees are grown and compared to determine the importance of each variable.

Misclassification rates are an important measure of model robustness for decision trees. An advantage of decision trees over logistic regressions include interpretability of output, robust handling of scalar and categorical data of any distribution (even with missing data), and inclusion of non-linear relations and interaction effects without additional specifications. As

with any analytic technique, decision trees are improved with high quality data and careful selection of relevant variables, and they are impaired by the omission of important variables.

The misclassification rate is valuable as a tool that allows us to compare the overall performance of various placement models. We are also able to further decompose misclassifications into false positive and false negatives. This further ability is important when we place a differential value on false negative vs. false positive. For instance, a false negative on an airport screening procedure (e.g., missing a bad actor with an explosive device) may be so important that we would prefer a procedure with a ratio of 100,000 false positives to one false negative resulting in many innocent people being thoroughly searched or placed on no-fly lists. In other situations (e.g., fraud detection in retail return situation; pregnancy tests) a false positive may be very problematic, so we would rather err on the side of incurring more false negatives than false positives.

Appendices B and C show examples of predicting success in transfer-level English and one-level below transfer math, respectively, using both logistic regression and decision trees based on preliminary Cal-PASS Plus data. Note that an individual student could be predicted to succeed by one model and not succeed with another model. One solution to this situation is to have several different predictions from a variety of models and calculate an average prediction. While more complex, combining multiple predictions can help mitigate particular biases inherent in different analytical techniques.

Establishing Validity

Standards set by the California Community College Chancellor's Office require placement tests to demonstrate a correlation of 0.35 between the placement instrument and an appropriate criterion in order to establish predictive validity (Glasnapp & Poggio, 2001). However, what is the correct criterion? As noted by Lagunoff, Michaels, Morris and Yeagley (2012), "neither tests nor test scores are validated; rather, 'it is the claims and decisions based on the test results that are validated' (Kane, 2006, p. 60)". An effort to determine the validity of a placement test must achieve clarity on the nature of the outcome that the placement is supposed to facilitate. Clearly, placement is intended to increase student success. But which specific form of student success? Should the outcome be placement at a level of the sequence that maximizes the chance of success in a single course? Or is placement intended to put a student into a course that maximizes his or her chances of completing the gatekeeper transfer-level course?

As a result of drop-outs and attrition from the course sequence, it is likely more damaging for students to be inappropriately placed at lower levels than for them to be inappropriately placed at a higher level, at least in terms of their chances of ever completing the sequence. However, some may feel it is better for students to repeat material they know than to be set up for failure and discouragement. The balance between false negatives and false positives in placement is

important to empirically analyze and for college personnel to discuss in detail when making changes to placement practices.

Algorithms, Placement Decisions, and the Role of Uncertainty

Placement decisions can be thought of as a series of evaluations about whether a student should be placed into a course of lower skill or moved into a course of higher skill. The deciding factor in this decision is the student’s likelihood of success vis-a-vis the decision criterion (i.e., the focal student success outcome). Whether the criterion is the prediction of success in the course or the prediction of successfully completing the gatekeeper course at the end of the sequence, the correct placement decision, logically and ethically, is the one that maximizes a student’s chances of success. The critical question, then, is what decision criterion should be used?

The process of maximizing a student’s chance of success can be conceived as a series of hypothetical placements. Beginning with the transfer-level, gatekeeper course, the odds of a student succeeding can be calculated for each hypothetically possible placement level. Similarly, the probability of the student ever completing the gatekeeper course at the end of the sequence can be calculated, as well. An example of such a matrix is given in Table 1, below.

Table 1. Hypothetical probability of success in the initial (placement) course and in the final transfer-level gatekeeper course for a given student at each level of placement

	Transfer level	One level below	Two levels below	Three levels below	Four levels below
Probability of success in initial course	0.45	0.55	0.65	0.75	0.75
Probability of success in gatekeeper course, given placement	0.45	0.45	0.40	0.25	0.20

Given the hypothetical predicted probabilities of success in Table 1, how should the student be placed? A rule set that seeks to maximize odds of success at the level of the initial placement course might place the student at three levels below transfer-level because that is the course in which the student has the highest probability of being successful (in this hypothetical case, it is tied with the course that is four levels below transfer-level). However, a decision rule that is based on maximizing the student’s chances of completing the sequence might place the student into the transfer-level course, if the rule set dictates that the student should be placed into the level that provides the greatest probability of completing the sequence. Alternatively, the student might be placed at two levels below if the rule set were constructed such that students were placed into a lower level in the event of tied predicted probabilities of success.

While Table 1 is based on hypothetical data, these types of prediction matrices can be produced for any actual student.

Data Quality

What to do in cases of missing data? Formulas and decision sets can be constructed based on any specific set of available information, but their power will be diminished by the loss of information. In some cases, usable “sub-models” can be created, if only one or two pieces of information is missing. However, when all transcript information is missing, then an alternative set of multiple measures will need to be used.

When students interact with the online placement test, they should be asked to complete a set of basic multiple measures questions so that a secondary multiple measures algorithm can be used when high school transcript information is not directly available. These questions may also be able to add to the predictive power of the transcript-based multiple measures model because they will be able to provide more recent information than will be available in the high school transcript database (e.g., information about senior year courses and grades).

Questions that could be included as additional multiple measures questions include:

- What was the last math/English course that you completed in high school?
- What was the grade you received in your last math/English course?
- How important is it to you to succeed in college?
- How important is it to your family and others in your life that you succeed in college?

These questions are among the most predictive of those that are typically in use across the CCC system (CCCCO, 1998; WestEd, 2012).

Caveats and Limitations

Predictive analytic tools need maintenance. As data sources change (e.g., the CST has been phased out and the new Smarter Balanced test will take its place), models will need to be updated in order to stay relevant. Regular validation and tuning, probably at least annually, should be anticipated. As it is possible for the original predictive relationships to shift over time, their validity must be reassessed in order to maintain confidence in the appropriateness of the model and associated placement rule set.

Current data sets do not systematically mark English and math sections that are accelerated. Acceleration is a fast-growing trend in the CCC system. It essentially allows some students to waive their placement and enroll in a higher level course, typically one-level below transfer. These accelerated pathways possibly create noise in the MMAP models due to what Calcagno and Long (2008) describe as “noncompliance.” The Chancellor’s Office MIS tables do not have

acceleration data elements as of yet. If we were able to tag accelerated course sections, they could be properly handled in the predictive analytic models.

Policies and Procedures

While the use of multiple measures may be mandated, implementation of multiple measures assessment is anything but uniform across California. For example, the California State University's Early Assessment Program (EAP) is now being accepted at some community colleges where a qualifying EAP score exempts students who meet CSU requirements from retaking the placement tests. Other colleges may also accept high scores on the SAT or ACT as well as AP credits for equivalency. As described above, some colleges accept completion of coursework at articulated high schools for placement purposes.

Under the local, decentralized governance structure of the California Community College system, each of the 112 colleges is required to pay for the administrative responsibility of assessing students in English, mathematics, and English as a Second Language (ESL). Each college is also responsible for selecting or developing its own assessments and determining the cut scores that correspond with various course placement levels (Venezia, Bracco & Nodine, 2010).

Each college sets its own test cut scores, meaning that students scoring below that level lack readiness for enrollment in college-credit courses across the state and, these "cut scores vary considerably. For example, in mathematics, cut scores to place into transfer-level mathematics courses can range anywhere between 20 points on the ACCUPLACER Transfer-Level Mathematics exam. Students can therefore receive different placements at two different colleges based on the same test scores, even within the same district (Venezia, Bracco & Nodine, 2010).

A concern among some community college counselors and faculty is the degree of alignment between high school and college curricula. One recommendation from the STEPS project and other similar work is for high school and community college faculty to collaboratively discuss course alignment (Hodara, Jaggars & Karp, 2012; Willett & Karandjeff, 2014). The student who passed the required English and math in high school, but does not perform well on the day of the placement test would be expected to benefit from alignment and articulation that encourages high school coursework to be utilized in college placement.

Variation in retake policies across colleges also exists and poses delays for many students. Venezia, et al. (2010) provided an example where the mean (arithmetic average) wait time statewide to retake a mathematics assessment was 160 days while the median was 73 days, meaning that half of wait times were over 73 days. The mean being higher than the median

indicates that a few colleges had very long waiting periods highlighting. Variation also exists in the acceptance of scores from other colleges within the system. While students may move from college to college to pick up the classes they need to reach their goal, only some colleges accept all scores, while others accept scores from colleges within the district or region, and some rarely accept scores from other colleges at all (Venezia, Bracco & Nodine, 2010).

Placement decisions typically fall under faculty purview. Faculty play a key role in any changes that will happen at their college or statewide. In a review of placement practices, Venezia, et al. (2010) found that discipline faculty members are typically responsible for establishing assessment and placement policies, while counselors implement the placement decisions by counseling students on the courses in which they should enroll.

A scan of multiple measures practices in California revealed that although the majority of CCCs apply multiple measures as part of a weighted algorithm, for a substantial minority of community colleges (about 25%), counselors are applying multiple measures qualitatively, based on their judgment (REL West for WestEd, 2011). With the advent of the Student Success Act of 2012 (SB 1456), the role of counselors is evolving and expanding. There is a greater focus on helping students complete an education plan and the other steps necessary to be matriculated, because only fully matriculated students will receive enrollment priority and access to the most in-demand classes. This expanded role may lead to a greater dependence on counselors in the student placement process.

The wide variation across the state makes the process of student placement multifaceted and complex. This document provides information on how to use predictive modeling to inform placement decisions in light of students' probability of success at the community college based on high school grades and test scores.

Further Research

The MMAP research team will continue to develop more complete and refined placement models in the coming months. As we progress along that pathway, some further areas of research include:

- Exploring models that combine multiple measures and test scores. By using test and transcript data together in a statistically optimized manner, it should be possible to improve placement accuracy over what is possible using either source alone.
- Exploring the current implementation planning efforts for the Smarter Balanced Assessment Consortium (Smarter Balanced) assessments, in order to assess how Smarter Balanced scores could be integrated into the CAI placement system.
- Evaluating how the adoption of Common Core standards may facilitate conversations across the state about increasing alignment between high school and community

college curricula and thereby streamline and/or improve the accuracy of placement processes.

- Continuing to explore and seek input on the appropriate student success criterion for establishing predictive validity of the placement system. As shown in Table 1, completion of gatekeeper math or English courses as a criterion will result in a very different placement system than one that seeks to optimize odds of success in a single course.

References

- Bailey, T., Jeong, D., & Cho, S. (2010). *Student progression through developmental sequences in community colleges* (CCRC Brief). New York: Columbia University, Teachers College, Community College Research Center.
- Belfield, C. & Crosta, P.M. (2012). *Predicting success in college: The importance of placement tests and high school transcripts* (CCRC Working Paper No. 42). New York: Columbia University, Teachers College, Community College Research Center.
- Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1994). *Classification and regression trees*. Chapman and Hall: New York, New York.
- Brown, R. & Conley, D. (2007). *Comparing state high school assessments to standards for success in entry-level university courses*. Eugene, OR: Center for Educational Policy Research. Retrieved Aug 30, 2014, from www.s4s.org/cepr.articles.php.
- Brown, R. & Niemi, D. (2007). *Investigating the alignment of high school and community college assessments in California*. San Jose, CA: National Center for Public Policy in Higher Education.
- Burdman, P. (2012). *Where to Begin? The Evolving Role of Placement for Students Starting College*. Boston, MA: Jobs for the Future.
- Calcagno, J. C., & Long, B. T. (2008). *The impact of postsecondary remediation using a regression discontinuity approach: Addressing endogenous sorting and noncompliance* (NCPR Working Paper). New York, NY: National Center for Postsecondary Research.
- California Community College Chancellor's Office (2014, July 16). *Common assessment initiative frequently asked questions* [PDF document]. Retrieved from <http://cccassess.org/downloads/category/10-2014-cai-documents-and-reports?download=38:faq>
- California Community College Chancellor's Office. (1998). *Multiple measures and other sorrows: A guide for using student assessment information with or instead of test scores* [PDF document]. Retrieved from <http://extranet.cccco.edu/Portals/1/SSSP/Matriculation/Assessment/MultipleMeasuresAndOtherSorrowsMarch1998.pdf>
- Fuenmayor, A., Hetts, J., & Rothstein, K. (2011). *Assessing assessment: Evaluating models of assessment and placement*. Long Beach, CA: Long Beach City College.
- Geiser, S., & Santelices, M.V. (2007). *Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes*. Research & Occasional Paper Series CSHE.6.07). Berkeley, CA: University of California, Center for Studies in Higher Education.

Glasnapp, D., & Poggio, J. (2001). *Standards, policies and procedures for the evaluation of assessment instruments used in the California community colleges* (4th Edition). Lawrence, KS: Center for Educational Testing and Evaluation, University of Kansas.

Hayward, C. (2011). The transfer velocity project: A comprehensive look at the transfer function. *Journal of Applied Research in the Community College*, 18(2), 21-32.

Hetts, J. J., & Fuenmayor, A. (2013). *Promising pathways to success: Using evidence to dramatically increase student achievement*. Long Beach City College, Long Beach, CA. Retrieved August 15, 2014 from <http://www.lbcc.edu/IE/Research/Researchdocs.cfm>.

Fuenmayor, A., Hetts, J.J., & Rothstein, K. (2012, April). *Assessing assessment: Evaluating models of assessment and placement*. Paper presented at the annual conference of the Research and Planning Group, Pomona, CA.

Hickson, J. & Dowdy, J. G. (2014). *Alternative Course placement: Practices at community colleges*. Education Advisory Board: Community College Forum.

Hodara, M., Jaggars, S. S., & Karp, M. M. (2012). *Improving developmental education assessment and placement: Lessons from community colleges across the country*. New York, NY: Community College Research Center, Teachers College, Columbia University.

Hughes, K., & Scott-Clayton, J. (2011). *Assessing developmental assessment in community colleges* (Working Paper No. 19). New York, NY: Community College Research Center, Teachers College, Columbia University.

Lagunoff, R., Michaels, H., Morris, P., & Yeagley, P. (2012). *A framework for evaluating the technical quality of multiple measures used in California community college placement*. Sacramento, CA: WestEd.

Scott-Clayton, J. (2012). *Do high-stakes placement exams predict college success?* New York, NY: Community College Research Center, Teachers College, Columbia University.

Scott-Clayton, J., & Rodriguez, O. (2012). *Development, discouragement, or diversion? New evidence on the effects of college remediation* (NBER Working Paper No. 18328). Cambridge, MA: National Bureau of Economic Research.

Seymour-Campbell Matriculation Act, California Education Code § 78210-78218 (1986).

Venezia, A., Bracco, K. R., & Nodine, T. (2010). *One-shot deal? Students' perceptions of assessment and course placement in California's community colleges*. San Francisco: WestEd.

REL West for WestEd. (2011). *Types of multiple measures used in California community college mathematics, English, and English as a second language course placement: Summary report of survey results*. San Francisco: WestEd.

Willett, T. (2013). *Student Transcript-Enhanced Placement Study (STEPS) Technical Report*. Berkeley, CA: The Research and Planning Group for California Community Colleges.

Willett, T. & Karandjeff, K. (2014). *Stepping up: Progression in English and math from high school to college*. Berkeley, CA: The Research and Planning Group for California Community Colleges.

Willett, T., Hayward, C., & Dahlstrom, E. (2008). *An early alert system for the remediation needs of entering community college students: Leveraging the California Standards Test*. Report 2007036. Encinitas, CA: California Partnership for Achieving Student Success.

Xu, D. (2011, March). *The impact of different levels of English remediation on underprepared students in community colleges*. Presentation at the Association for Education Finance and Policy, Boston, MA.

APPENDIX A: About MMAP

Goal of Multiple Measures Assessment Project

The Common Assessment Initiative (CAI), also known as CCCAssess, has established a goal of having a common assessment system ready for implementation across the entire California Community College system by the end of 2015 (CCCCO, 2014). The CAI is evaluating the validity and usefulness of standardized assessment instruments from several major publishers. Additionally, the CAI created the Multiple Measures Assessment Project (MMAP) research team, tasking the group with evaluating the predictive power of that category of evidence that is known generally as “multiple measures.” Multiple measures typically involve gathering background data from students via a questionnaire, interview, or survey, and then combining that information with the results of a standardized placement test, such as the Accuplacer or Compass tests (Lagunoff, Michaels, Morris & Yeagley, 2012; Seymour-Campbell Matriculation Act, 1986).

MMAP is a collaborative effort among Cal-PASS Plus, the RP Group, and the California Community Colleges Chancellor’s Office that strives to build a data warehouse and communications strategy to support community colleges in facilitating placement using multiple measures assessment. The overarching goal of the project is to fundamentally change the landscape of student success in order to:

- enable more efficient student placement and transitions from K-12 to college.
- support informed changes to K-12 and college curricula, instruction, and support services related to college academic preparation and course taking.
- reduce costs associated with basic skills courses, for both colleges and students.
- create more equitable outcomes for high-potential minority and low-income students as they enter and complete community college and university degrees and certificates.

To realize this overarching goal, MMAP has four primary objectives:

1. Develop a secure, large, and robust data warehouse within Cal-PASS Plus to collect, store and analyze multiple measures including assessment data from COMPASS and Accuplacer, College Board data including SAT, ACT, EAP, and AP tests, course, grade, and testing data from K-12 schools, and student demographic data from the CCCApply application for California Community Colleges.
2. Research, analyze and validate known multiple measures data points using predictive analytics, such as decision trees, to identify new data points that can serve as effective multiple measures and seek feedback on the outcomes from the Multiple Measures Assessment working group.

3. Develop online tools and an easy-to-navigate user interface, hosted by Cal-PASS Plus to allow colleges to use predictive multiple measures to improve their placement processes and increase student success.
4. Develop support tools based on best practices and conduct training for practitioners to embed the data, tools, and multiple measures in local college placement procedures and decisions.

Pilot Colleges

Fourteen community colleges from across the state representing wide geographical differences as well as student populations in which they serve have agreed to participate as pilot colleges for MMAP. Pilot colleges will review and provide feedback on the user interface for placement and provide comparative analysis of MMAP placement compared to their current system in regards to where students are placing. The pilot colleges include:

- Allan Hancock
- Bakersfield College
- Cañada College
- Contra Costa College
- Cypress College
- Foothill-De Anza Community College District (Foothill and De Anza Colleges)
- Fresno City College
- Irvine Valley College
- Chabot-Las Positas Community College District (Chabot and Las Positas Colleges)
- Rio Hondo College
- San Diego City College
- Santa Barbara City College
- Santa Monica College
- Sierra College

APPENDIX B: Example pilot analyses for predicting success in transfer-level English

The data used in the following analyses are based on a set of test data from Cal-PASS Plus used primarily to screen for data quality issues and to test modelling procedures. The models below are for illustrative purposes only and should not be used for actual placements.

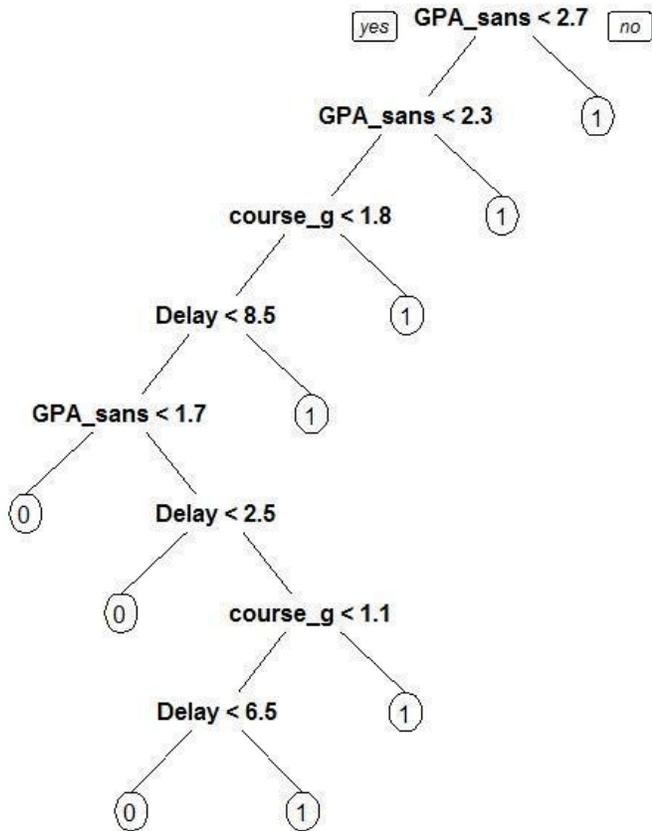


Figure B1. Decision tree predicting success in transfer-level English, misclassification=29% (n=98,483).

1= Predicted success; 0 = Predicted non-success

GPA_sans is a student's cumulative high school GPA excluding grades in English.

course_g is a student's grade in their most recent high school English course.

Delay is the number of primary terms between last high school English course and first college English course.

Note that GPA_sans was the most important predictor variable as determined by the random forest aggregated bootstrap method.

Table B1. Logistic regression coefficients for predicting success in transfer-level English (n = 86,318; missing cases=12,165).

Variable	Coef.	Odds Ratio	SE	Sig.	95% CI Lower Bound	95% CI Upper Bound
Delay	-2E-05	1.00	0.003	0.994	0.995	1.005
CBEDS_rank	1E-02	1.01	0.004	0.008	1.003	1.017
course_g	3E-01	1.29	0.011	0.000	1.272	1.316
A2GB	6E-02	1.06	0.017	0.000	1.029	1.097
cst_ss	2E-03	1.00	0.000	0.000	1.002	1.002
GPA_sans	5E-01	1.67	0.021	0.000	1.633	1.717
_cons	-2E+00	0.14	0.008	0.000	0.124	0.156

Nagelkerke/Cragg & Uhler's R square = 0.113

Note that the square root of the pseudo R square (i.e., the multiple R, analogous to the bivariate Pearson's r coefficient) is 0.34, which nearly attains the 0.35 criterion threshold established by Glasnapp & Poggio (2001) in the Chancellor's Office standards manual.

Delay is the number of primary terms between last high school English course and first college English course.

CBEDS_rank is the level of a student's last high school English course with higher values indicating more advanced courses.

course_g is a student's grade in most recent high school English course.

A2GB indicates a student met their English "A-G" UC/CSU eligibility requirement.

cst_ss is the scaled score of a student's English CST.

GPA_sans is a student's cumulative high school GPA excluding grades in English.

APPENDIX C: Example pilot analyses for predicting success in one level below transfer-level math

The data used in the following analyses are based on a set of test data from Cal-PASS Plus used primarily to screen for data quality issues and to test modelling procedures. The models below are for illustrative purposes only and should not be used for actual placements.

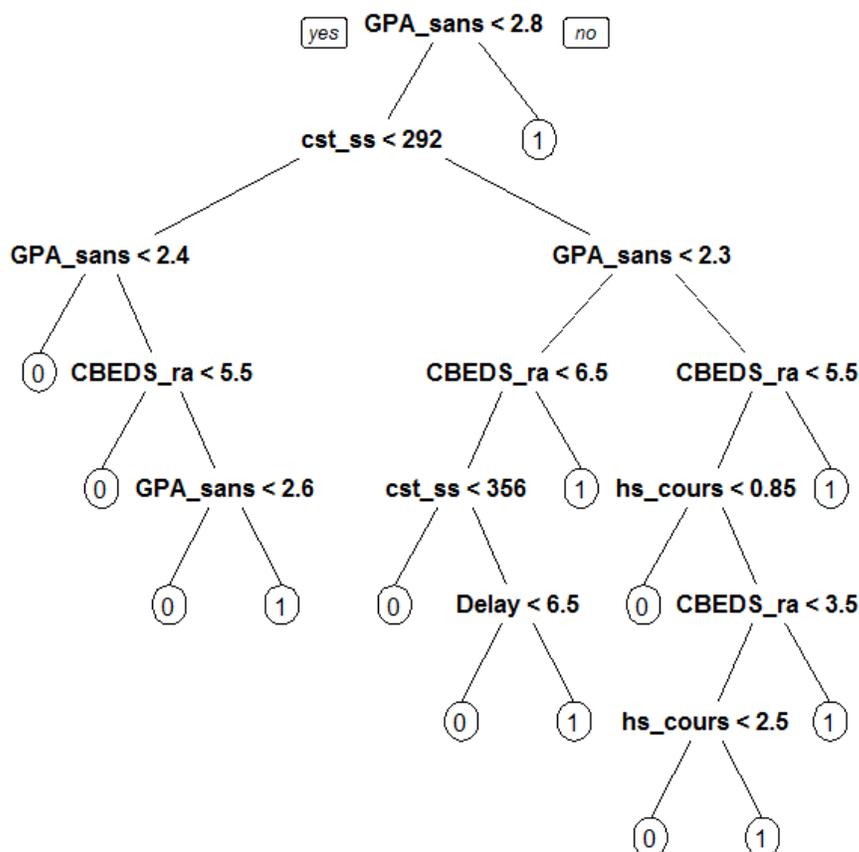


Figure C1. Decision tree predicting success in one level below transfer-level math, misclassification=36%

1 = Predicted success; 0 = Predicted non-success

GPA_sans is a student's cumulative high school GPA without grades in math.

hs_course is a student's grade in most recent high school math course.

Delay is the number of primary terms between last high school math course and first college math course.

Note that GPA_sans was the most important predictor variable as determined by the random forest aggregated bootstrap method.

Table C1. Logistic regression coefficients for predicting success in one level below transfer-level math

Variable	Beta	Odds Ratio	SE	Sig.	95% CI Lower Bound	95% CI Upper Bound
Delay	0.034	1.035	0.003	0.000	1.028	1.042
CBEDS_rank	0.210	1.252	0.007	0.000	1.217	1.252
hs_cours	0.210	1.23	0.008	0.000	1.214	1.254
A2GB	0.103	1.11	0.032	0.001	1.042	1.181
cst_ss	0.002	1.00	0.000	0.000	1.002	1.002
GPA_sans	0.493	1.64	0.014	0.000	1.592	1.683
_cons	-3.245	0.04	0.061	0.000		

Nagelkerke/Cragg & Uhler’s Pseudo-R square = 0.115

The variable with the largest odd ratio was GPA_sans, which was also the most important variable in the decision tree.

Note that the square root of the pseudo R square (i.e., the multiple R, analogous to the bivariate Pearson’s r coefficient) is 0.34, which nearly attains the 0.35 criterion threshold established by Glasnapp & Poggio (2001) in the Chancellor’s Office standards manual.

Delay is the number of primary terms between last high school math course and first college math course.

CBEDS_rank is the level of a student’s last high school math course with higher values indicating more advanced courses.

hs_cours is a student’s grade in most recent high school math course.

A2GB indicates a student met math “A-G” UC/CSU eligibility requirement.

cst_ss is the scaled score of a student’s math CST.

GPA_sans is a student’s cumulative high school GPA without grades in math.