



College Readiness Mathematics Test (CRMT) Working Meeting
University of Washington, Roosevelt Commons B, 416
April 7, 2008, 10:00-3:00

Meeting synopsis:

1. Welcome and report on current project status
 2. Report on WI08 pilot test results
 3. Implementation at K-12 and CTCs
 4. Setting the college readiness cutoff (the "single bright line")
1. **Welcome and report on current project status.** Nana Lowell, Director, OEA, welcomed the group and summarized project progress to date. There have been three changes to the original plan as reflected in the [CRMT development timeline and renewal cycle](#).

Development of new General MPT. The initial project proposal specified that the CRMT would be created by expanding the content of the existing Intermediate Math Placement Test (MPT-I) to cover the College Readiness Mathematics Standards (CRMS). The revised test would serve both as an Intermediate placement test and as a test for college readiness. However, at the June 2007 working group meeting it was determined that the content of the existing MPT-I should remain unchanged to optimize placement into precalculus courses, and a new broader test should be created for general math placement.

Switch to three- or four-alternative items. Existing APTP tests are comprised of five-alternative multiple-choice items and new CRMT items were constructed in the same format. However, results of the autumn 2007 pretest revealed that most items did not have four functioning distractors. Based on these findings, we selected the best pretest items and then constructed three-alternative and four-alternative versions of those items. These were then used to make new tests that were distributed to students in a winter pilot test. Item reliabilities were found to be equivalent for the three- and four-alternative formats, but many of the four-alternative items had non-functioning distractors. We noted that revising APTP test items to three-alternative format would have additional advantages as described below, and that this format is cited as superior in the testing literature.

Analysis of data from the APTP database showed that existing MPT-I and MPT-A items also had a limited number of functioning distractors and we concluded that all APTP tests should be revised to either three- or four-alternative format. The choice between these formats must be made by the CRMT working group and APTP governing council, but in any case additional empirical information is needed regarding the functioning of three- and four-alternative MPT-I and MPT-A items.

Delayed introduction of the new test. We had initially expected to first administer the revised MPT-I during the 2008 statewide testing; however, we determined instead to use this administration to examine MPT-I and MPT-A items in three- and four-alternative formats. This change was driven by the need to modify the format of existing APTP tests, but also allows us to develop and test logistical procedures for regular item pretesting as part of an ongoing test renewal cycle (see link above). The cycle will begin each fall with the creation of new items, followed by a preliminary items pretest in winter, and a final pretest during spring statewide testing. Testing time during statewide testing will be extended from 60 to 75 minutes to allow the

inclusion of 6-7 pretest items. These items will not count toward students' test scores. Given the expected volume of MPT-G testing at the CTCs, we expect to introduce a new version of this test each year. We may create new versions of the MPT-I and MPT-A only every other year, but this will depend on the volume of use.

Other changes to the APTP system. In addition to introducing a new test to the APTP battery, and implementing an ongoing test renewal cycle, we are encouraging more formal, standardized test administration procedures. Drafts of two new test administration manuals have been created. The [Testing Center manual](#) will be used for administrations throughout the year at both two-year and four-year college testing centers. The [Testing Site manual](#) will apply to administrations during May and June statewide testing and for other contracted testing sites such as may be established to test high school juniors. All test administrators (whether at Testing Centers or Sites) will be asked to sign a [Confidentiality Agreement](#) stating that they accept responsibility for ensuring security of the tests during storage and administration.

The standard APTP answer sheet will be modified to correspond to the reduced number of response alternatives, and reflect modifications to the APTP testing database. The University of Washington is contributing one full-time programming year to creating a new APTP testing database to score, report, and archive APTP test scores, improve the current online registration process for test applicants, and provide new online screens for test administrators. Social security numbers will no longer be used to link test scores to student names; we will instead use a combination of demographic information collected online for students who preregister for a particular testing session, and on the answer sheet for students who do not. There was considerable discussion among working group members whether to include gender and ethnicity in the information collected to facilitate student identification and to permit analysis of test bias. The group decided that asking for this information might itself affect student test performance, and that examination of test bias could be undertaken instead by linking to demographic information from other data sources. We have modified the draft answer sheets reviewed by the group to create a revised [Pre-registered Answer Sheet](#) and [Walk-in Answer Sheet](#).

2. **Report on WI08 pilot test results.** Debbie McGhee, Research Scientist, OEA, summarized results of the instrument pilot test conducted in January 2008. This was followed by a discussion among the working group of the relative merits of three- versus four-alternative items. The instrument pilot test was a sequel to the items pretest conducted in autumn 2007.

AU07 items pretest. The autumn pretest assessed the characteristics of individual items and their relationship to a homogeneous content area. Items were combined into test books each of which assessed a single College Readiness Mathematics Standard. The purpose of the pretest was to identify items for inclusion in the new MPT-G and determine whether additional item-level tests needed to be carried out. An important pretest finding related to the number of functioning distractors for each item. APTP items historically have been written as multiple-choice items with five answer alternatives (one correct answer and four distractors) and this format was used in developing new items for the MPT-G. Pretest results clearly indicated that most items did not have four functioning distractors.

WI08 instrument pilot test. The purpose of the winter pilot test was to examine the functioning of whole tests made up of items from all Standards, and to provide information regarding the optimal number of distractors for APTP test items. Items that performed well during the autumn pretest were assembled into two parallel versions of the MPT-G. Each version was presented in both three- or four-alternative item format. [Results of the pilot test](#) indicated the two parallel versions tested were of approximately equal difficulty, versions made up of three-alternative items tended to be somewhat easier than those made up of four-alternative items, and all versions were of adequate internal consistency. (Cronbach's alpha ranged from .74 to .77 consistent with the heterogeneity of the Standards.) Low discrimination indices were found for four items (two on

each of the two parallel versions) and we used the opportunity of the group meeting to obtain feedback on possible improvements to those items.

Three- versus four-alternative items. Item distractors are the incorrect alternatives in a multiple-choice test item. Their purpose is to present a possible solution that reveals a misconception of the material and so distinguishes examinees who know the subject matter from those who don't. Non-functioning distractors were those incorrect alternatives that were chosen by very few examinees or that didn't distinguish between examinees who scored well on the total test and those who didn't. Based on these criteria, most items were found to have only two functioning distractors. There was considerable group discussion about whether the new MPT-G (and consequently the MPT-I and MPT-A) should consist of three- or four-alternative items. There are several arguments in favor of using the three alternative format: 1) It is very difficult to create four functioning distractors. Not only do existing MPT-I and MPT-A tests show only two functioning distractors, this was also true of the new MPT-G items despite our efforts to create four. Future item development would be facilitated by use of the three alternative format; 2) The testing literature supports use of three-alternative items as superior to four-alternative items; 3) The results of our own pilot test showed the two formats to be equal in terms of item reliability; 4) Examinees can respond to a larger number of three-alternative items during the same period of time than four-alternative items. This may allow us to increase the total number of items on the test and so increase its overall reliability; and 5) The test was somewhat difficult and use of three-alternative items would provide a little more "floor" for test scores and be more encouraging to students. The mean score for students in general math classes was less than 60% for the two three-alternative versions, and just over 50% for the two four-alternative versions.

The primary argument in favor of using four-alternatives is that students, parents, and educators are accustomed to this format. Additionally, although there are several advantages to using three-alternative items, the magnitude of those advantages is difficult to determine. However, after considerable discussion, the majority of the working group concurred that the MPT-G should be presented in three-alternative format. Because all APTP tests will be presented in the same format, this decision must be affirmed by the APTP governing council.

- 3. Implementation at K-12 and CTCs.** Nana Lowell summarized what is known so far about implementation of the college readiness test with high school students, and described some of the steps that have been taken to begin working to identify and address implementation issues at the CTCs. Bill Moore, Director of the Transition Math Project, has convened a separate group to investigate and discuss CTC implementation issues. Their initial conversations supported 'decoupling' placement at the CTCs from college readiness testing.

K-12 implementation. There has not yet been extensive discussion of implementation of the college readiness test for high school students. As noted above, the initial development plan called for revision of the MPT-I to provide college readiness testing in conjunction with testing for placement into college-level courses. An expanded version of the revised test would be created to provide diagnostic information for eleventh-grade students regarding specific areas of ability or deficiency relative to the college readiness standards. Although the deadline for implementation of the college readiness test is the same for both the CTCs and high school students (autumn 2009), the requirement for K-12 implementation is subject to legislative funding.

Preliminary conversations with OSPI have indicated that we are well positioned to meet the requirements of the bill if funding is awarded. At this point, it is thought that an expanded, diagnostic test might not be needed so the test used for high school could be the same as that developed for post-secondary students. The annual testing volume might be expected to be approximately 40,000, the number of high school juniors who self-identify as college bound. This testing could be integrated into the APTP testing year by establishing specific testing dates and locations in the same manner as currently used by APTP for the spring statewide testing (see the draft [Testing Site manual](#)). High school juniors would register for the test using the APTP online

registration system and make payment using vouchers issued by OSPI. Given the intensity of discussions about math education throughout the state, college readiness testing for high school students would be conducted in a fairly high-stakes environment. We have addressed this possibility by increasing the standardization and security of existing APTP procedures, and establishing an ongoing test renewal process that can generate one or more new versions of the test each year. If high school testing is implemented we also will convene a standing technical advisory committee following the model used by OSPI for WASL testing. Nana Lowell sat in on the most recent (March 14) meeting of OSPI's National Technical Advisory Committee and will attend future meetings as well.

CTC implementation. The focus of the project so far has been on issues around the test itself (test structure, item development, and pretesting). The project timeline now calls us to turn our attention to issues of implementation.

The functions of college readiness testing and course placement were initially combined to reduce the testing load on post-secondary students (they would take a single test rather than two) and to ensure that college readiness scores represent students' best work. While combining the two functions protects the validity of college readiness scores, it also adds a significant level of complexity in implementation due to differences in the way placement testing is carried out by two- and four-year schools. Many students at four-year schools initially enter during autumn term and APTP tests 6,000 fall entrants throughout the state each year in May and June. The remaining 4,000 test administrations take place in campus testing centers. Our current understanding is that the CTCs do not admit large groups during autumn quarter and consequently campus testing centers administer placement tests individually throughout the year. Preliminary indications are that many two-year schools utilize commercially available computer-adaptive tests that have considerable logistical advantages over the paper-pencil tests used by APTP: 1) a single test provides placement into the entire range of math courses from below college-level through calculus; 2) the same testing session also places students into other subject areas; and 3) the testing program computes and reports student scores immediately following placement. We plan to obtain more complete and accurate information about CTC placement testing by conducting telephone interviews with all CTC test administrators over the course of the next month (see draft [interview protocol](#)). We will ask about current procedures, satisfaction with current tests, and the feasibility of administering paper-pencil tests, noting that we expect to be able to deliver the test online (but not in adaptive format) within one year. We also will provide each test administrator with our draft [Testing Center manual](#) and ask whether their campus might be interested in offering the APTP tests starting in autumn 2008. Schools that implement the test one year early will provide valuable information on implementation logistics and in return we will assist them in setting placement cutoffs by carrying out statistical analysis of student scores and grades.

At the same time that OEA has been making these preparations, Bill Moore has convened a separate group of CTC representatives to investigate issues around implementation of college readiness testing at two-year schools. After several discussions, that group is considering a recommendation to "de-couple" college readiness and placement testing at the CTCs. These schools would use other tests than the MPT-G test for math placement, but would meet the requirement for college readiness testing by recognizing MPT-G scores for students who happen to have taken it. This proposal led to considerable discussion among working group members, in particular with reference to the impact on four-year schools from the lack of integration between placement decisions at two- and four-year institutions.

4. **Setting the college readiness cutoff (the "single bright line").** The question of how to set a single college readiness cutoff score is a difficult one and entered into the conversation several times throughout the day. Although there are recognized statistical methodologies to set placement cutoffs by relating test scores to subsequent course grades, the same is not true of setting a college readiness score. Jerry Gillmore, Director Emeritus, OEA outlined four possible

approaches: 1) empirical (compare test scores to some external criteria such as success in specific courses); 2) rational (specify a priori the percentage of students who should pass each item); 3) practical (consider the implications for course enrollment of different cutoffs); and 4) political. He went on to note that there will be a high degree of arbitrariness in setting the cutoff regardless of which approach or combination of approaches is chosen.

Although the discussion did not resolve how to set the “bright line” we did identify several questions that need to be addressed:

- Would students identified as college ready in the eleventh grade be required to take additional math coursework?
- Could students who take the college readiness test in the eleventh grade use the results for college placement?
- How long should placement scores be valid?
- What happens if a student is identified as college ready but does not place into any college-level course at a particular institution?

Attending:

Patrick Averbeck	Edmonds CC	Jane Lane	EWU
Linda Bolte	EWU	Bridget Lewis	Nine Mile Falls SD
Jim Brady	Spokane Falls CC	David Lippman	Pierce CC
Bradford Dallas	Nooksack Valley HS	Nana Lowell	OEA
Angela Davis-Unger	OEA	Debbie McGhee	OEA
Jerry Gillmore	OEA	Bill Moore	SBCTC
Gary Glaze	Spokane Falls CC	Bev Parnell	Yakima Valley CC
Vauhn Foster Grahler	TESC	Ginger Warfield	UW
Cinnamon Hillyard	UW-Bothell	Heidi Ypma	Whatcom CC
Russ Killingsworth	SPU		

Not attending:

Katy Absten	Olympic ESD	Jeanette Martin	WSU
Stuart Boersma	CWU	Stephen Mitchell	UW
Linda Brown	Spokane IEL	Jon Peterson	OEA
Linda Cave	WWU	Jan Ray	Seattle Central CC
Jackie Coomes	EWU	Bill Thelen	CWU
Boo Drury	OSPI	Dave Thielk	Central Kitsap SD
Tom Henderson	CWU	Tjalling Ympa	WWU
Michael Lundin	CWU		

Guests:

Cindy Morana	COP
Randy Spaulding	HECB