ETS® Major Field Tests

Guide to Score Interpretation

December 2022

Table of Contents

Section 1: Introduction

These guidelines for the Major Field Tests (MFT) are designed to help faculty, administrators, and staff get the most from their MFT administrations and the results they provide. This guide includes:

- appropriate and inappropriate uses of the MFT;
- information on the development, administration types, and appropriate usage of the MFT scores;
- best practices for assessing student learning outcomes using Major Field Tests;
- help in interpreting score reports; and
- tips for creating an assessment process that provides opportunities to obtain data that can be used for improving student learning.

Assessing student outcomes in higher education has received increasing emphasis and attention in recent years. Outcomes assessment can relate to measuring student progress in accredited major programs, assessing learning achievements at graduation, evaluating postgraduate activities, and identifying early career success. Determining the impact of higher education on these outcomes is of utmost importance to institutions, policy makers, and many other stakeholders. Doing so requires high quality information about what students know, think, and can do.

It was in response to inquiries into student learning — specifically, the need voiced by undergraduate institutions for valid, reliable measures of the outcomes of instruction in the disciplines — that Educational Testing Service® (ETS®) and the Graduate Record Examinations® (GRE®) Board developed the Major Field Tests in 1989. Development of the Major Field Tests was modeled on the development of the Graduate Record Examinations Subject Tests. However, unlike the GRE Subject Tests, the Major Field Tests do not serve as a predictor of graduate school success, but instead are designed to measure the basic knowledge and understanding achieved by undergraduates by the final year of their major field of study. Each test is revised approximately every 3–5 years. Experienced teaching faculty members, representing all relevant areas of a discipline, participate in determining test specifications, questions, and types of scores reported. ETS assessment experts subject each question to rigorous tests of sensitivity and reliability. In addition, every effort is made to include questions that assess the most common and most important topics and skills within each major field of study as supported by curriculum surveys sent to institutions in the United States.

Assessing Critical Knowledge with Major Field Tests

Major Field Tests (MFT) are designed to assist institutions and faculty members in assessing student learning outcomes. The MFTs are comprehensive undergraduate and MBA outcomes assessments designed to measure the critical knowledge and understanding obtained by students in a major field of study. Major Field Tests go beyond the measurement of factual knowledge by helping educators evaluate students' ability to analyze and solve problems, understand relationships, and interpret material from their major field of study.

Faculty members at colleges and universities use Major Field Tests to measure their students' academic achievement and growth, and to assess the educational outcomes of their programs. In addition, academic departments use Major Field Tests to evaluate their curricula and to measure the progress of their students. The tests also provide students with an opportunity to test their own knowledge within a field of study and compare their performance to students in their program, and to all other students taking the MFT.

Overview of Major Field Test Uses

When developing Major Field Tests, ETS test developers identified many appropriate ways in which institutions could use the assessments to demonstrate student learning. Major Field Tests can be used as stand-alone assessments in a specific major and/or in conjunction with an institution's internal assessments. Provided all applicable guidelines are adhered to, particularly the use of multiple sources of information when assessing student learning outcomes, scores from the MFT can also help faculty and other stakeholders.

- Document and describe student achievement of program-specific knowledge in terms of both benchmarking and trend analyses over time
- Provide evidence of improvements in student learning
- Identify performance levels of a specific group of interest
- Facilitate conversations and pinpoint areas of strength regarding pedagogy, curricula, and educational interventions
- Inform faculty development/training opportunities
- Integrate outcomes assessment data with other noncognitive and co-curricular indicators of student success (e.g., study habits, achievement motivation)
- Support accreditation and accountability initiatives by describing students' ability to meet institutional and program-level learning outcomes
- Provide cross-institutional and other comparative data (e.g., Major, Ethnicity, Transfer Student)

For many institutions using Major Field Tests, determining whether the assessment is appropriate for their program involves the following steps:

1. Determining the purpose(s) for which the program wanted to assess their students.
2. Reviewing the content description of the MFT and comparing it to the program's curriculum.
3. Obtaining a confidential review of a test form and reviewing it with faculty.
4. Soliciting feedback from students after the first administration of the MFT.

Following these steps can help to ensure that the Major Field Tests are an appropriate measure of the learning outcomes students are expected to gain as a result of your program and to help set program expectations accordingly.

Despite the rigorous MFT development process, which includes input from experts in their field, a test that may be quite appropriate for one department in a specific discipline may be inappropriate for another. Judging the appropriateness of a test for a particular department

or program is an important first step to take when considering the use of Major Field Tests. The appropriate faculty members and/or administrators should undertake a content review to determine whether the content and coverage of the tests are consistent with the content covered in your program and your program's expectations of students majoring in that field at your institution. The Test Description document on the Major Field Tests website shows the blueprint of the content covered and the approximate percentage of coverage. In addition, you may request a copy of the test for your review. These review copies are made available to you for review by you and your colleagues via email with a link to the test, which can be accessed for 30 days. You are responsible for the security and confidentiality of all test content.

To obtain a review copy of any of the tests, an institution must complete the Confidential Review Copy Request Form available on the Major Field Tests website under the Content tab. You may also request a copy of the form by calling an ETS Advisor at 1-800-745-0269.

The following guidelines summarize important considerations for appropriate use of the MFT assessments and provide examples of normally appropriate uses.

As an indicator of trends

Some institutions use Major Field Tests to look for year-to-year changes in the ability of their graduating seniors. Major Field Tests provide statistically-equated scaled scores. Scores that are scaled and equated ensure that minor changes in test content that may impact the difficulty of any test form are accounted for such that all test forms remain comparable over time. The MFT is typically administered to the graduating student because the test covers the breadth of the discipline. Therefore, in this scenario, institutions compare the performance of different cohorts of students over time. Another option for investigating trends includes investigating performance for both incoming students as well as graduating students over time, also known as pre-test and post-test design.

As a basis for comparisons with other institutions

Some institutions use Major Field Tests to determine how their students' skills and knowledge compare with the skills and knowledge of students at similar institutions. It is important that these comparisons involve students at the same point in their educational careers; for the MFT, typically students in their final year of study. The information about the performance of students at other institutions can come from either of two sources: the Comparative Data Guide, which is updated annually to ensure that it is representative of all types of institutions and programs using MFT; or the optional Custom Comparative Data report, which permits the user institution to select ten or more other institutions of interest. The scaled scores are particularly useful for this purpose, although the MFT Assessment Indicators can also be used, since either type of score would be converted to a percentile using data tables.

As a tool for critically examining teaching and learning practice

Some institutions use Major Field Tests data to determine not only demonstration of critical knowledge but also to identify areas in which students could use further reinforcement. Subscores and Assessment Indicators help identify areas where students performed well and those that may require improvement. Moreover, your program can use the MFT Item Information Report, which dissects your program's performance on the MFT on a question-by-question basis — an excellent tool for opening up a dialogue with faculty about the application of MFT results while promoting the least intrusive curricular changes possible. Use of the MFT in this manner allows for the development of an assessment process that focuses on continuous improvement.

As a tool for ensuring students graduate prepared

Some institutions test their students at a point in their academic career (such as at the end of the junior year) that would allow the program time to analyze areas in which a cohort of students may need improvement as identified by Major Field Tests. They can either advise students to seek out educational activities and resources to improve those skills or tailor their capstone experience at reinforcing some of these areas.

As an aid in recruitment

Some colleges administer Major Field Tests to identify those students whose skills and knowledge improved substantially. They may identify information in their students' educational background — information available at the time of admission — that distinguishes these students from those whose skills did not improve. This information can help the institution focus its recruitment efforts on the types of students who are likely to benefit from the institution's instructional program. Moreover, publishing your MFT success stories for your program — on your program website or in the local media — can attract more students and different kinds of students, further diversifying your applicant pool and show the value of your program to local employers.

Appropriate Uses of the Major Field Test

- Benchmark and Trend Analysis
    - Internal — Track your cohorts' performances over multiple years and compare students within your program
    - External — Compare your students' total score performance over time against the performance of other institutions
- Accreditation/accountability initiatives
    - Assurance of Learning — Demonstrate your students' ability to meet learning outcomes within your discipline
    - Continuous Improvement — Track and monitor the impact of curricular changes
- Cohort/subgroup analysis/comparison
    - Identify performance of a specific group of interest
    - Compare across different groups (e.g., Major, Gender, Ethnicity, Transfer Status, Academic Minor)
- Comparison to non-cognitive skills such study habits or achievement motivation
- Curriculum analysis/discussion
- Determine if expected learning outcomes are being met
- Effects of different course requirements on different groups of students

According to the Standards for Educational and Psychological Testing (AERA, APA, & NCME, 2014), test users (i.e., institutions) are responsible for collecting validity evidence for any uses of the test other than those recommended above.

Inappropriate Uses of Major Field Tests

Each Major Field Test is developed for use at the program level. Therefore, data is best interpreted in the aggregate, across students and courses within a program. Inferences made about individual mastery of overall discipline knowledge or sub-discipline knowledge should be avoided. We further emphasize that the Major Field Tests were not designed for making individual decisions about student knowledge within the major field of study. Specifically, the MFT was not developed to provide individual students with feedback on mastery of critical knowledge within the major.

Major Field Tests should never be used at the sole criteria for decisions at the individual student level. Test scores provided by the MFT should always be used in conjunction with other criteria when making decisions about an academic program or individual students. Departments and institutions are strongly cautioned against making the achievement of a certain score or percentile on the Major Field Tests a necessary condition for a student's graduation.

Institutions are encouraged to motivate their students to take the MFT seriously in order to produce results that are truly reflective of your students' abilities and, therefore, of your program's effectiveness. A combination of intrinsic and extrinsic methods may be necessary to adequately motivate your students, and some experimentation over time may be required to determine the combination of incentives that are most effective with your students. While it may seem that the ideal solution to motivating students is to simply have the MFT serve as a final exam in the capstone experience or to implement a cut score that students must attain in order to graduate, this solution creates other undesirable consequences. Students who must "pass" the MFT in order to graduate may respond to this pressure with any number of efforts that would invalidate the test results. At best, students would likely "cram" for the assessment, which would result in an inaccurate picture of the effectiveness of your program. (After all, an outcomes assessment is intended to measure what students learned over a course of study, not what they learned in the days leading up to the exam.) At worst, students may attempt testing improprieties such as collaboration among students during the test or between test administrations (if conducted over several days) or by bringing unauthorized test aids to the testing session (such as "crib notes"). Testing improprieties such as those described here would require ETS to cancel MFT scores for all students involved. Should these improprieties extend beyond your institution, it may require cancellation of MFT scores for all participating institutions. In order to maintain the security and the validity of the test for all participating institutions, ETS reserves the right to indefinitely prohibit the use of MFT at any institution encouraging students in these improprieties by establishing and enforcing cut scores.

Major Field Tests are designed for programs in the United States and are developed and validated based on samples of U.S. students whose primary language is English. For this reason, the MFT may be inappropriate for students who are not native speakers of English, for programs that are not taught in English, or regions other than the United States. To ensure the validity of the test for your students, please refer questions about your specific population to an ETS representative. Use of Major Field Tests outside of the United States is expressly prohibited without the written consent of ETS. Based on a 2012 study1, only the MFT–Business (undergraduate) has been validated for use outside the United States, and only for institutions with English as the primary language and a curriculum based on a U.S. curriculum.

Section 2: The Major Field Tests Outcomes Assessments

General Description of Major Field Tests

Major Field Tests were designed to assist higher education institutions and academic programs in assessing student knowledge within the academic major. Each MFT is a comprehensive outcomes assessment designed to measure the critical knowledge and skills commonly obtained by college students in their major program of study. Each test evaluates students' ability to analyze and solve problems, understand relationships, and interpret material typically seen within their major field of study. Each assessment goes through a rigorous test development process in order to provide reliable and valid data on student attainment of critical knowledge within the major.

Key features of the Major Field Tests include:

- Data that provides information about student learning
- Results that can inform improvements to educational programming within the majors.
- Results that can be used to meet the accountability demands for regional accreditation bodies such as the Southern Association of Colleges and Schools and discipline specific accreditors such as the Association to Advance Collegiate Schools of Business.

Major Field Tests Content

Each MFT assessment was developed by a panel of subject matter experts, based on core curriculum content identified in a curriculum survey. The content specifications for the MFT reflect the critical knowledge gained in the curriculum as supported by a survey sent to institutions across the U.S.

To ensure content currency, Major Field Tests are periodically reviewed to ensure that they remain valid, fair, and current to the discipline. MFT forms are "refreshed" periodically. New forms are released in September. The scores for new test forms can only be reported once a sufficient sample size has been received and the scores can be compared across different forms.

The curriculum survey, sent to all department chairs for a relevant subject across the U.S., reviews the test content to see if there are any noticeable shifts in distribution of content. If there is a significant shift in the distribution of a subfield, the scale will be broken, and the new form will no longer be comparable to the old form.

[1] Ling, G. (2013). Repurposing a business learning outcomes assessment to college students outside of the United States: validity and reliability evidence.(ETS Research Report No. RR-13-40) Princeton, NJ. Educational Testing Service.

In addition to factual knowledge, the tests evaluate students' abilities to:

- analyze and solve problems,
- understand relationships, and
- interpret graphs, diagrams, and charts based on materials from the specific field of study.

Note: All MFT items are designed to test analytical skills, which can be derived from knowledge of the concept.

The ETS website provides content description and sample items for all Major Field Tests on the Test Content Overview page. Each subject has a link that will take you to the specific subject test blueprint and sample questions. This information can be used to help determine the alignment between departmental or programmatic student learning outcomes and the corresponding MFT assessment instrument.

Logistical Characteristics of Major Field Tests

Major Field Tests are administered in a proctored environment using either the paper-and-pencil or online versions. MFT assessments are offered in over a dozen disciplines including:

Business	Science, Technology, Engineering and Mathematics (STEM)	Social Sciences	General
• Associate's Degree in Business • Bachelor's Degree in Business • Master of Business Administration	• Biology • Chemistry • Computer Science • Mathematics • Physics	• Criminal Justice • Economics • Political Science • Psychology • Sociology	• Music • Literature in English

Specific information about the topics covered in each major can be found in webpage links in the table above. Each MFT is broken into two equally timed sections. This is helpful for those institutions using class time for testing.

Table 1. Summary of logistical characteristics for the Major Field Tests

Major Field Tests	Administration Time (minutes)	Number of Items
Biology	120	150
Chemistry	120	100
Computer Science	120	66
Mathematics	120	50
Physics	120	70
Economics	120	90
Criminal Justice	120	150
Literature in English	120	150
Music	120	129
Political Science	120	130
Psychology	120	140
Sociology	120	140
Associate's Degree in Business	120	100
Bachelor's Degree in Business	120	120
Master of Business Administration	180	124

Section 3A: Technical Guidelines for Use of Major Field Tests Outcomes Assessment Scores

Overview

The following guidelines have been adopted by the Major Field Tests program staff to provide information about the appropriate use of MFT scores when assessing institutional and programmatic student learning outcomes. They are also intended to provide considerations for fair and best practices when using assessment scores. Adherence to the technical guidelines is important.

Policies and Guidelines for Appropriate Use of Major Field Tests Information

Although limitations and cautions apply to the use of any student learning outcomes measure, ETS staff have a particular obligation to inform users of the appropriate uses of scores from the Major Field Tests and to identify and rectify instances of misuse. ETS highly encourages score users to become knowledgeable about the validity of score uses and interpretations.

The following policies and guidelines are available to users of the Major Field Tests:

Score users. Higher education institutions and their students are considered score users. MFT program staff retain the right to make exceptions to this policy in special circumstances.

Validity. The general appropriateness of using scores from the Major Field Tests to measure critical knowledge and skills has been established by research studies carried out by ETS. Major Field Test scores may be appropriate for some other purposes, but it is important for institutions and programs to validate their use for those purposes. The ETS Global Higher Education Division staff can advise institutions on different processes and strategies for conducting validity studies and gathering necessary validity evidence.

Confidentiality. Major Field Test scores, whether those of an individual or aggregated for an institution, are confidential and can be accessed only by authorization of the individual or institution. All Major Field Tests data are confidential and are reported only to the administering institution. Score information aggregated across institutions is provided as comparative data; however, individually identifiable information is available only to the administering institution. Information about an institution gathered through the testing program will not be released in any form attributable to or identifiable with the institution unless ETS has received written authorization to do so. The confidentiality of student information and scores should be recognized and maintained. ETS suggests that institutions obtain a general written authorization from students stating that certain faculty members and others who are directly concerned with the students' education will have access to students' scores. ETS maintains the same degree of confidentiality of student data as it does institutional data.

Use of scores in aggregated form. Use of Major Field Test scores as a measure for ranking or rating undergraduate programs, institutions, university systems or states is

strongly discouraged except when the scores are used as one indicator among several appropriate indicators of educational quality.

Consequences as a result of student performance. Use of Major Field Tests scores to provide a grade for a student is strongly discouraged given that the assessment modules are not developed to assess course level material. Use of MFT scores for graduation requirements is also strongly discouraged.

Description of Scores from Major Field Tests

MFT scores provide institutions with the tools for making appropriate inferences about their students' demonstration of learning, and to make informed changes to educational interventions. To appropriately use data from Major Field Tests, stakeholders must understand the various types of scores made available to them, what these scores mean, and appropriate uses/interpretations of these scores.

Types of Data

The following describes the background information data, scaled scores, total scaled scores, subscores, and assessment indicator scores provided for the Major Field Tests.

Background Information. Background information is collected for each student for the purpose of gathering data in group form about students' backgrounds, academic preparation, and demographics. Answers to these questions do not affect a student's test scores, but responses are summarized and reported as part of the scoring services.

Note that score data (such as means and standard deviations) are not reported by these various demographic categories. In order to perform such analyses, you may create additional questions of interest to your program or institution and include these in the "Additional 50 questions" option or as subgroup questions. The data collected for demographic questions you create can then be used for score analysis by demographic subgroup. (See Subgroups below.)

- **Scaled Scores.** The Major Field Tests use scaled scores to summarize and report performance. A scaled score is a conversion of the numerical raw score achieved on the test (e.g., 50 of 80 test questions answered correctly) to a score in the predetermined range used for the MFTs. Scaled scores are used to ensure consistency in reporting and considering scores from tests that are composed of different numbers and types of test questions but display an overall similarity of use and purpose.

  As part of the equating process, the scaled scores on Major Field Tests are computed in a way that adjusts for the difficulty of the questions, to make them comparable to the corresponding scaled scores on previous editions of the same MFT titles.

  However, the different subscores and assessment indicators are not comparable to each other. For example, on an MFT Psychology, a scaled score of 50 in Clinical and Abnormal Psychology does not mean the same thing as a scaled score of 50 in Developmental Psychology.

- **Total Scaled Scores.** This score provides an indication on overall performance on the MFT. It is the estimated statistical representation of a student's critical knowledge in the discipline, as represented by the test content. Higher scores indicate an estimate of higher critical knowledge than lower scores indicate.

  Example of a correct interpretation of an individual student's total score:
  After a review of the tests and the curriculum, Student A's score on the MFT does/does not meet expectations.

  Example of an incorrect interpretation of an individual student's total score:
  Student B needs to remediate Course XX due to poor performance on the MFT.

- **Subscores.** These scores represent achievement in broad content areas within the field and reflect students' strengths or weaknesses by area within their major. The number of questions on the exam and the breadth of the subfield determine if a reliable subscore can be reported for an individual. Because subscores require approximately 30 test questions contributing to them in a specific subfield, subscores are not reported for all MFT titles. When subscores are available for an MFT title, those subscores are individually reliable for each student. Subscores are reported on a scale of 20–100. Note: These are not equated form to form, so institutions should use caution when including these in longitudinal studies.

  Example of a correct interpretation of an individual student's subscore:
  On this assessment, Student A has demonstrated more critical knowledge in the area of Cell Biology than Student B.

  Example of an incorrect interpretation of an individual student's subscore:
  Student A has mastered Cell Biology.

- **Assessment Indicator Scores.** These scores represent even more refined content areas than subscores. Assessment Indicators result from clustering test questions that pertain to a particular subfield within a major field of study. Assessment Indicators report the *average percent of correct answers* in a particular subfield of the discipline for all students tested, so you can determine if your students are having difficulty with specific clusters of questions. When using the assessment indicators, it is important to first review the questions within that subfield area of the test before determining what percent correct is most important for your students.

  From form to form, Assessment Indicators vary in the number of questions that contribute to their score. Because so few questions contribute to these scores, Assessment Indicators are not considered statistically reliable to report for individual students. Therefore, score reports for performance on the Assessment Indicators are reported only for groups of students.

Types of Score Reports

As part of the test purchase, institution users receive access to all reports without a subscription fee inside the Program Workshop portal. Examples of these are described below. Students who do not complete 50 percent of the assessment are excluded from aggregated data in score reports and comparative data.

Certain reports allow users to aggregate and disaggregate data at many levels, depending on their research needs.

The need for these reporting options depends on the size and makeup of the sample of students at each institution. ETS advises users that, while disaggregation is helpful in analyzing your data, extremely small samples of students cannot yield meaningful data. However, with the online portal, users can combine cohorts to ensure they meet reporting minimums. For all but the Item Information Report, MFT reports require a minimum of 5 students. The Item Information Report requires a minimum of 20.

Individual Score Report. This report provides the student's overall score and subscore (where applicable), the percentile rank of the individual's score and the average score range. The percentile rank indicates the percentage of test takers whose performance is lower than the reported score. The average score range illustrates the middle 50 percent of the test takers. The scores provide an estimation of the amount of critical knowledge in the domain that has been demonstrated by the student. Higher scores indicate the demonstration of more critical knowledge.

Students and faculty members can use this report to determine the relative position of this student's critical knowledge among his peers at his/her institution or among all test takers.

Departmental Summary of Total Test and Subscores Report. This report provides the range of possible total scores and subscores for the MFT assessments. The number of your program's students within each range of scores is provided, as well as the percent of students within your program scoring below each range. This information provides the user with a frequency distribution of scores for students in their program.

Faculty members can use this report to obtain an idea of the range in critical knowledge among the students in their program. The institution can use this report to help determine if this range in student scores is expected given the program's curriculum and student profiles.

Departmental Summary of Assessment Indicators Report. This report shows the average of all correct answers in each category from all students taking the test. Assessment Indicators can be used effectively for curriculum improvement. Performance in each Assessment Indicator subdomain can be compared to the Test Description on the Major Field Tests website to begin to determine specific concepts that students are expected to have mastered in each subdomain.

The percent correct scores for the Assessment Indicators cannot be meaningfully compared to each other across sub-discipline areas. Using an example from biology, a percent correct score of 20 in Biochemistry and Cell Energetics does not mean the same thing as a percent correct score of 20 in Molecular Biology and Molecular Genetics. However, comparative data can be used to show relative performance in each subdomain area; an Assessment Indicator can be converted to a percentile and compared to other programs using the same MFT test title.

This report is best used when faculty members have had an opportunity to review the test form prior to administering it to students. Doing so allows for the opportunity to estimate the expected performance of their students in each of these areas. The results shown in this report can then be used as an indication of whether or not students are meeting expectations.

Departmental Demographic Summary. This report provides the user with a demographic summary of the students tested within the program. The report provides both counts and percentages of students by demographic category.

This report can be used to show characteristics of the students in your program. It may allow you to identify subgroups of students whose scores you'd like to analyze, or assist you in evaluating if your population and scores have changed dramatically over time.

Comparative Data Reports. Updated annually, the Comparative Data report is available for every operational MFT test form. It provides you with the relative position of your student or institution against the entire test-taking population for that specific test form (or version). When a new form of a Major Field Tests title is introduced, there is a delay between publishing the test and publishing comparative data for that test. Comparative data for the new test form cannot be compiled until a representative number of programs have administered that new test.

Even though a new test form is statistically equated to the previous test form at the Total Score level, caution should be used when making judgments about performance on the new test form using comparative data from the old test form because of the potential for minor differences in the population using each test form.

This report is best used to identify your program's rank among other institutions that are using the Major Field Tests. Comparative data allows institutions to evaluate their program's performance on the ETS® Major Field Tests relative to the larger group of test takers at other institutions. For the Major Field Tests, the comparative data for each test is based on the scores of all senior-level students who took the most recent form of a test and who are from institutions where at least five students were tested. The data is comparative rather than normative, since the sample of institutions and students does not represent all possible types of institutions and departmental curricula. This means that your scores are interpretable with reference to other institutions that selected to use the MFT and are not reflective of where you stand against all students at all institutions in the U.S.

Some institutions use this data to encourage students to outperform previous cohorts of students at the same institution.

Subgroup Reports. When looking to analyze subgroups of your population, there are two approaches you can use to report the data. The first is the Subgroup Report. This allows you to ask two (2) questions each with eight (8) answer options to filter the standard reports by each subgroup. Alternatively, cohorts can be set up and managed for your subgroups and either individual reports or a combined report for overall review can be run. These reporting options prove very beneficial when trying to narrow down your analysis. It provides all roster and summary reports listed above for each subgroup. It is important to have enough valid results in each subgroup to make it meaningful.

Faculty members can use this report to determine the performance of various subgroups of students. When there are subgroups of students whose performances are expected to vary, this report can help with exploring those differences even further.

Custom Comparative Data Reports. An institution can select ten or more other institutions from the user list in order to create a reference group of meaning to them. Examples of reference groups of interest could include: programs in the same state or accreditation region; programs accredited by the same discipline-specific accreditation

body; programs of similar caliber, academic reputation, or against whom they compete for students; or Minority Serving Institutions. The report will include all of the same data tables in the same formats as the published comparative data tables for that subject.

Faculty members use this report to determine their relative performance among peer institutions. This report is best used when the institutions included in the report conduct fairly similar assessment processes.

Item Information Reports. An institution can receive descriptive item-level data for a particular cohort in order to further pinpoint strengths and weaknesses in the curriculum. Item-level data includes the percent correct of your institution, the percent correct of the national comparative population, the percent omitted/percent not reached, and the mapping of which items contribute to which subscores/assessment indicators.

This report is best used in conjunction with a copy of the test. Faculty members can use this report to help gauge expectation of student performance on each test question of the MFT. Where students are not meeting expectations, programs can further investigate the impact of their curriculum on student critical knowledge of the discipline. This report is available in MS Excel to allow deeper analysis, such as creating a column for a gap between your institution and the nation, highlighting cells above/below a certain threshold for deeper review, and filtering for content areas to focus on.

Design Your Own Analysis Reports. This report is a data file that provides all administration, biographic, and student performance information. This data file is provided to allow institutions the flexibility of conducting their own analyses on student data. For added flexibility and potential integration to various analytic tools, this report is available in MS Excel. This allows institutions to integrate student information from their databases such as courses taken, SAT/ACT, and GPA.

Faculty members use this report to conduct in-depth analyses of student data. This report provides faculty the opportunity to manage their own data files and is often used by programs who conduct research on student learning within their discipline.

Score Reporting

Users of paper-and-pencil MFT can access their scores via their Major Field Tests online account (Program Workshop). Scores for reports will be available within 10 business days of receipt of a school's completed answer sheets.

For the online MFT format, an administrator at the institution must let the system know that the administration is over (also known as "closing the cohort"), after which all reports are immediately available through the Major Field Tests online account.

The only exceptions, to both the paper-and-pencil and online test schedules, occur when ETS publishes a new exam or when an exam form changes.

Releasing a New Form — Score Reporting Timetable

The availability of scores for a new test form depends on the time it takes to gather a sufficient quantity of data. The data from new forms must be analyzed to ensure comparability to previous test forms. Attaining a representative sample for a new test form may take a few months. Once this sample is available, the equating process can begin, and scores can be determined. During this wait, the Major Field Tests program provides a

preliminary score range for your students taking the test. This roster of ranges can be run directly out of the Program Workshop portal. These ranges are based on prior versions of the form. You may use the current MFT Comparative Data Guide to evaluate total scaled scores from the new test form in order to determine how your students did on the new form as compared to the population that took the previous form. This comparison can be made because the total scaled scores are equated from form to form, and the new and old test forms are comparable.

Description of Test Formats

ETS Major Field Tests are offered with many flexible options to support your academic programs. There are no preset test administration dates, enabling you to schedule testing at the convenience of your faculty and students. Institutions can choose between proctored paper-and-pencil and online testing on-campus, and off-campus with a remote proctor.

The Major Field Tests are multiple-choice exams that require two hours to administer (three hours for the MBA test).[2] Both the paper-and-pencil and online versions must be administered in a secure, proctored environment.

Departments or institutions choose when and where to give these tests; however, the tests are normally administered during the students' final year when they have completed the majority of courses in the major. Many institutions administer the tests as part of the requirements of a capstone course. Other institutions employ a pre-test and post-test model, testing students upon declaring their major and then again in their capstone course to measure change in performance.

Note: Our Assessment experts are available to assist your institution in developing a pre-test/post-test design.

Refer to the following table for basic format and features of the Major Field Tests.

Table 2. Major Field Tests Administration Features

Delivery methods	• Paper-and-pencil • Online
Administration methods	• Proctored on-campus administration for both paper-and-pencil and online versions • Proctored off-campus (remote proctored) administration for online versions
Payment methods	• Institution-pay • Student-pay
Test lengths	• Two hours — Associate and Bachelor's degree programs • Three hours — MBA program

[2] Up to five additional minutes may be necessary to permit students to complete a Biographical Information Questionnaire.

Minimum number of test takers to produce aggregate scores	• Five
Integrated with Learning Management Systems	• Integrates with any IMS LTI-based Learning Management Systems.

To ensure access to free technical support for the online version of the Major Field Tests, it is recommended that testing be scheduled during Technical Support's business hours: Monday-Friday, 8 a.m.–6 p.m. Eastern Time. You can contact Technical Support at **1-800-514-8491**.

Statistical Information for Major Field Tests Scores

The reliability of the MFT Total Scores are shown in Table 3 in the Appendix.

Reliability Coefficients and Standard Errors of Measurement

Ensuring the reliability of test scores is a major component of the development process for Major Field Tests. The reliability of MFT scores allows academic programs to have confidence that student performance is an accurate and precise indicator of their knowledge. Best practices in academic program assessment suggest that programs should have a good understanding of the reliability of the scores they intend to interpret and use.

Reliability is an indicator of the precision or consistency of test scores. It refers to the extent to which scores obtained on a specific form of an assessment, administered under one particular set of conditions, can be generalized to scores obtained on other forms of the assessment or administered under other conditions. Reliability can also be viewed as an indicator of the extent to which differences in test scores reflect true differences in the knowledge or ability being tested, rather than random variation caused by such factors as the form of an assessment, the time of administration, or the scoring method.

The theoretical concepts of "truth" and "error" are important in the study of reliability. Theoretically, if a student takes an infinite number of equivalent editions of a test, the scores obtained would vary but would cluster around the student's true score. The true score would be the average score over the infinite number of replications. A true score is not necessarily a correct score. A consistent or systematic error that occurs in every sample would be counted as truth. Error can be thought of as the random fluctuations from sample to sample, or the differences between a person's true score and the obtained test scores (the sample scores). When a test is administered to a group, the variance in the distribution of scores is due partly to real differences in ability and partly to random errors (error variance). Reliability is the proportion of total variance attributed to true variance.

The reliability estimate for the MC scores is computed using the Kuder and Richardson (1937) formula 20 (KR-20), as shown below. For the total test, a desired level of reliability is .90 or higher. The reliability of a test is directly related to the number of items contributing to the final scores.

$$r_{X} = \frac{n}{n-1}\left[1-\frac{\sum_{i=1}^{n} p_i q_i}{\sigma_X^2}\right]$$

where

n = number of items in the test,

p_i = proportion of correct responses to item i,

q_i = 1- p_i , and

σ_X^2 = variance of total scores on the test.

No test can be perfectly reliable because a test is only a sample from a larger population of possible questions and possible times and conditions of administration. There are differences between a true score and an obtained score. These differences are called errors of measurement. A way of characterizing the typical amount of error is called standard error of measurement (SEM). The SEM is a statistic that indicates the standard deviation of the differences between observed scores and their corresponding true scores. The SEM of a raw score is computed from the reliability estimate and the standard deviation (SDx) of the scores by the formula:

$$SEM_X = SD_X \sqrt{1 - r_X}$$

Due to the requirement of the MFT score reporting timeline, a convenient sample — data from the first students to test on that form — was used for equating and reliability analyses. Examinees who responded to less than 50% of the items and whose best language is not English were excluded from the analyses. Furthermore, the reliability coefficients and standard errors of measurement (SEM) were calculated using data for seniors only.

Table 6 gives the mean scaled scores, reliability and SEM for the total score as well as each subscore for 15 subject tests. For each test, a minimum of 200 students in a subgroup is required for these statistics to be reported.

Description of Score Equating for Major Field Tests

Test score equating is the process of statistically adjusting scores on subsequent tests forms to account for minor changes in test difficulty over time. This allows different forms of the same test to remain comparable over time. The Major Field Tests equating process uses common questions — or an anchor block — to form a linkage between the previous test form and the new test form. This statistical process allows the total score from different versions of the same test to be compared. Common-item equating is one of the most accurate methods of ensuring comparability of scores. The accuracy of equating is dependent on the number of questions contributing to the score.

For this reason, only the MFT Total Score has enough test questions contributing to it to create this anchor block. Subscores within each Major Field Test are not on the same scale as the total score and should not be compared across forms. Assessment Indicators — which represent further granularity on a subject but are comprised of fewer items — cannot be statistically equated across versions of the test. Therefore, Subscores and Assessment Indicators are not fully comparable from one version of the test to another.

Test Administration Guides and Documents

- Test administration procedures and other information can be found in the manuals available inside the Program Workshop portal. Types of material available include:
- Test Administration Manual – provides step-by-step instructions for paper-and-pencil test administrations
- Processing Request Form – should be included with completed tests when returning them for scoring. Use a separate form for each subject area.

- Test Administration Report Form (within Test Administration Manual) – used to report procedural irregularities during testing
- Supervisor's Questionnaire and Comment Form (within Test Administration Manual) – used to provide comments on the test for future improvements
- Proctor scripts for online and paper versions – provides step-by-step instructions for proctoring online and paper test administrations
- Quick Start Guides and Videos

Adding Test Questions

For institutions that have program content that is not included on the Major Field Test or have biographical information questions they would like to include for targeted reporting, Major Field Tests offer the ability to add up to 50 locally authored questions to the end of either the paper-and-pencil or online on-campus MFT. These questions can serve a variety of useful data collection functions:

- Additional questions about a subdiscipline within the major that may already be covered by the MFT but that your program wishes to assess in greater depth
- Additional questions about a subdiscipline within the major that is not covered by the MFT but that is a learning objective for your program
- Additional questions about a discipline tangentially related to the discipline covered by the MFT
- Additional questions to assess core mission and learning goals for your institution or program that are not in any way related to the major covered by the MFT
- Survey questions about student characteristics or other non-cognitive attributes

Examples might include:

- Adding different types of Accounting questions to the MFT Business test
- Adding questions about a specific programming language to the MFT Computer Science test
- Adding Biology questions to the MFT Chemistry test (or adding Chemistry questions to the MFT Biology test) to accommodate biochemistry programs
- Adding questions in non-cognitive areas such as social responsibility, cultural awareness, ethics, or teamwork

Institutions receive a score report at no extra cost that includes a summary of student responses to each question. The 50 additional questions are reported in aggregate, as frequencies of students selecting each of the response options for each test question. This data can then be used to calculate percent correct scores for the additional questions.

There is no need to inform ETS about the intended scoring key for these additional questions since data is generated about every answer choice. The exact response selection is available to institutions with the Design Your Own Analysis report.

Please note that a score for an individual student cannot be generated from these questions, as the validity and fairness of questions authored by you and your colleagues

cannot be confirmed by ETS. Moreover, these additional questions will not contribute to your students' overall Major Field Tests scores; such customization would render MFT scores incomparable across programs and institutions. Optional questions written by your institution for your institution are not comparable to questions written by other institutions. Therefore, ETS cannot report comparative data about these locally authored questions.

The addition of locally authored questions may lengthen test time. The institution should make the determination about how much additional test time students should receive to complete these questions in the same manner they would for other assessments they design themselves.

Testing Accommodations

Students with Disabilities

Make arrangements for administering the ETS Major Field Tests to students with disabilities in the same way you would make arrangements for administering other important local tests.

For students who have difficulty filling in the bubbles on answer sheets, we suggest that students who are able to do so circle their answers in the test book. After the testing session, the test administrator must transfer the answers to an answer sheet to be processed by ETS. The student's test book should be attached to the answer sheet.

The Major Field Tests online system offers a zoom capability to magnify the screen display and enlarge fonts, and to reverse foreground and background colors (changing black-on-white to white-on-black and vice versa) when desired. It should be noted that when using magnification or enlarged fonts, excessive scrolling is required, which could place the student at a time disadvantage.

For those students not testing online, the Major Field Tests are also available in braille and large print, and on audio. We recommend that requests for tests in a large-print format be made six weeks in advance, and that requests for braille or audio cassette format be made as far in advance as possible, preferably at least three months ahead.

Results of the paper-and-pencil, large-print or braille exam can be combined with results of the online tests for the rest of that institution's population.

Students Requiring Additional Time

A student should only be granted additional testing time for a Major Field Test if the same student receives similar accommodations for other exams and has previously demonstrated and documented the need for the accommodation.

For students requiring additional time, time may be granted at the discretion of the test administration at the institution. Test administrators can configure the MFT online system to permit additional testing time for a student. The student's countdown clock will be adjusted accordingly. Please follow the instructions outlined in the Test Administration Manual.

Section 3B: Important Validity and Statistical Considerations

Overview

The following validity and statistical considerations have been adopted by the Major Field Test program staff as part of the Technical Guidelines for use of the MFT outcomes assessments. These considerations are informed by best practices in educational and psychological testing as well as best practices in higher education learning outcomes assessment. Making note of these considerations could be beneficial to your process for assessing student learning.

Validity Considerations

Use Multiple Sources of Information

A best practice when engaging in student learning outcomes assessment is the use of multiple sources of information about student learning when making decisions. Triangulating multiple data points allows for the creation of a more holistic picture of students' knowledge, skills, abilities, and development.

Moreover, broadening your assessment plan to include key skill areas by incorporating one of the HEIghten Outcomes Assessment modules (Critical Thinking, Written Communication, Quantitative Literacy, Civic Competency & Engagement, and Intercultural Competency & Diversity) can further inform your program assessment. HEIghten Outcomes Assessment is available from Territorium with official content from ETS. Scores from any one of these assessments of core skills can provide further context to student performance on the Major Field Tests. For instance, institutions may consider using a direct assessment of critical knowledge within a discipline in conjunction with direct measure of critical thinking. This would allow programs to explore and describe relationships between specific content knowledge and core skills.

The use of multiple sources also includes the consideration of the various aspects of the curriculum or co-curricular activities that are expected to improve student learning. These sources may include the types of courses students have taken, credits earned, transfer status, major, prior academic achievement, and honors or developmental activities that may impact student learning.

Conduct Validity Studies

Institutions and programs using Major Field Tests to assess student learning for purposes of program improvement or for accountability are encouraged to assess the validity of the inferences they are marking from learning outcomes data. An example may be to assess incoming first-year students' content knowledge before they experience any major or program courses and again years later, after they have completed most or all of their major or program coursework. For instance, if your biology program student's learning outcomes align with the Biology MFT assessment, it would be reasonable to expect a positive relationship between assessment performance and biology course grades. Students who have completed more biology courses should, in theory, perform better on the Biology MFT assessment compared to students who have completed none or fewer of their biology program classes.

Institutions using a locally developed biology instrument, in addition to the Major Field Test, could conduct convergent validity studies (i.e., evaluating the relationship between two

measures of biology content knowledge). Students' scores on the locally developed biology instrument should be positively related to their scores on the Biology MFT assessment; however, stakeholders would not expect the relationship to be perfectly strong (i.e., a perfect correlation) because the two assessments are likely defining and measuring biology skills somewhat differently.

Various types of validity studies may help determine the alignment of each Major Field Test to an institution's learning outcomes. These validity studies can help aid in the interpretation of student performance on the Major Field Test. ETS Global Higher Education Division staff can advise institutions on different processes and strategies for conducting validity studies and gathering necessary validity evidence.

Conduct Reviews of Major Field Tests Content

Although the development of each Major Field Test was informed by committees of assessment developers and content experts from higher education institutions, the match between each assessment and the curriculum at any given intuition may not be exact and may vary over time. Institutions are encouraged to periodically review the assessment instrument content in order to verify the appropriateness of the content for their programs.

Statistical Considerations

Avoid Decisions Based on Small Score Differences

Small differences in MFT scores (as defined by the standard error of measurement [SEM] for score differences) should not be used to make distinctions among assessed students. SEMs vary for each MFT assessment and are available in the Appendix.

Avoid Decisions Based on Small Samples of Students

The size and representativeness of the students participating in your assessment processes will impact the inferences that can be made from the Major Field Tests. If assessment of all students is not an option, institutions and programs should ensure that their sample size is large enough to be a stable reflection of student performance, and representative of their student population.

Institutions are also encouraged to report the demographic information for their assessed students in addition to the overall demographic information for their entire student body. This allows for your internal and external stakeholders to ascertain any important demographic differences and/or similarities between the assessed students and the larger, overall student population.

Use the Appropriate Percentile Ranks when Comparing Students

Percentile ranks are provided on score reports and can be used to compare institutions' and students' relative performances among the measures. Percentile ranks indicate the percent of assessed students or institutions in a group who obtained scores below a specified score. For instance, students scoring at the 98th percentile rank scored better than 98 percent of the other students who took the assessment. Percentile ranks should be compared only if they are based on the same reference population. When Percentile Ranks are updated, they will be available in Program Workshop, the portal used to access your ETS assessments.

Important Limitations for Major Field Tests Scores

A test contains only a sample of the tasks that students are expected to be able to do. On a different sample of tasks designed to measure the same skills, the same students might perform somewhat differently. Information provided on the score reports enables the user of the scores to determine how much the scores could be expected to differ if a different set of tasks were used. For this reason, institutions and programs should consider multiple sources of information.

The Major Field Tests measure a specific collection of knowledge. When the Major Field Test is used to evaluate student learning in an academic program, it should be used in conjunction with other information. It should never be used as the sole means for evaluating the effectiveness of a program or the educational progress of the students.

Any Major Field Test has two primary limitations:

- It does not and cannot measure all the knowledge and skills students attain at an institution, and
- It is an inexact measure; consequently, the scores should not be used to make high-stakes decisions regarding individual students. Inferences about student knowledge and skills should be focused at the group level.

The methods used to standardize each MFT assessment allows for possibility of comparing student performance within and across institutions. The Custom Comparative Data Guide allows you to choose the group of institutions to which you'd like to be compared. Within any category of institutions, those that use MFT assessments are not likely to be representative of all institutions in that category. This is due to the fact that not all institutions choose to use the MFT as their outcomes assessment. Also, the numbers of students assessed and the sampling procedures vary from one institution to another, and it is impossible to verify that the students tested at each institution are representative of all the students within a program at the relevant class level.

Section 4: Practical Guidelines for the Use of Major Field Tests Scores for Demonstrable Student Learning Improvement

Overview

The Major Field Tests program provides institutions with tests that have established fairness, reliability, and validity for their use in learning outcomes assessment. The quality of these assessments allows them to be used to meet accountability demands and help inform curricular changes for the purpose of improving student learning.

However, the quality of the data obtained from these assessments is not only impacted by ETS's ability to build fair, valid, and reliable assessments, but is also impacted by the assessment processes instituted by your program.

The following assessment processes and practices are provided to assist institutional stakeholders with implementing best practices when assessing student learning at the program level.

In this section "Tips" accompany each practical consideration. These tips offer actionable advice to institutions engaged in the process of learning outcomes assessment. To that end, these tips will help institutions:

- understand the purpose of your assessment,
- engage faculty members in assessment processes,
- improve the quality of the results obtained from learning outcomes assessment,
- modify pedagogies and curricula, and
- demonstrate student learning improvements related to MFT assessment results, among other outcomes.

ETS hopes that institutional stakeholders can use these practical guidelines, in conjunction with the Major Field Tests, as thought leadership to promote innovation, facilitate conversations, support educators, and empirically demonstrate learning improvements on campus.

Practical Considerations

Focus on the program, not the individual courses.

The Major Field Tests were developed to gain a broad understanding of student content knowledge within their academic major. These assessments were not developed to measure the impact of curriculum of any specific course or program. By assessing this broad content knowledge users of these assessments recognize that their student performance can be compared to students that experience different curricula. This framing of assessment acknowledges that students are not only competing for opportunities in graduate education or workforce with peers in their institutions, but with peers from across the nation. By focusing on program level assessment, we are able to leverage information about student learning that goes beyond the information already gained by faculty members in their

course assessments. Using the MFT allows institutions to contextualize their programs impact on student learning beyond course level assessments.

TIP: Review the assessment to note any misalignment between what is taught in your courses and what is being assessed by the test. Set appropriate expectations based on your understanding of the curriculum, the students you serve and the content of the test. Your expectations may differ across the various content areas based on your programs' curriculum.

TIP: Where there are areas of misalignment, take advantage of the ability to add your own test questions to the MFT.

TIP: Assessing only the content covered in your program may not provide the full picture of the value of your program to your students and their lifelong learning or professional goals.

Engage faculty members and students in the assessment process.

Assessment works best when everyone in the teaching and learning process understands that assessment data is being used to modify pedagogies and curricula for improving student learning. When faculty members are able to answer the questions they have about student learning with assessment data, that data is more likely to be used to improve pedagogies and curricula. Likewise, students may have more incentive to engage in the assessment process when they know that their data will be used to meaningfully inform improvements at the program or university level.

TIP: Consider what questions about student learning you want to answer at the outset. Then focus improvement efforts on one or two targeted learning outcomes that faculty are interested in examining and improving.

TIP: Share a brief summary of results with faculty and students; discuss if and/or how faculty and students would find this useful for informing pedagogical and curricular changes and enhancing student learning.

TIP: Survey faculty members to see what kind of educational activities they would add if students were not performing at expectations on the learning outcomes as measured by the MFT assessments. "Crowd-source" potential actions to be taken to improve student learning.

TIP: If faculty buy-in is a concern, start with a dedicated "core nucleus" of faculty who want to work collaboratively with colleagues, assessment practitioners, faculty developers, students, etc., to improve learning.

TIP: Disseminate assessment data in a variety of easily accessible and understandable formats to both students and faculty.

Frame assessment processes as research and scholarship opportunities.

To promote learning improvement, it is essential to align assessments with student learning outcomes, pedagogy, and curricula. In doing so, institutions can begin to frame assessment processes as research opportunities. Often, faculty have research questions regarding the effectiveness of teaching and learning practices that can be addressed through assessment

methodologies. That is, assessment processes can be used to answer important research questions related to student learning.

TIP: Encourage faculty to use assessment processes to document and share important findings related to student learning at disciplinary conferences, in peer-reviewed publications, etc.

TIP: Most faculty are intrinsically interested in teaching and learning. Leverage this interest by framing assessment in terms of research questions faculty want to answer about student learning.

TIP: Assessment can be framed as research that is driven by previous results and/or informed by research from applicable literature (including focus groups, faculty experiences, interview, anecdotal evidence, etc.)

TIP: Focus on one or two key learning outcomes that are of interest to your faculty members. Assessment can become resource intensive when many assessment initiatives are occurring simultaneously.

Provide adequate support systems and recognitions for faculty.

Using assessment results to influence and support continuous learning improvement requires adequate support from administrative stakeholders, as well as expertise from assessment professionals and faculty developers. ETS offers webinars, resources, and support for learning improvement using the Major Field Tests.

TIP: Consider recognizing faculty members or programs/units within your institution for contributing to assessment initiatives that lead to improved student learning.

TIP: Opportunities for support or recognition may include (but are not limited to): Provost's or President's award for assessment practices, involvement in assessment and learning improvement activities included as part of applicable tenure or promotion experiences, stipends for time spent promoting and facilitating learning improvement projects, supporting faculty attendance/participation in summer institutes or workshops related to learning improvement, incorporating learning improvement into pre-existing assessment/reporting processes like program reviews, etc.

Ensure Major Field Tests meet purpose(s) for assessing students.

When selecting your learning outcomes assessment, make sure the assessment meets the program's purpose(s) for assessing its students. Particular attention should be paid to the curricula or co-curricular programs that will impact the skills measured by the Major Field Tests. By focusing on the extent to which the assessment meets your purposes, you can set goals appropriately and supplement the assessments, where needed. In addition, ETS advises you to determine when the learning outcomes are addressed in the curriculum; this information will help with score interpretation and learning improvement.

TIP: Review the following content and item information for your MFT assessment. Also review your curriculum map and substantive co-curricular program or initiatives.

TIP: Request a Confidential Review Copy of the module and review as a department.

TIP: Map assessment module characteristics and/or content back to your student learning outcomes and/or curricula to ensure modules are aligned with educational experiences at your institutions and are meeting your intended assessment purposes.

Examine the current pedagogies and curricula.

MFT assessment results can be used to influence and inform intentional changes to pedagogies and curricula at your institution. Having intimate knowledge of your curriculum is imperative to do so successfully. To make intentional and informed changes, you should understand the pre-existing pedagogies and curricula that were in place (e.g., how would you describe the educational experiences students had that likely contributed to their performance on the assessment modules).

TIP: Collaborate with faculty who teach in major or program courses to determine/ outline where in the curriculum students are currently learning specific knowledge, skills, and abilities. Does this differ from where faculty feel students should be learning specific knowledge, skills, and abilities?

TIP: Consult with faculty who teach in major or program courses to determine how various specific knowledge, skills, and abilities are being taught, what teaching strategies are being used, what additional pedagogical strategies faculty may want to pilot, and how might you modify current pedagogies or curricula or create new ones to give students better/more opportunities to learn specific knowledge, skills, and abilities.

TIP: If faculty want to pilot new pedagogical strategies or re-design courses, they will most likely need support from administration and appropriate faculty development/ training opportunities, in addition to adequate time.

TIP: It's important to consider co-curricular activities and opportunities to learn when examining the current pedagogies and curricula.

Use Major Field Tests scores as part of a customized assessment plan.

MFT program staff continue to work with partnering institutions of higher education to implement best practice tips for creating customized assessment plans — for both accountability and improvement — that are applicable to all colleges. Institutions and programs can use the Major Field Tests as part of a custom assessment solution that best fits their needs. This may involve the need for locally developed assessments in addition to the Major Field Tests, which can serve a complementary role to the locally developed assessments.

TIP: Institutions and programs should take advantage of the transparency ETS has provided in the development of the Major Field Tests to ensure that institutions can identify potential areas for augmentation.

TIP: Each assessment module allows users to append their own assessment tools. This process allows you to manage and combine multiple data points from your customized assessment plan to meet your institutional or programmatic needs.

Make sure the group of students to be assessed will provide the necessary information.

If the purpose of assessing student learning is to make inferences about the performance of groups of students, it is important to assess an adequate number of students from your total population or from the subgroups you've identified. These students should be selected in such a way that the students assessed from each group are representative of the group or groups you wish to make inferences about. The best way to accomplish this is to assess all students. However, that is often the least efficient. Nevertheless, it is particularly

important not to limit the assessment to students who volunteer to be assessed, unless the program wants information that applies only to those students. Programs should think carefully about the demographic information the students provide as well. This is the information that will be used to split students into different subgroups for comparison. Obtaining more demographic information allows for more comparisons to be made.

TIP: When assessing the quality of your data, begin by comparing your institution's or program's demographics to the demographics of the group of students tested. This allows you to place assessment results into appropriate contexts.

Make sure students are motivated to take each assessment and motivated to do well.

Student motivation in learning outcomes assessment is a serious concern. If the students are not motivated to do well on the assessment, their scores will not reflect their actual skill levels. Therefore, the students' scores will not accurately reflect the impact of educational interventions (e.g., pedagogies, curricula) on their learning. In addition, conclusions or inferences drawn from students' scores may be inaccurate.

Although the Major Field Tests do not include a measure of student motivation, several psychometrically sound instruments exist for examining students' self-reported levels of motivation. For instance, if students self-reported giving low effort on the modules, institutions could potentially have justification for excluding their data from subsequent analyses.

Finding the most appropriate motivation technique is a matter of finding the incentive that speaks to your students and is best aligned with your institutional culture. Best practices indicate that often a combination of incentives may be required to encourage students to do their best on assessments. The most effective combinations of motivational incentives strive to achieve a delicate balance between the extrinsic (cash, prizes, giveaways) and the intrinsic (or largely academic, where pride takes an important role).

ETS offers the following "best practices" for motivating students taking the Major Field Tests. These motivation ideas come from MFT user institutions; ETS does not necessarily endorse these methods. It is widely acknowledged that the same approaches and incentives will not work for every program or institution. It is a matter of finding the incentive that speaks to your students and which is best aligned with your institutional culture.

There are two objectives in motivating students in your assessment efforts: getting them to take the test and getting them to take the test seriously. Note that many of the suggested incentives that follow may only serve the purpose of getting students to take the test (if taking the test is largely a voluntary effort for your program/institution). "Best practices" indicate that often a combination of incentives may be required to encourage students to "do their best" on an assessment like Major Field Tests.

With any form of motivation technique comes the possibility of introducing sampling bias — the undesirable effect of drawing a sample that may not be representative of your overall population. (Sampling is not recommended for the MFT except for programs graduating extremely large populations in a particular major.) A choice of a specific incentive may draw to your test administration a larger percentage of a particular type of student than the actual percentage of that same type of student represented in your total population. (For example, monetary incentives may disproportionately draw students from certain socio-economic backgrounds, scholarships may draw solely from those students intending to

attend graduate school, and purely academic incentives may disproportionately draw from either high performing students who enjoy academic challenges or lower performing students who require extra credit to pass a course in order to graduate.) It is important to achieve a balance of motivation efforts that result in a representative sample so that you can make inferences about the larger population based on that sample.

TIP: Consider providing students with information on how well they perform against their peers at your institution. In addition, make students aware of resources your campus offers (specific classes, on-campus tutors, co-curricular opportunities, etc.) to help them improve the kinds of knowledge that are assessed by MFT assessments.

TIP: If using longitudinal assessment design, tell students they will be assessed again and that they will be able to see the progress they have made in developing this important knowledge over time.

TIP: Offer incentives for performance, such as a reception with esteemed colleagues at your institution or letters or recommendation from the dean or chairperson of the program.

TIP: Illustrate how the results from the assessment are used to show the performance of the program. Also show students how the results are used to assist in improving the curriculum.

TIP: Tell students that employers value the knowledge assessed by the MFT and tell them why.

TIP: Request a motivation analysis from ETS to identify examinees that may respond rapidly, which could indicate that they are not motivated. Also, consider measuring student motivation using a self-report measure.

TIP: Determine the stakes of the assessment and clearly communicate those to students. For instance, an example of a low-stakes motivation technique might be to place a hold on students' records and course registration if they failed to complete the assessments. Another technique might include a requirement for low performing students to meet with their program head and academic adviser to discuss their performance on the assessment. Yet another technique could be to simply provide instructions to students that explain the importance of these results for your institution.[3]

[3] Liu, O. L., Bridgeman, B., & Adler, R. M. (2012). Measuring learning outcomes in higher education: Motivation matters. *Educational Researcher*, *41*(9), 352–362.

Use rigorous data collection methodologies and appropriate data analyses.

Both cross-sectional and longitudinal data from Major Field Tests can be useful for assessing student learning. However, longitudinal data at both the student level and program level (i.e., pre-test/post-test, pre-intervention/post-intervention, etc.) from MFT assessments allow institutional stakeholders to gauge student development over time and evidence improvements in learning (e.g., from pre-intervention or "baseline" to after students experience the modified or redesigned educational intervention or experience).

TIP: Use trained proctors and protocols or scripts when administering assessments. This ensures standardization across assessment sessions within the institution. Moreover, proctor scripts can be modified to convey the importance of student engagement in the assessment process.

TIP: Collect pre-test data before implementing new pedagogies or curricula, making pedagogical or curricular modifications, etc. Also, it is important to collect post-test data using the same assessment modules and data collection processes, after students have experienced the new or modified pedagogies/curricula. The combination of MFT pre- and post-test data allow faculty to demonstrate learning improvement.

TIP: Note the stability of your pretest data or data collected at the freshmen level. If your freshmen performance is stable over time, you may want to shift your assessment resources to your post-test, or target larger subgroups of freshmen and junior/seniors.

TIP: Keep faculty research questions at the forefront of data analysis and interpretation.

TIP: Share results with students and faculty. Also, consider sharing results back to external stakeholders like ETS, especially when Major Field Test scores are used to demonstrate student learning improvements. In the future, ETS can highlight these results on their website as institutional case studies.

Use Results of Major Field Tests to design or modify pedagogies and curricula.

When MFT assessments are well-aligned with your major or program-level student learning outcomes, results can be used to help faculty implement new types of pedagogies or curricula. They can also aid faculty with pedagogical and curricular modifications. For instance, you can request descriptive, item-level data for a particular cohort of students to further pinpoint strengths and weaknesses in the curriculum and make modifications accordingly. If students are not meeting expectations for a particular learning outcome, based on assessment results, you can target pedagogical and curricular modifications to courses/learning experiences that emphasize the content or skills related to that learning outcome.

TIP: There are two ways to go about identifying educational activities that can be implemented in the curriculum to improve learning. First is by consulting the research literature on each the skills you are measuring. Second, crowd source ideas from faculty members who understand your curriculum and your students.

TIP: Consult with faculty development experts to help faculty articulate their program theory, re-design courses, examine course scaffolding, implement evidence-driven pedagogies, etc.

Contextual Factors Impacting Major Field Tests Score Interpretations

In addition to the aforementioned considerations and tips for best practices in assessment, faculty should consider certain contextual factors when making decisions or modifications based on Major Field Tests results. These include the following:

Number of major or program courses students have completed.

The Major Field Tests are intended to measure program or discipline-specific skills (e.g., biology, psychology, etc.). It is important to consider how many applicable courses students have completed prior to taking the assessment. Students who have completed more hours of coursework in these types of major or program courses should (in theory) earn higher scores on the Major Field Tests, compared to students who have completed none or fewer major or program courses. Administering the Major Field Tests to students before they have completed their major or program coursework and again after they have completed most or all of their major or program coursework can help your institution demonstrate the extent to which your major or program curriculum positively influences students' growth or development over time.

Knowledge or skills students already have.

Another important contextual factor to consider is the knowledge, skills, and abilities that students bring with them to their educational environments. Different students will enter your institutional programs and courses with varying degrees of discipline-specific skills. Longitudinal data collection methodology can help track student development and acknowledge the skills or abilities that students enter your institution with. However, this methodology requires users to understand retention or attrition in their institution or program.

Use of the Major Field Tests can aid your institution in understanding what knowledge students enter with, and how much they have gained through major or program coursework. That is, the Major Field Tests can be used as part of longitudinal research designs to assess students' knowledge and skills prior to them beginning any major or program coursework, and again once they have successfully completed most or all of their major or program classes.

Special groups such as transfer, non-traditional, and developmental students.

As higher education is expanding, institutions must serve an increasingly diverse population of students, including those who have transferred from two- or four-year institutions, those who entered higher education from "non-traditional" pathways having previous learning experiences, and those who require developmental educational experiences to achieve success. These are important subpopulations of students to assess. Institutions need to gauge these students' skills and abilities, while also understanding how to help improve their learning. Moreover, ignoring or excluding these subpopulations of students may limit or compromise the interpretations or conclusions institutions can draw from their assessment results. Institutions should consider creating infrastructures and using methodologies that allow for the assessment of these students.

Appendix A

Table 1
MFT Test Takers' Statistics
(based on retired forms)

Test	Total Number of Institutions	Total Number of Test Takers	The Highest Percent of Students within Institution of the Total Students
Biology	444	19,067	2.63%
Business	607	102,097	1.86%
Chemistry	251	8,187	5.17%
Computer Science	405	9,422	2.27%
Criminal Justice	260	16,045	6.71%
Economics	79	2,181	6.14%
Literature	177	5,445	4.28%
Math	328	6,806	3.57%
Music	106	1,908	11.32%
Physics	190	2,608	3.26%
Political Science	149	6,445	4.92%
Psychology	289	12,023	2.57%
Sociology	150	4,639	5.97%
Associate Business	85	6,321	11.68%
MBA	266	23,963	3.68%

Table 2
Summary of Classical Item Difficulty and Item Discrimination Indices
(based on retired forms)

				Item Difficulty		Item Discrimination	
	Test	**Form Code**	**N of Items[1]**	**Mean**	**SD**	**Mean**	**SD**
1	**Associate Business**	4DMF	94	48.45	19.31	0.24	0.08
2	**Biology**	4GMF	145	46.90	17.98	0.28	0.10
3	**Business**	4GMF	115	50.55	20.56	0.25	0.09
4	**Chemistry**	4HMF	100	46.68	14.36	0.33	0.10
5	**Computer Science**	4HMF	64	43.25	17.67	0.32	0.11
6	**Criminal Justice**	4GMF	145	53.33	20.33	0.31	0.09
7	**Economics**	4EMF	88	57.16	19.11	0.31	0.07
8	**Literature**	4HMF	148	48.89	19.25	0.29	0.10
9	**MBA**	4FMF	117	52.26	18.55	0.27	0.11
10	**Mathematics**	4IMF	50	31.45	13.01	0.28	0.09
11	**Music**	4AMF	127	51.35	17.93	0.28	0.07
12	**Physics**	4IMF	70	45.96	19.72	0.29	0.12
13	**Political Science**	4HMF	129	59.46	20.30	0.32	0.10
14	**Psychology**	4GMF	137	54.72	16.86	0.33	0.08
15	**Sociology**	4IMF	137	49.51	19.90	0.31	0.10

[1] N of Items in table above represent only the questions that contribute to scores

Table 3
Reliability Coefficients and Standard Errors of Measurement (SEM)
(based on retired forms)

Reliability Coefficients and Standard Errors of Measurement (SEM)				
Test (Form)	**Items[1]**	**Examinees**	**Reliability Coefficient**	**SEM**
Associate Business (4DMF)	94	4,523	0.82	5.26
Biology (4GMF)	145	36,474	0.92	3.83
Business (4GMF)	115	122,687	0.87	4.93
Chemistry (4HMF)	100	3,959	0.92	4.25
Computer Science (4HMF)	64	4,086	0.87	5.63
Criminal Justice (4GMF)	145	9,778	0.93	3.79
Economics (4EMF)	88	5,437	0.89	5.31
Literature (4HMF)	148	3,089	0.93	4.55
MBA (4FMF)	117	28,818	0.89	5.22
Mathematics (4IMF)	50	2,828	0.77	8.56
Music (4AMF)	127	6,102	0.91	4.45
Physics (4IMF)	70	895	0.84	6.01
Political Science (4HMF)	129	3,473	0.93	3.70
Psychology (4GMF)	137	21,268	0.94	3.79
Sociology (4IMF)	137	2,129	0.93	3.15

Note: MBA includes all test takers; Associate Business includes students with 31–90 credit hours.

[1] N of Items in table above represent only the questions that contribute to scores

Table 4
Scaled Score Means and Standard Errors of Measurement (SEM)
for Male and Female Examinees
(based on retired forms)

Scaled Score Means and Standard Errors of Measurement (SEM)							
Test (Form)	**Total Items[1]**	**N**		**Mean Score**		**SEM**	
		M	**F**	**M**	**F**	**M**	**F**
Associate Business (4DMF)	94	1,590	2,824	549.56	545.21	5.26	5.25
Biology (4GMF)	145	14,051	21,432	156.61	151.19	3.77	3.87
Business (4GMF)	115	64,522	53,698	154.86	150.22	4.95	4.87
Chemistry (4HMF)	100	1,884	1,984	152.61	143.71	4.27	4.21
Computer Science (4HMF)	64	3,477	496	151.00	143.38	5.64	5.56
Criminal Justice (4GMF)	145	5,045	4,380	155.07	151.92	3.75	3.82
Economics (4EMF)	88	3,855	1,474	159.48	153.30	5.28	5.38
Literature (4HMF)	148	846	2,131	154.39	153.40	4.55	4.55
MBA (4FMF)	117	15,384	12,393	251.75	245.78	5.21	5.19
Mathematics (4IMF)	50	1,428	1,325	160.07	151.36	8.37	9.01
Music (4AMF)	127	2,776	3,279	152.24	147.99	4.43	4.45
Physics (4IMF)	70	729	149	152.58	145.36	6.00	5.99
Political Science (4HMF)	129	1,872	1,473	155.96	148.28	3.64	3.76
Psychology (4GMF)	137	4,906	15,675	158.12	155.18	3.80	3.79
Sociology (4IMF)	137	618	1,425	149.56	148.32	3.18	3.13

Note: MBA includes all test takers; Associate Business includes students with 31–90 credit hours.

[1] N of Items in table above represent only the questions that contribute to scores

Table 5
Scaled Score Means and Standard Errors of Measurement (SEM) for African American and White Examinees
(based on retired forms)

Scaled Score Means and Standard Errors of Measurement (SEM)							
Test (Form)	**Total Items[1]**	**N**		**Mean Score**		**SEM**	
		White	**Afr. Am.**	**White**	**Afr. Am.**	**White**	**Afr. Am.**
Associate Business (4DMF)	94	3,335	611	548.44	540.49	5.28	5.21
Biology (4GMF)	145	25,139	4,026	155.68	141.73	3.75	4.08
Business (4GMF)	115	85,884	16,574	155.17	143.02	4.95	4.78
Chemistry (4HMF)	100	2,618	585	150.88	137.92	4.28	4.09
Computer Science (4HMF)	64	2,981	355	152.63	134.80	5.65	5.35
Criminal Justice (4GMF)	145	6,195	2,299	157.40	144.84	3.73	3.87
Economics (4EMF)	88	3,926	328	159.44	143.88	5.28	5.35
Literature (4HMF)	148	2,402	356	156.24	137.96	4.55	4.42
MBA (4FMF)	117	18,863	3,742	253.18	237.75	5.29	4.96
Mathematics (4IMF)	50	2,185	251	157.33	143.11	8.59	8.90
Music (4AMF)	127	4,998	447	151.01	137.50	4.46	4.35
Physics (4IMF)	70	712	35	152.70	134.66	6.02	5.57
Political Science (4HMF)	129	2,451	538	155.64	140.78	3.64	3.82
Psychology (4GMF)	137	15,183	3,008	158.70	143.88	3.84	3.61
Sociology (4IMF)	137	1,269	536	152.12	140.80	3.16	3.12

Note: MBA includes all test takers; Associate Business includes students with 31–90 credit hours.

[1] N of Items in table above represent only the questions that contribute to scores

Table 6
Scaled Score Means, Reliability and Standard Errors of Measurement (SEM)
(based on retired forms)

Test	Score	Items[1]	Examinees	Mean Score	Reliability Coefficient	SEM
Associate Business (4DMF)	Total	94	4,523	546.76	0.82	5.26
Biology (4GMF)	Total	145	36,474	153.31	0.92	3.83
	Subscore 1	30	36,474	53.32	0.75	6.66
	Subscore 2	28	36,474	53.10	0.75	6.37
	Subscore 3	48	36,474	53.19	0.76	6.75
	Subscore 4	39	36,474	53.00	0.77	6.38
Business (4GMF)	Total	115	122,687	152.75	0.87	4.93
Chemistry (4HMF)	Total	100	3,959	148.04	0.92	4.25
	Subscore 1	28	3,959	48.10	0.77	7.35
	Subscore 2	30	3,959	48.36	0.81	6.37
	Subscore 3	26	3,959	48.16	0.77	7.19
	Subscore 4	30	3,959	48.19	0.76	7.16
Computer Science (4HMF)	Total	64	4,086	150.03	0.87	5.63
Criminal Justice (4GMF)	Total	145	9,778	153.60	0.93	3.79
Economics (4EMF)	Total	88	5,437	157.71	0.89	5.31
	Subscore 1	40	5,437	58.24	0.81	6.77
	Subscore 2	34	5,437	56.35	0.74	8.08
Literature (4HMF)	Total	148	3,089	153.61	0.93	4.55
	Subscore 1	102	3,089	53.52	0.90	5.30
	Subscore 2	46	3,089	53.71	0.79	7.59
	Subscore 3	78	3,089	53.59	0.90	5.37
	Subscore 4	50	3,089	53.36	0.78	7.92
MBA (4FMF)	Total	117	28,818	249.04	0.89	5.22
Mathematics (4IMF)	Total	50	2,828	155.87	0.77	8.56
Music (4AMF)	Total	127	6,102	149.94	0.91	4.45
	Subscore 1	41	6,102	50.81	0.74	7.51
	Subscore 2	40	6,102	49.42	0.82	6.24
	Subscore 3	46	6,102	49.90	0.78	6.59

(continued)

Table 6
Scaled Score Means, Reliability and Standard Errors of Measurement (SEM)
(based on retired forms)

Test	Score	Items[1]	Examinees	Mean Score	Reliability Coefficient	SEM
Physics (4IMF)	Total	70	895	151.23	0.84	6.01
	Subscore 1	38	895	51.11	0.76	7.23
	Subscore 2	32	895	51.48	0.68	8.62
Political Science (4HMF)	Total	129	3,473	152.62	0.93	3.70
	Subscore 1	58	3,473	52.71	0.87	5.02
	Subscore 2	24	3,473	52.66	0.74	7.08
	Subscore 3	29	3,473	52.50	0.76	6.72
Psychology (4GMF)	Total	137	21,268	155.77	0.94	3.79
	Subscore 1	27	21,268	54.99	0.76	7.29
	Subscore 2	22	21,268	55.83	0.71	7.74
	Subscore 3	25	21,268	56.22	0.72	7.92
	Subscore 4	30	21,268	55.35	0.80	6.56
Sociology (4IMF)	Total	137	2,129	148.59	0.93	3.15
	Subscore 1	45	2,129	48.48	0.82	5.29
	Subscore 2	38	2,129	48.19	0.80	5.31

[1] N of Items in table above represent only the questions that contribute to scores

Table 6a
Scaled Score Means and Standard Errors of Measurement (SEM) for African American and White Examinees
(based on retired forms)

Test	Score		N		Mean Score		SEM	
	Name	Items[1]	Afr. Am.	White	Afr. Am.	White	Afr. Am.	White
Associate Business (4DMF)	Total	94	611	3,335	540.49	548.44	5.21	5.28
Biology (4GMF)	Total	145	4,026	25,139	141.73	155.68	4.08	3.75
	Subscore 1	30	4,026	25,139	44.51	54.97	6.97	6.58
	Subscore 2	28	4,026	25,139	44.96	54.72	6.71	6.31
	Subscore 3	48	4,026	25,139	43.04	55.22	6.97	6.66
	Subscore 4	39	4,026	25,139	41.07	55.67	6.87	6.23
Business (4GMF)	Total	115	16,574	85,884	143.02	155.17	4.78	4.95
Chemistry (4HMF)	Total	100	585	2,618	137.92	150.88	4.09	4.28
	Subscore 1	28	585	2,618	37.95	51.02	6.92	7.44
	Subscore 2	30	585	2,618	41.07	50.27	6.20	6.41
	Subscore 3	26	585	2,618	38.83	50.81	6.86	7.25
	Subscore 4	30	585	2,618	38.25	51.05	6.88	7.20
Computer Science (4HMF)	Total	64	355	2,981	134.80	152.63	5.35	5.65
Criminal Justice (4GMF)	Total	145	2,299	6,195	144.84	157.40	3.87	3.73
Economics (4EMF)	Total	88	328	3,926	143.88	159.44	5.35	5.28
	Subscore 1	40	328	3,926	44.24	59.90	6.82	6.72
	Subscore 2	34	328	3,926	46.34	57.64	8.02	8.06
Literature (4HMF)	Total	148	356	2,402	137.96	156.24	4.42	4.55
	Subscore 1	102	356	2,402	37.93	56.11	5.14	5.31
	Subscore 2	46	356	2,402	40.10	56.03	7.45	7.58
	Subscore 3	78	356	2,402	37.52	56.35	5.31	5.37
	Subscore 4	50	356	2,402	40.71	55.42	7.58	7.95

(continued)

Table 6a
Scaled Score Means and Standard Errors of Measurement (SEM) for African American and White Examinees
(based on retired forms)

Test	Score		N		Mean Score		SEM	
	Name	Items[1]	Afr. Am.	White	Afr. Am.	White	Afr. Am.	White
MBA (4FMF)	Total	117	3,742	18,863	237.75	253.18	4.96	5.29
Mathematics (4IMF)	Total	50	251	2,185	143.11	157.33	8.90	8.59
Music (4AMF)	Total	127	447	4,998	137.50	151.01	4.35	4.46
	Subscore 1	41	447	4,998	41.36	51.71	7.28	7.54
	Subscore 2	40	447	4,998	37.65	50.47	6.14	6.24
	Subscore 3	46	447	4,998	38.60	50.77	6.30	6.61
Physics (4IMF)	Total	70	35	712	134.66	152.70	5.57	6.02
	Subscore 1	38	35	712	34.91	52.58	6.88	7.25
	Subscore 2	32	35	712	37.54	52.66	7.69	8.66
Political Science (4HMF)	Total	129	538	2,451	140.78	155.64	3.82	3.64
	Subscore 1	58	538	2,451	42.47	55.55	5.26	4.92
	Subscore 2	24	538	2,451	41.45	55.35	7.23	7.01
	Subscore 3	29	538	2,451	42.39	54.93	6.86	6.65
Psychology (4GMF)	Total	137	3,008	15,183	143.88	158.70	3.61	3.84
	Subscore 1	27	3,008	15,183	45.30	57.41	6.90	7.38
	Subscore 2	22	3,008	15,183	46.23	58.19	7.40	7.81
	Subscore 3	25	3,008	15,183	45.55	58.85	7.73	7.96
	Subscore 4	30	3,008	15,183	44.59	57.97	6.31	6.62
Sociology (4IMF)	Total	137	536	1,269	140.80	152.12	3.12	3.16
	Subscore 1	45	536	1,269	40.60	52.13	5.17	5.32
	Subscore 2	38	536	1,269	40.88	51.54	5.23	5.33

[1] N of Items in table above represent only the questions that contribute to scores

Table 6b
Scaled Score Means and Standard Errors of Measurement (SEM) for Male and Female Examinees
(based on retired forms)

Test	Score		N		Mean Score		SEM	
	Name	Items[1]	Female	Male	Female	Male	Female	Male
Associate Business (4DMF)	Total	94	2,824	1,590	545.21	549.56	5.25	5.26
Biology (4GMF)	Total	145	21,432	14,051	151.19	156.61	3.87	3.77
	Subscore 1	30	21,432	14,051	51.43	56.24	6.74	6.54
	Subscore 2	28	21,432	14,051	51.78	55.20	6.43	6.28
	Subscore 3	48	21,432	14,051	51.19	56.28	6.78	6.69
	Subscore 4	39	21,432	14,051	50.99	56.12	6.47	6.25
Business (4GMF)	Total	115	53,698	64,522	150.22	154.86	4.87	4.95
Chemistry (4HMF)	Total	100	1,984	1,884	143.71	152.61	4.21	4.27
	Subscore 1	28	1,984	1,884	44.32	52.10	7.27	7.39
	Subscore 2	30	1,984	1,884	45.01	51.91	6.36	6.37
	Subscore 3	26	1,984	1,884	43.73	52.83	7.06	7.28
	Subscore 4	30	1,984	1,884	44.16	52.40	7.09	7.19
Computer Science (4HMF)	Total	64	496	3,477	143.38	151.00	5.56	5.64
Criminal Justice (4GMF)	Total	145	4,380	5,045	151.92	155.07	3.82	3.75
Economics (4EMF)	Total	88	1,474	3,855	153.30	159.48	5.38	5.28
	Subscore 1	40	1,474	3,855	54.59	59.70	6.89	6.70
	Subscore 2	34	1,474	3,855	52.16	58.03	8.14	8.05
Literature (4HMF)	Total	148	2,131	846	153.40	154.39	4.55	4.55
	Subscore 1	102	2,131	846	53.33	54.25	5.31	5.29
	Subscore 2	46	2,131	846	53.43	54.54	7.58	7.60
	Subscore 3	78	2,131	846	53.50	54.10	5.38	5.37
	Subscore 4	50	2,131	846	53.17	54.00	7.92	7.92
MBA (4FMF)	Total	117	12,393	15,384	245.78	251.75	5.19	5.21

(continued)

Table 6b
Scaled Score Means and Standard Errors of Measurement (SEM) for Male and Female Examinees
(based on retired forms)

Test	Score		N		Mean Score		SEM	
	Name	Items[1]	Female	Male	Female	Male	Female	Male
Mathematics (4IMF)	Total	50	1,325	1,428	151.36	160.07	9.01	8.37
Music (4AMF)	Total	127	3,279	2,776	147.99	152.24	4.45	4.43
	Subscore 1	41	3,279	2,776	48.79	53.15	7.49	7.47
	Subscore 2	40	3,279	2,776	47.43	51.78	6.27	6.18
	Subscore 3	46	3,279	2,776	48.64	51.41	6.57	6.58
Physics (4IMF)	Total	70	149	729	145.36	152.58	5.99	6.00
	Subscore 1	38	149	729	46.02	52.33	7.19	7.22
	Subscore 2	32	149	729	45.75	52.73	8.52	8.63
Political Science (4HMF)	Total	129	1,473	1,872	148.28	155.96	3.76	3.64
	Subscore 1	58	1,473	1,872	48.89	55.64	5.13	4.91
	Subscore 2	24	1,473	1,872	48.65	55.76	7.19	6.98
	Subscore 3	29	1,473	1,872	48.32	55.77	6.79	6.63
Psychology (4GMF)	Total	137	15,675	4,906	155.18	158.12	3.79	3.80
	Subscore 1	27	15,675	4,906	54.34	57.50	7.27	7.37
	Subscore 2	22	15,675	4,906	55.21	58.15	7.74	7.73
	Subscore 3	25	15,675	4,906	56.10	56.96	7.92	7.90
	Subscore 4	30	15,675	4,906	55.00	56.82	6.56	6.54
Sociology (4IMF)	Total	137	1,425	618	148.32	149.56	3.13	3.18
	Subscore 1	45	1,425	618	48.15	49.71	5.25	5.37
	Subscore 2	38	1,425	618	47.93	49.05	5.27	5.37

[1] N of Items in table above represent only the questions that contribute to scores

Table 7
Correlations among Subscores
(based on retired forms)

Major Field Test / Subscore			Subscores			
	Items[1]	**Examinees**	**(1)**	**(2)**	**(3)**	**(4)**
Biology (4GMF)	145	36,474				
(1) Cell Biology	30		-	0.72	0.68	0.61
(2) Molecular Biology and Genetics	28		0.97	-	0.65	0.61
(3) Organismal Biology	48		0.91	0.87	-	0.65
(4) Population Biology, Evolution, and Ecology	39		0.82	0.82	0.87	-
Chemistry (4HMF)	100	3,959				
(1) Physical Chemistry	28		-	0.67	0.75	0.81
(2) Organic Chemistry	30		0.86	-	0.67	0.71
(3) Inorganic Chemistry	26		0.98	0.86	-	0.77
(4) Analytical Chemistry	30		1.00	0.91	1.00	-
Economics (4EMF)	88	5,437				
(1) Microeconomics	40		-	0.72		
(2) Macroeconomics	34		0.93	-		
Literature (4HMF)	148	3,089				
(1) Literature 1900 and Earlier	102		-	0.81	0.96	0.86
(2) Literature 1901 and Later	46		0.97	-	0.85	0.81
(3) Literary Analysis	78		1.00	1.00	-	0.75
(4) Literary History and Identification	50		1.00	1.00	0.90	-
Music (4AMF)	127	6,102				
(1) Listening Comprehension	41		-	0.64	0.61	
(2) Written Theory	40		0.83	-	0.70	
(3) Written History	46		0.82	0.88	-	

(continued)

Table 7
Correlations among Subscores
(based on retired forms)

Major Field Test / Subscore			Subscores			
	Items[1]	**Examinees**	**(1)**	**(2)**	**(3)**	**(4)**
Physics (4IMF)	70	895				
(1) Introductory Physics	38		-	0.70		
(2) Advanced Physics	32		0.99	-		
Political Science (4HMF)	129	3,473				
(1) United States Government and Politics	58		-	0.73	0.71	
(2) Comparative Politics	24		0.95	-	0.74	
(3) International Relations	29		0.90	1.00	-	
Psychology (4GMF)	137	21,268				
(1) Learning and Cognition (including Language, Memory and Thinking)	27		-	0.69	0.67	0.75
(2) Perception, Sensory, Physiology, Comparative and Ethology	22		0.95	-	0.66	0.67
(3) Clinical, Abnormal, and Personality	25		0.94	0.95	-	0.70
(4) Developmental and Social	30		0.99	0.92	0.97	-
Sociology (4IMF)	137	2,129				
(1) Core Sociology	45		-	0.87		
(2) Critical Thinking	38		1.00	-		

Note. This table shows observed correlations and correlations adjusted for unreliability. Correlations above and to the right of the diagonal have been observed; the correlations shown in the lower left are estimates of each pair of subscores if there were no measurement error.

[1] N of Items in table above represent only the questions that contribute to scores

Table 8
Summary of Demographic Information Reported by Test
(based on retired forms)

	Associate Business (4DMF)		Biology (4GMF)		Business (4GMF)	
	N	**%**	**N**	**%**	**N**	**%**
Number of Examinees	4,956		39,834		134,948	
Gender						
Males	1,792	37	15,246	39	70,574	54
Females	3,046	63	23,446	61	59,140	46
Ethnic Subgroup						
American Indian/Alaskan Native	33	<1	354	<1	813	<1
Asian/Pacific American	77	2	2,659	7	5,925	5
Black/African American	656	14	4,354	11	17,623	13
Mexican American/Chicano	83	2	762	2	2,895	2
Puerto Rican	34	<1	247	<1	778	<1
Latin American/Other Hispanic	108	2	1,813	5	4,914	4
White	3,531	76	27,140	69	91,180	69
Other	129	3	1,792	5	7,386	6
Transfer						
No	3,852	81	28,671	73	80,552	61
Yes	902	19	10,361	27	51,245	39
Enrollment Status						
Full Time	3,190	67	36,046	92	115,666	88
Part Time	1,558	33	3,110	8	16,523	12
English as Best Language						
Equal	163	3	2,171	6	6,679	5
No	235	5	2,607	7	9,440	7
Yes	4,360	92	34,303	88	116,008	88
Undergraduate GPA						
<1.00	1	<1	2	<1	21	<1
1.00 - 1.99	15	<1	42	<1	166	<1
2.00 - 2.49	294	6	1,446	4	6,946	5
2.50 - 2.99	1,098	23	7,372	19	35,271	27
3.00 - 3.49	1,702	36	15,336	40	52,763	41
3.50 - 4.00	1,625	34	14,425	37	34,957	27
Major GPA						
<1.00	841	18	4	<1	17	<1
1.00 - 1.99	18	<1	140	<1	184	<1
2.00 - 2.49	332	7	1,696	4	4,673	4
2.50 - 2.99	98	2	7,675	20	25,972	20
3.00 - 3.49	2,638	57	15,377	41	55,570	43
3.50 - 4.00	717	15	12,912	34	41,621	33

	Chemistry (4HMF)		Computer Science (4HMF)	
	N	**%**	**N**	**%**
Number of Examinees	4,505		4,447	
Gender				
Males	2,137	49	3,783	88
Females	2,263	51	536	12
Ethnic Subgroup				
American Indian/Alaskan Native	31	<1	40	<1
Asian/Pacific American	331	7	256	6
Black/African American	643	15	381	9
Mexican American/Chicano	60	1	90	2
Puerto Rican	25	<1	14	<1
Latin American/Other Hispanic	206	5	138	3
White	2,943	66	3,169	72
Other	194	4	296	7
Transfer				
No	3,465	78	3,076	70
Yes	957	22	1,288	30
Enrollment Status				
Full Time	4,105	93	3,870	89
Part Time	328	7	485	11
English as Best Language				
Equal	278	6	223	5
No	463	10	273	6
Yes	3,681	83	3,863	89
Undergraduate GPA				
<1.00	0	<1	0	<1
1.00 - 1.99	3	<1	6	<1
2.00 - 2.49	148	3	278	6
2.50 - 2.99	795	18	1,105	26
3.00 - 3.49	1,610	37	1,621	38
3.50 - 4.00	1,814	42	1,270	30
Major GPA				
<1.00	0	<1	1	<1
1.00 - 1.99	11	<1	9	<1
2.00 - 2.49	170	4	165	4
2.50 - 2.99	864	20	842	20
3.00 - 3.49	1,620	38	1,563	37
3.50 - 4.00	1,639	38	1,609	38

	Economics (4EMF)		Literature (4HMF)	
	N	**%**	**N**	**%**
Number of Examinees	6,334		3,302	
Gender				
Males	4,440	71	901	28
Females	1,770	29	2,283	72
Ethnic Subgroup				
American Indian/Alaskan Native	17	<1	13	<1
Asian/Pacific American	446	7	39	1
Black/African American	370	6	371	11
Mexican American/Chicano	91	1	48	1
Puerto Rican	29	<1	11	<1
Latin American/Other Hispanic	282	5	81	2
White	4,425	72	2,560	79
Other	501	8	120	4
Transfer				
No	5,079	82	2,241	69
Yes	1,099	18	1,008	31
Enrollment Status				
Full Time	5,797	94	2,982	92
Part Time	388	6	256	8
English as Best Language				
Equal	360	6	52	2
No	744	12	160	5
Yes	5,077	82	3,037	93
Undergraduate GPA				
<1.00	0	<1	0	<1
1.00 - 1.99	0	<1	2	<1
2.00 - 2.49	218	4	118	4
2.50 - 2.99	1,155	19	583	18
3.00 - 3.49	2,427	40	1,053	33
3.50 - 4.00	2,312	38	1,459	45
Major GPA				
<1.00	0	<1	0	<1
1.00 - 1.99	15	<1	2	<1
2.00 - 2.49	189	3	65	2
2.50 - 2.99	983	16	363	11
3.00 - 3.49	2,451	41	1,030	33
3.50 - 4.00	2,380	40	1,699	54

	MBA (4FMF)		Mathematics (4IMF)	
	N	**%**	**N**	**%**
Number of Examinees	40,825		3,145	
Gender				
Males	19,380	55	1,581	52
Females	15,810	45	1,482	48
Ethnic Subgroup				
American Indian/Alaskan Native	230	<1	26	<1
Asian/Pacific American	1,869	6	143	5
Black/African American	4,063	12	274	9
Mexican American/Chicano	764	2	44	1
Puerto Rican	154	<1	10	<1
Latin American/Other Hispanic	1,214	4	73	2
White	20,800	63	2,375	76
Other	3,710	11	162	5
Transfer				
No	27,970	88	2,326	75
Yes	3,978	12	779	25
Enrollment Status				
Full Time	20,866	65	2,819	91
Part Time	11,300	35	282	9
English as Best Language				
Equal	2,458	8	124	4
No	3,512	11	270	9
Yes	26,360	82	2,704	87
Undergraduate GPA				
<1.00	7	<1	1	<1
1.00 - 1.99	4	<1	4	<1
2.00 - 2.49	403	1	94	3
2.50 - 2.99	3,696	12	418	14
3.00 - 3.49	12,312	40	1,054	34
3.50 - 4.00	14,507	47	1,500	49
Major GPA				
<1.00	8	<1	0	<1
1.00 - 1.99	2	<1	4	<1
2.00 - 2.49	66	<1	114	4
2.50 - 2.99	820	3	499	17
3.00 - 3.49	8,461	33	1,035	35
3.50 - 4.00	16,347	64	1,340	45

	Music (4AMF)		Physics (4IMF)	
	N	**%**	**N**	**%**
Number of Examinees	6,768		1,002	
Gender				
Males	3,043	45	812	83
Females	3,658	55	171	17
Ethnic Subgroup				
American Indian/Alaskan Native	34	<1	4	<1
Asian/Pacific American	147	2	39	4
Black/African American	480	7	38	4
Mexican American/Chicano	55	<1	9	<1
Puerto Rican	38	<1	3	<1
Latin American/Other Hispanic	122	2	54	5
White	5,472	83	789	80
Other	206	3	47	5
Transfer				
No	5,363	81	799	81
Yes	1,287	19	186	19
Enrollment Status				
Full Time	6,206	93	903	91
Part Time	465	7	88	9
English as Best Language				
Equal	146	2	36	4
No	548	8	94	10
Yes	5,956	90	859	87
Undergraduate GPA				
<1.00	0	<1	0	<1
1.00 - 1.99	2	<1	0	<1
2.00 - 2.49	97	1	33	3
2.50 - 2.99	859	13	163	17
3.00 - 3.49	2,329	35	364	37
3.50 - 4.00	3,356	51	422	43
Major GPA				
<1.00	0	<1	0	<1
1.00 - 1.99	0	<1	5	<1
2.00 - 2.49	37	<1	35	4
2.50 - 2.99	429	7	166	17
3.00 - 3.49	2,127	32	335	35
3.50 - 4.00	3,992	61	421	44

	Political Science (4HMF)		Psychology (4GMF)		Sociology (4IMF)	
	N	**%**	**N**	**%**	**N**	**%**
Number of Examinees	3,863		23,082		2,297	
Gender						
Males	2,075	56	5,272	24	658	30
Females	1,636	44	17,068	76	1,552	70
Ethnic Subgroup						
American Indian/Alaskan Native	25	<1	160	<1	33	1
Asian/Pacific American	77	2	595	3	39	2
Black/African American	602	16	3,175	14	566	25
Mexican American/Chicano	57	2	375	2	71	3
Puerto Rican	28	<1	201	<1	9	<1
Latin American/Other Hispanic	122	3	779	3	77	3
White	2,692	71	16,385	72	1,367	60
Other	193	5	1,026	5	105	5
Transfer						
No	2,833	75	15,094	67	1,547	69
Yes	964	25	7,541	33	705	31
Enrollment Status						
Full Time	3,548	94	20,604	91	2,066	91
Part Time	236	6	2,072	9	202	9
English as Best Language						
Equal	153	4	770	3	70	3
No	305	8	1,449	6	131	6
Yes	3,320	88	20,498	90	2,059	91
Undergraduate GPA						
<1.00	0	<1	0	<1	0	<1
1.00 - 1.99	10	<1	29	<1	7	<1
2.00 - 2.49	211	6	1,115	5	188	8
2.50 - 2.99	858	23	5,077	23	658	30
3.00 - 3.49	1,412	38	8,842	39	862	39
3.50 - 4.00	1,249	33	7,362	33	515	23
Major GPA						
<1.00	0	<1	0	<1	0	<1
1.00 - 1.99	6	<1	46	<1	3	<1
2.00 - 2.49	104	3	743	3	101	5
2.50 - 2.99	564	15	3,641	17	443	20
3.00 - 3.49	1,469	40	8,613	39	871	40
3.50 - 4.00	1,530	42	9,001	41	761	35

Table 9
Speededness Analyses
(based on retired forms)

Test	Percent Completing 100% of the Test	Percent Completing 75% of the Test	Variance Index of Speededness	Number of Items reached by 80% of Candidates	Total Number of Items	Number of Test Takers
Associate Business (4DMF)	97.48	99.23	0.15	100	100	4954
Biology (4GMF)	95.41	99.47	0.08	150	150	39834
Business (4GMF)	97.01	99.64	0.07	120	120	134956
Chemistry (4HMF)	95.98	99.60	0.04	100	100	4505
Computer Science (4HMF)	95.21	99.78	0.03	66	66	4450
Criminal Justice (4GMF)	95.90	99.10	0.12	150	150	10408
Economics (4EMF)	96.73	99.79	0.03	90	90	6334
Literature (4HMF)	91.88	98.88	0.12	150	150	3302
MBA (4FMF)	96.33	98.97	0.17	124	124	40851
Mathematics (4IMF)	94.28	99.55	0.06	50	50	3145
Music (4AMF)	93.32	99.42	0.06	129	129	6767
Physics (4IMF)	97.11	99.80	0.03	70	70	1002
Political Science (4HMF)	97.98	99.56	0.05	130	130	3863
Psychology (4GMF)	97.80	99.52	0.05	140	140	23084
Sociology (4IMF)	96.44	99.52	0.05	140	140	2301