2014-06-01

Patterns of Student Curricular Experience in Psychology as Predictors of Performance on the ETS Major Field Test

Jason J. van der Horst
Brigham Young University - Provo

Patterns of Student Curricular Experience in Psychology as Predictors of Performance on the ETS Major Field Test

Jason van der Horst

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Bruce L. Brown, Chair
Richard R. Sudweeks
Dawson W. Hedges
Jeffrey S. Reber
Shawn Gale

Department of Psychology

Brigham Young University

July 2014

ABSTRACT

Patterns of Student Curricular Experience in Psychology
as Predictors of Performance on the
ETS Major Field Test

Jason van der Horst
Department of Psychology, BYU
Doctor of Philosophy

The purpose of this dissertation is to evaluate the relationship between student performance and their performance on the Major Field Test (MFT). The MFT purports to adequately assess student mastery and achievement in the college major, in this case psychology. The major advantages of the MFT over internally-created instruments are its standardized content, its established national norms, and its connection to the Graduate Record Examination (GRE). The MFT is the most widely used standardized test for learning outcomes assessment within psychology departments.

The first hypothesis, that MFT scores are reflective of summary curricular values (i.e. GPA), was not supported when ACT composite scores are regressed out. ACT composite score by itself is predictive of MFT performance which casts doubt on its claim to be reflective of achievement in one's college program. The results of the second hypothesis, regarding prediction of MFT scores from grades in specific courses within the major, provided positive support for the use of the MFT test. In this second analysis, I found higher multiple R-squared values for predicting MFT scores from specific course grades with R-squared values substantially stronger than the ACT bivariate regression of hypothesis 1. This helps to support the claim that the MFT measures specific achievement within one's major. The results for the third hypothesis, prediction of MFT from particular mix of courses taken in the major, were somewhat supportive. Prediction of MFT scores was found to be strongest for the subscale area Perception and Physiology, and the strongest predictor of these scores (t value of 3.78) is student completion of the Brain, Behavior, and Cognition course group.

Keywords: psychology, learning outcomes, assessment, curriculum

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Patterns of Student Curricular Experience in Psychology as Predictors of Performance on the ETS Major Field Test

Modern higher education systems are an essential step in many career paths. Current policies outline the need to evaluate the outcomes of student learning to provide external accountability for use of funds and also an internal method of improving teaching and learning. Individual departments define the expected learning outcomes for students completing specific degree programs. Assessing expected versus observed learning outcomes will aid in many departmental decision-making processes. The purpose of this dissertation is to evaluate the relationship of the Major Field Test (MFT) as a standardized measure of student summative performance and achievement within the psychology major.

The BYU Psychology Department has five learning outcomes that are examined in a threefold continuous improvement cycle consisting of measurement, statistical analysis, and planning. The purpose of this dissertation is to focus solely on the first of the five learning outcomes, knowledge of the discipline of psychology, as measured by the Major Field Test. It also focuses on the second step of the continuous improvement cycle, statistical analysis of patterns, specifically to identify the curricular predictors that predict knowledge outcomes in the field of psychology. The results of this dissertation are expected to be of value to the BYU psychology department in refining the learning outcomes evaluation process. However, the fundamental purpose of this study as a dissertation is to evaluate the usefulness of the Major Field Test (MFT) in assessing student summative performance and achievement within the psychology major, and to demonstrate methods of predicting learning outcomes that can be applied more generally to other similar academic departments.

History of Learning Outcome Assessments

Higher education in the United States has changed greatly since the first institutions of higher education were started in the 1600's. Higher education not only serves as a means of establishing lifelong learning patterns, but also as a fundamental portal of entry into a variety of highly desirable career paths. As a well-rounded education has become increasingly important for employment preparation, college enrollment numbers are increasing along with educational costs (Bogue, 2000).

Increased enrollment in and costs of higher education were contributing factors that led to the Joint Declaration of the European Ministers of Education in 1999, which established the "Bologna Process". Initially, the purpose of this process was to strengthen the competitiveness and attractiveness of the European higher education system, with the concept of quality assurance taking center stage from the outset (Adam, 2004). Over the past ten years the European Ministers of Education have broadened this agenda to include other areas such as modifying the undergraduate/postgraduate degree structure into a three-cycle system (bachelor, master, doctorate) with an emphasis on learning outcomes. The ministers organized the European Higher Education Area (EHEA) in 2010 with the following priorities:

- Social dimension
- Lifelong learning
- Employability
- Student-centered learning
- Education, research and innovation
- Mobility
- Data collection
- Multidimensional transparency tools
- Funding.

Many of these priorities, and also the central idea of learning outcomes, have also become important priorities across the Atlantic within the United States.

The United States Department of Education (2006) published a report and several issue papers that call for reforms in the country's higher education system in six key areas: (a) access, (b) cost and affordability, (c) financial aid, (d) learning, (e) transparency and accountability, and (f) innovation. In the area of Transparency and Accountability, the report outlines that the Department of Education:

> …believe(s) that improved accountability is vital to ensuring the success of all the other reforms we propose. Colleges and universities must become more transparent about cost, price, and student success outcomes, and must willingly share this information with students and families. Student achievement, which is inextricably connected to institutional success, must be measured by institutions,… should be made available to students, and reported publicly in aggregate form to provide consumers and policymakers an accessible, understandable way to measure the relative effectiveness of different colleges and universities.

This statement regarding the vital nature of improved accountability outlines a clear need for accurately assessing student learning outcomes.

Universities use four focus areas of learning outcomes: (a) those used in individual teaching events; (b) those specified for modules or short courses; (c) course-level outcomes; and (d) those outcomes specified for entire degree programs (Hussey & Smith, 2008). On the program level, these statements outline the knowledge, skills, and attitudes that learners should have after successfully completing a program, which are also known as expected learning outcomes. The EHEA priorities (2004) and US Department of Education report (2006) outline the need for assessment of student learning outcomes for each degree level. Accountability comes in measuring the alignment between expected learning outcomes and observed learning outcomes, which are the actual knowledge, skills, and attitudes learners have obtained upon completion of a specific program. Measuring this alignment is not a new context and evidence of this goes back to the 1900s.

Joseph Mayer Rice (1857-1934) was a physician, journal editor, education critic, and originator of comparative methodology in educational research. Notably, in 1892, Rice undertook an exhaustive survey of public schools from the East Coast to the Midwest. Rice published a series of articles on urban education in the magazine *The Forum* in 1892 and 1893. These articles were used by parents to petition politicians to improve public school processes and facilities.

Learning outcome assessments of academic programs emerged again during the late 1970's as universities looked for ways to improve academic excellence and create a competitive advantage in response to declining student enrollments (Gatson, 1991; Pace, 1979). State governments and accrediting agencies began to consider requiring outcome assessments to receive public funding during the mid-1980's (Ewell et al., 1990; Spangehl, 1987; Thrash, 1990). Over the past three decades the assessment of student learning outcomes has progressed steadily in scope and sophistication. Today, the assessment of student learning outcomes has evolved to serve two main purposes: (a) summative external accountability for use of funds; and (b) a formative internal method for improving teaching and learning.

Federal government, state legislatures, students, parents, and the regional accreditation bodies, among others, increasingly demand that institutions of higher education demonstrate that they meet their stated educational missions and goals and that the evidence they provide is objectively and continuously gathered and reported (Dugan and Hernon, 2002). Accountability is increasingly important when coupled with rising enrollments and the recent worldwide economic downturn that threatens the affordability of higher education (Kuh, 2010).

In a coordinated effort to evaluate student learning outcomes, bodies governing higher education have been working to address this need for accountability. They continually review

and revise the standards that institutions use to demonstrate effectiveness. At the national level, both the American Association for Higher Education (AAHE) and the Council for Higher Education Accreditation (CHEA) have sought to increase awareness of the need for measuring and reporting student learning outcomes. However, because of the decentralized structure of accreditation of higher education in the United States, it has become the responsibility of the regional higher education institutional accreditation associations to develop standards by which to hold institutions accountable through evaluation, including the imposition of student learning outcome measures (McMurtrie, 1999). These measures include direct measures, indirect measures, and comparative data.

Measurement Tools

There are many ways that institutions of higher education can gather information regarding student performance, some of which have been widely used by universities for years. Such student statistics include graduation rates, retention rates, transfer rates, and employment rates for a graduating class. These institution outcomes are used to track internal change in year-to-year performance, and also as a comparative measure with other institutions (Frye, 1999). Such outcomes reflect what the institution has accomplished but do not reflect what (or how much) students have learned. Courses, credits, certificates, and degrees are important proxies for student accomplishment, but they are only proxies. It is the knowledge and skills that students take with them when they graduate that yields the personal, economic, and societal benefits promised by higher education (Kuh, 2010).

Direct Measures: One of the main ways institutions of higher education measure learning outcomes is by standardized examinations. The appeal of standardized tests is partly due to convenience, but also the ability to compare scores to a national average. The most

common criticism of standardized examinations is that in order to provide effective guidance, tools for gathering student measures need to be specific to the context of each university (Donald, 2002; Mentkowski, 2000).

Popular alternatives to standardized examinations include curriculum-embedded assessments and student portfolios. Under the curriculum-embedded approach, faculties collectively identify specific assignments located at key points in a curricular sequence that can be used to examine particular learning outcomes at particular levels of performance. Doing this requires the institution to create its own version of a qualifications framework and requires secondary reading of student responses by faculty who did not teach the class in which the response was generated (Lewis, 2010). Under the portfolio process, students post examples of their work in an accessible electronic medium, grouped under learning outcomes specified by the institution or program, as evidence that they have mastered each area. This evidence is then evaluated by faculty using specially developed rubrics (Tartwijk et al., 2007). The main criticism with this type of measure is the time necessary to obtain meaningful data.

Indirect Measures: In addition to such direct measures of student learning outcomes, there are several types of indirect measures, such as the results of commonly administered surveys of currently enrolled students and recent graduates, as self-report measures of knowledge gained. Results drawn from such surveys can be used in combination with results of direct assessments of learning outcomes to target what needs to be improved in student populations. The National Survey of Student Engagement (NSSE), for instance, is regularly administered to students in many of America's institutions of higher education. This self-report measure is used to collect information at hundreds of four-year colleges and universities about student participation in programs and activities that institutions provide for their learning and personal

development. The results provide an estimate of how undergraduates spend their time and what they gain from attending college (Ewell, 2010). Of course, direct evidence of student learning is of paramount importance in assessing learning outcomes, but indirect evidence from self-report sources such as these can also provide much valuable contextual information in evaluating the educational experience.

Comparative Data: Using assessment to improve teaching and learning can be considerably enhanced if assessment results can be benchmarked against established standards or across institutions. Such benchmarking not only enables institutions and programs to know where they stand, but also allows them to identify potential "best practices" from which they can learn.

Learning Outcomes Within Psychology

Within the area of psychology, the American Psychological Association (APA) Board of Education Affairs (BEA) commissioned a task force to describe a set of learning goals and outcomes for student performance at the completion of the baccalaureate degree. This task force published the "Undergraduate Psychology Major Learning Goals and Outcomes" report (APA, 2002). Subsequently, BEA charged the task force to develop a document to address assessment strategies based upon these learning goals and outcomes. The first edition of the "The Assessment CyberGuide for Learning Goals and Outcomes" was created in 2002. The document outlines ten goals and suggested learning outcomes that represent reasonable departmental expectations for the undergraduate psychology major across educational contexts. The goals are divided into two major categories:

(1) knowledge, skills, and values consistent with the science and application of psychology, and

(2) knowledge, skills, and values consistent with liberal arts education that are further developed in psychology.

In 2006, the APA Council of Representatives adopted a revised version of the task force's initial report under the title "APA Guidelines for the Undergraduate Psychology Major".

In 2009, the Assessment Cyberguide Revision Task Force updated the CyberGuide to reflect current practice, and revised the links. This second edition of the CyberGuide (Pusateri and Halonen, 2009) serves as a companion resource for implementing the "APA Guidelines for the Undergraduate Major in Psychology". These resources are designed to aid psychology department faculty in designing the most appropriate and effective assessment plans. Within the CyberGuide are several potential sources of data. The CyberGuide discusses the advantages and disadvantages of each source of data, and provides some general recommendations on each data source. These data sources are: course data, individual projects/performance assessment, summative performance assessment, self-assessment/reflection, collaboration, interviews and surveys, and archival measures. Overall they recommend that:

> …the disadvantages of the use of standardized tests can be minimized with some additional planning. Embedding the capstone test in an existing course will enhance student motivation since the student may take the experience more seriously. When student performance can also be tied to course grading, maximum motivation to do well is likely. Describing how well the existing test matched the required curriculum will encourage faculty support and student cooperation.

While all of these measures can play an essential role they are not all used universally. The most common measure used across psychology programs are standardized tests.

The most widely used standardized test for learning outcome assessments in psychology is the Educational Testing Service (ETS) Major Field Test (MFT) for Psychology (see Table 1). Development of the MFT began in 1989 and was modeled after the Graduate Record Examination (GRE) Subject Test. While the GRE is designed to predict graduate school success,

the MFT is designed to measure the basic knowledge and understanding achieved by senior undergraduates. The MFT for Psychology has been used by more than 350 institutions of higher education over the past five years. The Psychology MFT consists of 140 multiple-choice questions designed to assess the most common and most important topics and skills within psychology (for a complete list of areas covered in the Psychology MFT see Appendix A). Questions are drawn from courses of study most commonly offered in undergraduate programs. Questions require students to identify theories, psychologists, methods, and other information from particular subareas. Additional questions require students to analyze relationships, apply principles, draw conclusions from experimental data, and evaluate experiments. In addition to the standard question pool, individual institutions may add up to 50 locally written questions.

Table 1
Percentages of Psychology Programs Using Particular Assessment Methods.

Assessment Method	All 4-Year Programs	Doctoral	Masters	Bachelors	Community College
Internal Instrument	46%	47%	42%	47%	26%
ETS Major Fields Test	25%	11%	29%	33%	1%
ACAT	7%	4%	8%	9%	0%
Other	22%	18%	21%	25%	13%
Senior Survey	3%	4%	3%	2%	0%
Course-Embedded Assessment	3%	3%	2%	4%	4%
Portfolio Review	3%	1%	7%	2%	1%
Thesis/Research Project	3%	0%	2%	5%	0%
Alumni Survey	2%	2%	0%	2%	0%
Exit Interview or Focus Group	2%	0%	2%	4%	0%
Employment/Grad School Admission	1%	0%	1%	1%	0%
None	20%	32%	14%	15%	34%

Note: Data for this table was gathered by APA from 2003-2005.

Results for the MFT include summary information for the entire group of test takers along with individual student scores. Overall student scores are reported on a scale of 120-200. Four additional subarea scores are reported on a scale of 20-100 and are provided for the following areas: (a) Learning and Cognition, (b) Sensory and Physiology, (c) Clinical, Abnormal, and Personality, and (d) Developmental and Social. When scoring student responses, ETS only scores correct answers so that students are not penalized for omissions or guesses. In order to be included in overall scores, participants need to complete at least 50% of the MFT questions. ETS also provides comprehensive national comparative data for the MFT, enabling institutions to evaluate student performance and compare performance to programs at institutions nationwide.

The MFT has been used in several learning outcome research articles. Most notable is Stoloff and Feeney's (2002) article that sought to predict MFT scores along with several other student outcome measures within their psychology program. Stepwise multiple regression found that the overall MFT score was well predicted, r^2=.47, by the combination of GPA (.44) and the completion of several courses: Social Psychology, Abnormal Psychology, and Biopsychology. Psychology GPA or any grades in particular psychology courses were not found to be significant predictors. All regressions performed by Stoloff and Feeney grouped summary curricular values along with specific course values. While providing overall predictability, this does not capture individual variable type relationship. Dolinsky and Kelly (2010) followed the analytic plan of Stoloff and Feeney and found similar results within their own psychology program. These findings show that MFT has a relation to a number of important college activities and measures. Given a larger psychology program, this study looks to

Both of these studies show how multiple regression analysis can be usefully applied to study a number of relevant college predictors and outcomes. Multiple regression analysis is also considered one of the most useful forms of statistical analysis for studying the effects of college on students in Pascarella and Terenzini's (2005) book, *How College Affects Students*. Given its proven utility in prior related research, the present study also employed multiple regression analysis to test its hypotheses.

Learning Outcome Assessment at BYU's Psychology Department

The Psychology Department at Brigham Young University (BYU) has defined five expected learning outcomes for students completing the Psychology BS degree. These learning outcomes are:

1. Demonstrate knowledge of the discipline of psychology in each of four areas: (a) learning and cognition, (b) sensation and physiological psychology, (c) clinical and personality, and (d) developmental and social psychology. (These are the four subareas of the MFT.)
2. Students will demonstrate that they understand and can apply basic research methods in psychology including research design, psychometrics, data analysis, and interpretation and writing of results in light of previous findings.
3. Students will be able to use computers and other research-related technology to collect, access, manage, and interpret research information.
4. Students will be able to critically reflect on the content of psychology as well as on disciplinary values.

5. Students will be able to apply the principles of psychology in an internship setting (399R), a teaching (410R) or community service setting (420R), or in a research mentoring setting (430R)

The first learning outcome is a *content outcome*; it deals with knowledge of the field of psychology. The assessment for this learning outcome is the MFT. The other four learning outcomes are *competency outcomes.* They deal with essential skills necessary for performance within the field. Learning outcomes 2-4 will be measured by a departmentally constructed assessment of skills in methodology, numeracy, and technology. Learning outcome 5 will be assessed by the Internship Profiling Questionnaire (IPQ).

In even numbered academic years (i.e. 2010-2011, 2012-2013, etc.) the content outcome (learning outcome 1) is measured using the MFT, or a more targeted internally-constructed instrument. In odd numbered academic years (i.e. 2011-2012, 2013-2014, etc.) three of the competency outcomes (learning outcomes 2-4) are measured. The fifth learning outcome is continuously measured throughout each semester/term as part of the internship class. This is part of a continuous improvement cycle of (1) measurement, (2) statistical analysis, and (3) planning. Outcomes from the continuous improvement cycle will be used to inform current decisions that the department may be facing and identify future areas of focus from the various aspects of the department. From a student's perspective, potential changes will be most apparent through the Psychology Map. It is through this document that the department specifies required courses, course clusters, and a suggested per-semester layout of when to complete various course options. Given its utility, the present study will be using the requirements outlined in the BYU Psychology BS map as part of its analysis, specifically course groupings (called clusters) and the order in which students complete their course.

This dissertation focuses solely on the first of the Psychology Department's learning outcomes, the content of the field of psychology as measured by the Major Field Test. By understanding the relationship between student performance, based on measures gathered by the university (i.e. grades, courses completed, GPA), and their performance on the MFT this study is the initial implementation for the first and second steps of the continuous improvement cycle, measurement and statistical analysis of patterns.

Objectives

Overall, the hypotheses listed below investigate the relationship between student outcomes and their performance on the MFT. Student performance will be broken into four specific areas: (a) summary curricular measures (i.e. GPA), (b) individual course grades, (c) number of courses completed within course groups (i.e. clusters defined by BYU Psychology Department), and (d) order in which students completed their course work. Student performance areas (a) and (b) are pulled from Stoloff and Feeney's work (2002). This study anticipates finding similar results based on area (a) summary curricular measures but, given a larger psychology program, I anticipate that area (b) will have stronger results. This dissertation adds to their work by also looking at the relationship between the four MFT subscores and the groups of courses outlined in the BYU Psychology Program Map (Developmental & Clinical Cluster, Social Cluster, Behavior Brain and Cognition Cluster, Methods, and other requirements). The BYU Psychology Program also outlines a particular order in which courses should be completed and this sequence of courses will be included in this dissertation.

In particular, I seek to identify the curricular predictors that account for strong outcomes in knowledge of the field of psychology, as measured by higher MFT scores, by addressing the following four research hypotheses:

H1: The 5 MFT scores will be strongly predicted by summary curricular values (Total GPA, BYU GPA, BYU Graded Credit Hours, Major Credit Hours, and Major GPA).

H2: The 5 MFT scores will be strongly predicted by grades in specific courses within the psychology major.

H3: The 5 MFT scores will be predicted by number of courses completed within course groups as outlined in the BYU Psychology Program Map.

H4: The 5 MFT scores will be predicted by the student's sequence of courses.

Method

Participants: Participants were Brigham Young University undergraduate students enrolled in a psychology capstone courses (Psych 399R, Psych 410R, Psych 420R, or Psych 430R) during fall semester 2010. Of the 142 students enrolled in a psychology capstone course during that semester, 108 of them completed the MFT. Of this, 104 participants were included in the analyses due to the fact that 3 participants were underclassmen and 1 was a Psych Minor. Participant selection is based on a new department policy requiring senior psychology students to participate in learning outcomes evaluation as part of their capstone course. Students who are currently enrolled in a capstone course that are not seniors may opt out of taking the MFT if they will complete a capstone course during their senior year. Student participation is required since the MFT constitutes the final examination for their capstone course. Demographic information for participants is outlined in Table 2.

Materials: The Major Field Test (MFT) subject matter test for psychology was completed by 104 participants. The MFT consists of 140 multiple choice questions drawn from courses of study most commonly offered in undergraduate psychology programs. Overall

student scores are reported on a scale of 120-200 with subscores reported on a scale of 20-100. Subscores are provided for each of the following areas: (a) Learning and Cognition, (b) Perception, Sensation, and Physiology, (c) Clinical, Abnormal, and Personality, and (d) Developmental and Social. The MFT takes an average of two hours to complete. All participants must complete this survey.

Procedure: The MFT was available to the students through the campus testing center from November 22nd through December 9th of 2010. Participants were informed of the test's availability by their capstone instructor.

Table 2
Demographic Characteristics for MFT Respondents

Demographic Variable	Mean ± StDev	Group	N
Total GPA	3.41 ± 0.42		
Major GPA	3.42 ± 0.43		
Age:	23.3 ± 4.6	19	4
		20	13
		21	18
		22	17
		23	16
		24	15
		>=25	21
Gender:		Female	63
		Male	41
Ethnicity		Asian/Asian-American	1
		Black/African-American	2
		Hawaiian/Part Hawaiian	2
		Hispanic-Latino	9
		White	89
		Not Specified	1

Additional Information

Additional curricular information for each participant was provided by Student Academic and Advisement Services (SAAS). Curricular data points for each participant include such things as course enrollments, course grades, overall credit hours, major GPA, and overall BYU GPA as well as course sequence information.

Results

The MFT data are combined with and analyzed with curricular data in four ways corresponding to the four research questions listed above: predictions of MFT scores from summary curricular values, from grades in specific courses, from particular mix of courses taken, and from time sequence of the courses.

Hypothesis 1: Summary Curricular Values

The first hypothesis states that the five MFT scores will give a strong and consistent reflection of summary curricular values. There are five specific summary curricular values that I will be targeting: Total GPA, BYU GPA, BYU Graded Credit Hours, Major Credit Hours, and Major GPA. In previous studies, this is generally evaluated with ACT composite score included as one of the predictors. I feel it is of more interest to analyze these predictors as part of two separate analyses. First, I will setup a canonical correlation procedure where the effects of the ACT scores are removed. After that I will analyze ACT score by itself as a predictor for comparison. This will allow us to compare college level summary variables to a pre-college summary variable.

The canonical correlation analysis is used as a filter to guide subsequent multiple regression analyses. It identifies the maximum *R-squared* value for all possible multiple regression analyses (criterion scores, Y, predicted from multiple X-score predictor values). I

begin this analysis by calculating the residual scores for the five summary curricular measures with the effects of ACT score removed. In order to do this I conducted five bivariate regressions, one separate regression for each of the summary curricular variables as an individual criterion with ACT as the predictor for each. From here I determined the residual value by subtracting the predicted value from the original value. These residuals from the five regressions were then combined to make the predictor data set with the influence of ACT scores regressed out.

This data set is summarized in Table 3, which shows the 10x10 intercorrelation matrix with a 5x5 Rxx submatrix of intercorrelations among the five criterion variables, a 5x5 Ryy submatrix of intercorrelations among the five predictor variables, and a 5x5 Rxy submatrix of correlation coefficients between each of the five criterion variables and each of the five predictor variables. This adjoined 10x10 correlation matrix is the basis of the first canonical correlation analysis. Notice in this adjoined correlation matrix that the five MFT scores (Total MFT and the four subscores) all have a fairly low correlation with each of the GPA predictor variables (Total GPA, BYU GPA, and Psychology GPA) when ACT score is regressed out.

Table 3

Correlation Matrix with the Five MFT Criterion Variables and the Residual Scores for the Five Summary Curriculum Predictor Variables in the Upper-Right Half Matrix, and Sample Size Corresponding to Each Correlation Coefficient in the Lower Left Half.

	Total	L & C	Phys	C & P	D & S	T GPA	GPA	CHrs	PsychHrs	PsychGPA
Total	1	0.8668	0.7917	0.8958	0.9075	0.1646	0.1401	-0.2084	-0.2726	0.0995
L & C	104	1	0.6524	0.7285	0.7189	0.1982	0.1912	-0.1841	-0.3148	0.1597
Phys	104	104	1	0.6261	0.6473	0.0356	0.0057	-0.1714	-0.2129	-0.0152
C & P	104	104	104	1	0.7756	0.0540	0.0336	-0.1946	-0.2286	-0.0157
D & S	104	104	104	104	1	0.1975	0.1742	-0.2423	-0.2877	0.1450
T GPA	103	103	103	103	103	1	0.9780	-0.3077	-0.1110	0.9093
GPA	101	101	101	101	101	100	1	-0.2059	-0.0911	0.9221
CHrs	99	99	99	99	99	99	96	1	0.4498	-0.2351
PsychHrs	100	100	100	100	100	100	98	97	1	-0.0807
PsychGPA	94	94	94	94	94	94	92	92	93	1

Note: Total = MFT Overall Score, L & C = Learning and Cognition MFT Subscore, Phys = Perception, Sensory, Physiology, Comparative, and Evolutionary MFT Subscore, C & P = Clinical, Abnormal, and Personality MFT Subscore, D & S = Developmental and Social MFT Subscore,T GPA = Total GPA, GPA = BYU GPA, C Hrs = Overall Credit Hours, Psych Hrs = Psychology Credit Hours. This correlation matrix has the Rxx triangular matrix for the five MFT Criterion Variables in the upper left, the Ryy triangular matrix of five curriculum Predictor variables in the lower right, and the Rxy matrix of correlations between Predictor Variables and Criterion Variables in the lower left.

Now that I have ACT scores regressed out, the canonical correlation analysis was conducted where the Y set of variables (the criterion variables) includes the 5 MFT measures (overall score and four subscores), and the X set of variables consists of five predictor variables: Total GPA, BYU GPA, BYU Graded Credit Hours, Major Credit Hours, and Major GPA—all with ACT composite score regressed out.

Table 4 contains the summary table for this canonical correlation analysis. The first three canonical correlations are 0.4315, 0.3505, and 0.2214. None of these canonical correlation coeffients are statistically significant for this three-latent-variable solution (see coefficients and corresponding *p* values in the lower left hand corner of the table). We can interpret this first canonical correlation (0.4315) like any other correlation coefficient, which allows us to find the percent of variance accounted for by squaring *r* (0.186). This means that no multiple *R-squared*

for any subsequent regression analysis will be larger than 0.186. Since the canonical correlations were not significant no follow-up regression analyses were conducted for the Summary Curricular Measures. These canonical correlations are, of course, made weaker than those that would be obtained from summary curricular variables without removing the effects of ACT scores. For comparison, this same analysis was carried out uncontrolled for ACT, and the first canonical correlation came out to be 0.6591 ($p = 0.0006$), substantially larger than when controlled for ACT.

Table 4

Canonical Correlation Summary Table with the Y Set of Major Field Test Variables (Criterion) at the Top of the Table, the X Set of five Summary Curricular Measures (Predictor) with ACT Score Regressed out Located at the Bottom of the Table

	Loadings			Squared Loadings				Uniqueness
	LV1	LV2	LV3	LV1	LV2	LV3	Total	U
Y Set (MFT)	*(eta1)*	*(eta2)*	*(eta3)*	*(eta1)*	*(eta2)*	*(eta3)*		
Total MFT Score	.2669	-.1106	-.1519	.0712	.0122	.0231	.1065	.8935
Learning & Cognition	.3370	-.0883	-.0678	.1136	.0078	.0046	.1260	.8740
Perception & Physiology	.1396	-.1516	-.1446	.0195	.0230	.0209	.0634	.9366
Clinical & Personality	.1632	-.1864	-.1340	.0266	.0347	.0180	.0793	.9207
Developmental & Social	.3037	-.1069	-.1304	.0922	.0114	.0170	.1207	.8793
		Sum of squares by columns:		.3232	.0892	.0835	.4959	4.5041
		Percents of sums of squares:		6.46%	1.78%	1.67%	9.92%	90.08%
X Set (Curricular Measures)	*(chi1)*	*(chi2)*	*(chi3)*	*(chi1)*	*(chi2)*	*(chi3)*		
Total GPA	.3056	.1851	-.0523	.0934	.0343	.0027	.1304	.8696
BYU GPA	.3091	.1912	-.0128	.0955	.0366	.0002	.1323	.8677
BYU Credit Hours	-.1888	.1823	.0679	.0356	.0332	.0046	.0735	.9265
Major Credit Hours	-.3238	.1912	.0149	.1048	.0366	.0002	.1416	.8584
Major GPA	.3017	.2047	.0229	.0910	.0419	.0005	.1334	.8666
		Sum of squares by columns:		.4204	.1825	.0083	.6112	4.3888
		Percents of sums of squares:		8.41%	3.65%	0.17%	12.22%	87.78%

	Coefficient	p Value	Multivariate Statistics	Index	p Value
First Canonical Correlation	.4315	.1613	Wilks' Lambda	.6660	.1613
Second Canonical Correlation	.3505	.4717	Pillai's Trace	.3771	.1620
Third Canonical Correlation	.2214	.7934	Hotelling-Lawley Trace	.4396	.1674
			Roy's Greatest Root	.2288	.0054

With the first analysis completed I will now look at how ACT composite score by itself predicts performance on the MFT. This is done by conducting five individual bivariate regressions, one for each of the MFT scores as the criterion and with ACT score as the only

predictor variable. The results for these regressions are summarized in Table 5 (reported in descending order of how well one can account for a specific MFT score using ACT score as the predictor). All of these predictive models were significant except for the Perception, Sensation & Physiology Subscore. The overall strength is low since the highest *R-squared* only accounts for 11.8% of the variance.

Table 5
Table of Summary Values from Five Bivariate Regressions: One Bivariate Regression Analysis for Each of the Five MFT Scores as the Criterion Variable and with ACT as the Predictor Variable

Criterion Variable	*F*	*p*	R^2
Regression 1: Learning & Cognition Subscore	12.05	0.0008	0.118
Regression 2: Total MFT score	7.85	0.0060	0.080
Regression 3: Developmental & Social Subscore	7.58	0.0071	0.078
Regression 4: Clinical & Personality Subscore	4.25	0.0423	0.045
Regression 5: Perception, Sensation & Physiology Subscore	2.41	0.1239	0.015

Hypothesis 2: Grades in Specific Courses

For the second research question, I evaluate how strongly the five MFT scores are predicted by grades in specific courses within the psychology major. I included the nine classes with the highest number of students who completed the course since beyond this point the course enrollment dropped significantly. Grades for these courses were represented on a 4 point GPA scale (4.0 = A, 3.7 = A-, etc.). For this analysis, the Y set of variables (the criterion variables) includes the 5 MFT measures (overall score and four subscores), and the X set of variables

consists of nine predictor variables, grades in these nine courses (listed below in order of enrollment size specified in parenthesis):

- Psych 210: A History of Psychology (103 students),
- Psych 101: Orientation to the Psychology Major (101),
- Psych 301: Psychological Statistics (99),
- Psych 302: Psychological Research Design and Analysis (99),
- Psych 304: Psychological Testing (94),
- Psych 341: Personality (74),
- Psych 111: General Psychology (72),
- Psych 306: Psychology of Gender (72), and
- Psych 350: Introduction to Social Psychology (66).

When pulling together the dataset for the course grades, the correlation matrix used for the canonical correlation was calculated one cell at a time so that it would accommodate multiple sample sizes (see Table 6). If built normally this would cause the analysis to have a much smaller sample size since it would have used the lowest *n* size of 30 for all the correlations.

Table 6

Correlation Matrix with the Five MFT Criterion Variables and Nine Course Grade Predictor Variables in the Upper-Right Half Matrix, and Sample Size Corresponding to Each Correlation Coefficient in the Lower Left Half

	Total	L & C	Phys	C & P	D & S	P210	P101	P301	P302	P304	P341	P111	P306	P350	P303	P342
Total	1	0.8668	0.7917	0.8958	0.9075	0.3793	0.1856	0.4750	0.2662	0.4230	0.3553	0.1568	0.3843	0.2397	0.2970	0.3599
L & C	104	1	0.6524	0.7285	0.7189	0.3801	0.1362	0.4289	0.2380	0.3344	0.2400	0.1156	0.3524	0.1679	0.1703	0.2123
Phys	104	104	1	0.6261	0.6473	0.3129	0.1322	0.3645	0.2062	0.3466	0.3212	0.1469	0.2982	0.2250	0.2466	0.2893
C & P	104	104	104	1	0.7756	0.3007	0.1700	0.3961	0.2197	0.3656	0.2594	0.1182	0.3141	0.2039	0.3645	0.3433
D & S	104	104	104	104	1	0.2382	0.1066	0.3869	0.2367	0.3746	0.3438	0.1512	0.3173	0.2109	0.2222	0.3918
P210	103	103	103	103	103	1	0.4249	0.4404	0.3448	0.5121	0.5356	0.4016	0.6761	0.4595	0.4025	0.2974
P101	101	101	101	101	101	100	1	0.3834	0.2862	0.3394	0.3610	0.3242	0.2080	0.1863	0.2087	0.3995
P301	99	99	99	99	99	99	96	1	0.4110	0.4433	0.4273	0.4692	0.4493	0.3515	0.3864	0.4906
P302	100	100	100	100	100	100	98	97	1	0.5494	0.3881	0.2780	0.3039	0.2815	0.3555	0.3395
P304	94	94	94	94	94	94	92	92	93	1	0.5492	0.4211	0.6282	0.4920	0.6380	0.5015
P341	75	75	75	75	75	75	72	73	74	72	1	0.3768	0.6965	0.4758	0.5674	0.3936
P111	71	71	71	71	71	70	70	69	68	64	53	1	0.3595	0.3914	0.6218	0.6738
P306	72	72	72	72	72	72	70	71	72	67	58	48	1	0.3605	0.6240	0.5957
P350	68	68	68	68	68	68	65	68	67	62	52	48	53	1	0.6195	0.3773
P303	61	61	61	61	61	61	58	59	60	59	48	40	45	40	1	0.5601
P342	51	51	51	51	51	51	50	51	51	49	45	37	40	33	30	1

Using the correlation matrix in Table 6 I find the first three canonical correlations are 0.9200, 0.5983, and 0.3915 (see Table 7). This is a much stronger canonical correlation than in the test of the first hypothesis. Since the first canonical correlation is 0.9200, this means that a multiple *R-squared* for any subsequent regression analysis could be as large as .846.

Table 7

Canonical Correlation Summary Table Based on Corrlation Matrix in Table 6 with the Y Set of Major Field Test Variables (Criterion) at the Top of the Table, and the X Set of Grades for Nine Courses (Predictor) at the Bottom of the Table

	Loadings			Squared Loadings				Uniqueness
	LV1	LV2	LV3	LV1	LV2	LV3	Total	U
Y Set (MFT)	*(eta1)*	*(eta2)*	*(eta3)*	*(eta1)*	*(eta2)*	*(eta3)*		
Total MFT Score	.2669	-.1106	-.1519	.0712	.0122	.0231	.1065	.8935
Learning & Cognition	.3370	-.0883	-.0678	.1136	.0078	.0046	.1260	.8740
Perception & Physiology	.1396	-.1516	-.1446	.0195	.0230	.0209	.0634	.9366
Clinical & Personality	.1632	-.1864	-.1340	.0266	.0347	.0180	.0793	.9207
Developmental & Social	.3037	-.1069	-.1304	.0922	.0114	.0170	.1207	.8793
		Sum of squares by columns:		.3232	.0892	.0835	.4959	4.5041
		Percents of sums of squares:		6.46%	1.78%	1.67%	9.92%	90.08%
X Set (Curricular Measures)	*(chi1)*	*(chi2)*	*(chi3)*	*(chi1)*	*(chi2)*	*(chi3)*		
Psych 210	.1814	.7315	.2708	.0329	.5351	.0733	.6413	.3587
Psych 101	.1637	.3780	.0413	.0268	.1429	.0017	.1714	.8286
Psych 301	.2940	.5883	.4326	.0864	.3461	.1871	.6197	.3803
Psych 302	.1655	.2568	.2836	.0274	.0659	.0804	.1738	.8262
Psych 304	.3352	.3557	.3957	.1124	.1265	.1566	.3955	.6045
Psych 341	.2871	.2126	.5030	.0824	.0452	.2530	.3806	.6194
Psych 111	.1208	.0856	.2138	.0146	.0073	.0457	.0676	.9324
Psych 306	.2260	.4663	.3760	.0511	.2174	.1414	.4099	.5901
Psych 350	.2187	.1687	.2275	.0478	.0285	.0518	.1280	.8720
		Sum of squares by columns:		.4818	1.5150	.9910	2.9878	6.0122
		Percents of sums of squares:		5.35%	16.83%	11.01%	33.20%	66.80%

	Coefficient	p Value	Multivariate Statistics	Index	p Value
First Canonical Correlation	.9200	<.0001	Wilks' Lambda	.0635	<.0001
Second Canonical Correlation	.5983	.0406	Pillai's Trace	1.5791	<.0001
Third Canonical Correlation	.5641	.3915	Hotelling-Lawley Trace	6.5910	<.0001
			Roy's Greatest Root	5.5077	<.0001

Just as the canonical correlation analysis had stronger values than the previous hypothesis I find stronger results in the follow-up multiple regression analyses as well. All of the regression analyses are significant (see Table 8) as well as all of the predictors within each regression.

Table 8
Table of Summary Values from Five Multiple Regressions: One Multiple Regression Analysis for Each of the Five MFT Scores as the Criterion Variable and with Grades from Nine Courses as the Predictor Variables

Criterion Variable	*F*	*p*	R^2
Regression 1: Clinical & Personality Subscore	1561.97	< .0001	0.632
Regression 2: Total MFT score	1084.74	< .0001	0.544
Regression 3: Developmental & Social Subscore	884.81	< .0001	0.494
Regression 4: Perception, Sensation & Physiology Subscore	576.02	< .0001	0.338
Regression 5: Learning & Cognition Subscore	381.26	< .0001	0.296

Overall, I found that we are better able to predict MFT scores from grades in specific courses than we can from summary curricular measures or ACT composite score. This finding supports the idea that psychology programs can use the MFT as a direct measure in assessing the alignment between expected learning outcomes and observed learning outcomes. While MFT doesn't seem to be a particularly good measure of overall college performance, it does seem to do better as a measure of achievement and performance in particular courses within the major.

The predictive capability for the regressions in this hypothesis are strongest for Clinical & Personality Subscore (*R-squared* = .632) and weakest for the Learning & Cognition Subscore (*R-squared* = .296).

Hypothesis 3: Mix of Courses Taken

For the third research question, I predict MFT performance, both overall and on subscores, with the pattern of particular courses completed by individual participants. Each of the courses completed by a participant was used to create a frequency count for the following course groups:

A. Developmental & Clinical,
B. Social,
C. Behavior, Brain, and Cognition,
D. Methods, and
E. Other Requirements (see Appendix B for course listings)

The Y set of criterion variables includes the five MFT measures. The X set of predictor variables consists of the frequency of courses each student took for each of the five course groups.

Table 9

Correlation Matrix with the Five MFT Criterion Variables and the Five Course Group Predictor Variables With a Sample Size of 104

	Total	L & C	Phys	C & P	D & S	DevClin	Social	BBandC	methods	otherreq
Total	1.0000									
L & C	0.8668	1.0000								
Phys	0.7917	0.6524	1.0000							
C & P	0.8958	0.7285	0.6261	1.0000						
D & S	0.9075	0.7189	0.6473	0.7756	1.0000					
DevClin	0.0721	0.0365	0.0462	0.1043	0.1597	1.0000				
Social	-0.1176	-0.1526	-0.1740	-0.0945	-0.0489	0.0021	1.0000			
BBandC	0.2337	0.2468	0.3689	0.1622	0.1732	0.0143	-0.0454	1.0000		
methods	0.1654	0.0605	0.0406	0.1924	0.2201	0.3555	0.2175	0.1663	1.0000	
otherreq	-0.1213	-0.1734	-0.0824	-0.1803	-0.0117	0.1372	0.1806	-0.0894	0.1769	1.0000

For this analysis the three canonical correlations are 0.4687, 0.3306, and 0.3210 (see Table 10). All four multivariate statistics are significant for this analysis: Wilks' Lambda = .5954, p = .0018; Pillai's Trace = .4769, p = .0020; Hotelling-Lawley Trace = .5661, p = .0017; and Roy's Greatest Root = .2197, p = .0002. The square of the first canonical correlation coefficient ($.4687^2$=.2197) indicates that no multiple *R-squared* will be larger than 0.220 for any subsequent regression analyses.

Table 10

Canonical Correlation Summary Table with the Y Set of Major Field Test Variables (Criterion) at the Top of the Table, and the X Set of Five Course Groups (Predictor) at the Bottom of the Table

	Loadings			Squared Loadings				Uniqueness
	LV1	LV2	LV3	LV1	LV2	LV3	Total	U
Y Set (MFT)	*(eta1)*	*(eta2)*	*(eta3)*	*(eta1)*	*(eta2)*	*(eta3)*		
Total MFT Score	.4200	-.0558	.7184	.1764	.0031	.5161	.6956	.3044
Learning & Cognition	.5448	-.2037	.5647	.2968	.0415	.3189	.6572	.3428
Perception & Physiology	.8028	.0083	.3271	.6445	.0001	.1070	.7516	.2484
Clinical & Personality	.2668	-.0987	.9334	.0712	.0097	.8712	.9522	.0478
Developmental & Social	.2453	.3248	.6727	.0602	.1055	.4525	.6182	.3818
		Sum of squares by columns:		1.2490	.1599	2.2657	3.6747	1.3253
		Percents of sums of squares:		24.98%	3.20%	45.31%	73.49%	26.51%
X Set (Curricular Measures)	*(chi1)*	*(chi2)*	*(chi3)*	*(chi1)*	*(chi2)*	*(chi3)*		
Developmental & Clinical	.0779	.8159	.4534	.0061	.6657	.2056	.8773	.1227
Social	-.4348	.2568	-.1046	.1891	.0659	.0109	.2659	.7341
Behavior Brain and Cognition	.8402	.0665	.1424	.7059	.0044	.0203	.7306	.2694
Methods	-.2484	.3244	.6909	.0617	.1052	.4773	.6443	.3557
Other Requirements	-.2159	.6429	-.5188	.0466	.4133	.2692	.7291	.2709
		Sum of squares by columns:		1.0094	1.2546	.9833	3.2473	1.7527
		Percents of sums of squares:		20.19%	25.09%	19.67%	64.95%	35.05%

	Coefficient	p Value	Multivariate Statistics	Index	p Value
First Canonical Correlation	.4687	.0018	Wilks' Lambda	.5954	.0018
Second Canonical Correlation	.3306	.0492	Pillai's Trace	.4769	.0020
Third Canonical Correlation	.3210	.0888	Hotelling-Lawley Trace	.5661	.0024
			Roy's Greatest Root	.2816	.0002

Notice that the first latent variable, LV1, is defined in the Y set of criterion variables primarily by the Perception, Sensation and Physiology Subscore, and in the X set of predictor variables primarily by the Behavior, Brain, and Cognition group of courses. It is reasonable that this predictor and this criterion should be strongly linked and from this information in the canonical correlation summary table, I expect that the strongest multiple regression analysis from this set of X and Y variables would have Perception and Physiology as the criterion variable, and that the Behavior, Brain, and Cognition group of courses would be the strongest of the predictors of it. Indeed, that is the case.

In performing the regression analysis for each MFT score I get the values listed in Table 11, in order of how well you can account for a specific MFT score using number of courses completed within course groupings:

Table 11
Table of Summary Values from Five Multiple Regressions: One Multiple Regression Analysis for Each of the Five MFT Scores as the Criterion Variable and with Five Course Groups as the Predictor Variables

Criterion Variable	*F*	*p*	R^2
Regression 1: Perception & Physiology Subscore	3.83	0.0033	0.1634
Regression 2: Clinical & Personality Subscore	2.39	0.0434	0.1086
Regression 3: Learning & Cognition Subscore	2.27	0.053	0.1040
Regression 4: Total MFT Score	2.26	0.0546	0.1033
Regression 5: Developmental & Social Subscore	1.79	0.1214	0.0838

The two MFT scores that are best predicted are the Perception, Sensation & Physiology subscore and the Clinical & Personality subscore. The follow-up multiple regression analysis results are given in Table 12 for Perception & Physiology and in Table 13 for Clinical & Personality. In previously reported regression analyses for the first two research questions, Total MFT score and Developmental & Social subscore had the two highest *R-squared* values. We now see these two criterion variables at the bottom of the list when we are predicting them from the number of courses taken in each of the five groups.

Table 12
Regression Analysis with the Y Set of Perception & Physiology Subscore (Criterion) and the X Set of Five Course Groups (Predictors)

Predictor	Coef.	Std. Err.	*t*	*p*
Developmental & Clinical	0.4484	1.0147	0.44	0.660
Social	-1.6621	1.0445	-1.59	0.115
Behavior Brain and Cognition	5.5214	1.4611	3.78	0.000
Methods	0.0957	2.5257	0.04	0.970
Other Requirements	-0.7591	2.4734	-0.31	0.760

Note: $F(5, 98) = 3.83$, $p = 0.0033$, $R^2 = 0.1207$

In the multiple regression for Perception & Physiology Subscore, the only significant predictor is the course group Behavior, Brain, and Cognition. This can be accounted for by the type of materials covered. Behavior, Brain, and Cognition is more biologically-based, and as such is less likely to be known or deduced from a student's general knowledge.

Interestingly, for the Clinical & Personality Subscore the best predictor, and the only statistically significant one, is the course group Other Requirements (non-cluster courses).

Table 13
Regression Analysis with the Y Set of Clinical & Personality Subscore (Criterion) and the X Set of Five Course Groups (Predictor)

Predictor	Coef.	Std. Err.	*t*	*p*
Developmental & Clinical	0.6996	1.3183	0.53	0.597
Social	-1.3684	1.3569	-1.01	0.316
Behavior Brain and Cognition	2.0152	1.8982	1.06	0.291
Methods	6.4818	3.2812	1.98	0.051
Other Requirements	-6.4262	3.2132	-2.00	0.048

Note: $F(5, 98) = 2.39$, $p = 0.0434$, $R^2 = 0.1086$

As I review the third research question, I find additional evidence of value in using the MFT as a measure of observed learning outcomes. While the extent to which I can predict each of the five

MFT scores by the mix of courses taken, as indicated by the R-squared values, isn't as great as that for the second hypothesis, we can glean several insights. Of the five MFT scores, the Perception & Physiology Subscore was best predicted. With this predictive model, we find the most logical set of classes to be the only significant predictor, namely the Behavior, Brain, and Cognition course set.

Hypothesis 4: Time Sequence of Courses

For the fourth research question, I predict MFT performance, both overall and on subscores, using information concerning the sequence in which particular courses are completed by individual students. Two analyses will be used to evaluate to what extent taking courses in the proper specified sequence predicts MFT performance. The first analysis will use at what point of a student's time at the university did they complete specific courses (i.e. did they take Psych 101 when they were a freshman or as a senior). The second analysis will compare these student points to the suggested completion points outlined by the BYU Psychology Department (i.e. Psych 101 should be completed during a student's first semester).

For the first analysis, the Y set of criterion variables includes the 5 MFT measures (overall score and four subscores), and the X set of predictor variables consists of the order and timing in which six courses were completed: Psych 101, Psych 111, Psych 210, Psych 301, Psych 302, and Psych 304. The values for the X set of variables are the percent of total credit hours completed before the student completed each specified course (see Table 14 for sample calculation of percentages, see Table 15 for sample data).

Table 14
Example Calculations for Generating Percent of Total Credit Hours Completed Prior to Each Course for Two Example Students.

Fall 2007	Wint 2008	Fall 2008	…	Fall 2011
15 C.Hrs.	15 CH	15 CH	…	15 CH
Psych 101 Psych 111 Bio 100 Wtrg 150 Rel A 121 Vastu 330	Math 110 Psych 210 AmHrt Rel A 122 GenElec	Psych 301 Hist 201 Rel A 211 GenElec	…	Psych Cluster Course 4 Psych Cluster Elective 2 Psych Cluster Elective 3 Letters GE course General Electives
No classes before, each course is assigned 0/120 = 0	CH completed before = 15/120 = .125 (assigned to each course listed above	30/120 = .25	…	105/120 = .875

Fall 2007	Wint 2008	Fall 2008	…	Fall 2011
14 CH	12 CH	15 CH		18 CH
CS 235 ECEN 124 HEPE 129 PHSCS 121 Rel A 121	CS 236 Engl 115 MFHD 160 REL A 122 REL C 130	CS 224 Hist 201 PHSCS 220 Psych 101 Psych 111 Rel A 211		Music 101 Psych 352 Psych 430R Psych 306 Psych 350 Psych 361
0/154.5 = 0	14/154.5 = .091	26/154.5 = .168		136.5/154.5 = .883

Note: Student S1 started the psychology major during his or her first semester and has the top set of listed credit hour percentages assigned. Student S2 starts the psychology major during his or her third semester and has the bottom set of credit hour percentages.

Table 15

Sample Data Table for Predicting MFT Performance from Timeframe Location for Each Course Completed.

Criterion Variables						Predictor Variables							
TOTAL SCORE	**1**	**2**	**3**	**4**		Psych 101	Psych 111	Psych 210	…	Psych 350	…	Psych 375	…
175	78	81	57	73	(S1)	0	0	.125		.625		.875	
168	70	81	53	59	(S2)	.168	.168	.265		.770		.770	
180	78	81	78	73		0.06	0.09	0.15		0.77		0.97	
172	55	60	74	76		0.18	0.22	0.26		0.53		0.83	
178	78	81	53	80		0.08	0.19	0.20		0.90		0.86	
147	48	53	45	38		0.12	0.13	0.18		0.89		0.72	
161	55	49	57	69		0.18	0.26	0.27		0.65		0.52	
145	48	49	49	38		0.17	0.20	0.21		0.96		0.68	
175	74	49	69	87		0.12	0.23	0.24		0.88		0.64	
169	55	64	61	63		0.03	0.09	0.19		0.81		0.68	
173	78	53	78	63		0.05	0.15	0.18		0.99		0.98	
172	70	77	69	69		0.12	0.17	0.18		0.59		0.75	
187	99	94	78	76		0.12	0.12	0.13		0.72		0.95	
188	82	68	82	91		0.02	0.13	0.20		0.91		0.61	
171	63	57	74	63		0.18	0.23	0.31		0.81		1.21	
…	…	…	…	…		…	…	…		…		…	

The first three canonical correlations are 0.4928, 0.4335, and 0.3367 (see Table 16). Only one of the four multivariate statistics is significant for this three-latent-variable solution, Roy's Greatest Root = 0.3207, $p = 0.0147$. When this is the case the Roy's Greatest Root is generally viewed as an anomaly so there is little justification for conducting follow-up regression analyses.

Table 16

Canonical Correlation Summary Table with the Y Set of Major Field Test Variables (Criterion) at the Top of the Table, the X Set of Course Sequence (Predictor) Located at the Bottom of the Table

	Loadings			Squared Loadings				Uniqueness
	LV1	LV2	LV3	LV1	LV2	LV3	Total	U
Y Set (MFT)	*(eta1)*	*(eta2)*	*(eta3)*	*(eta1)*	*(eta2)*	*(eta3)*		
Total MFT Score	.1893	.5392	-.4993	.0358	.2907	.2493	.5759	.4241
Learning & Cognition	.1795	.3734	-.3693	.0322	.1394	.1364	.3080	.6920
Perception & Physiology	.6753	.3460	-.4723	.4560	.1197	.2231	.7988	.2012
Clinical & Personality	-.0434	.3392	-.3253	.0019	.1151	.1058	.2228	.7772
Developmental & Social	.1373	.8176	-.2546	.0189	.6685	.0648	.7521	.2479
		Sum of squares by columns:		.5448	1.3334	.7794	2.6576	2.3424
		Percents of sums of squares:		10.90%	26.67%	15.59%	53.15%	46.85%
X Set (Curricular Measures)	*(chi1)*	*(chi2)*	*(chi3)*	*(chi1)*	*(chi2)*	*(chi3)*		
Psych 210	-.0328	.9009	-.0273	.0011	.8116	.0007	.8134	.1866
Psych 101	-.4965	.5682	-.3305	.2465	.3229	.1092	.6786	.3214
Psych 302	-.0237	.2654	-.3119	.0006	.0704	.0973	.1683	.8317
Psych 301	-.3967	.3735	-.0259	.1574	.1395	.0007	.2975	.7025
Psych 304	-.1565	.3400	-.4386	.0245	.1156	.1924	.3325	.6675
Psych 111	-.1514	.4170	.2964	.0229	.1739	.0879	.2847	.7153
		Sum of squares by columns:		.4529	1.6339	.4882	2.5750	3.4250
		Percents of sums of squares:		7.55%	27.23%	8.14%	42.92%	57.08%

	Coefficient	p Value	Multivariate Statistics	Index	p Value
First Canonical Correlation	.4928	.2746	Wilks' Lambda	.5369	.2746
Second Canonical Correlation	.4335	.5287	Pillai's Trace	.5594	.2707
Third Canonical Correlation	.3367	.8257	Hotelling-Lawley Trace	.6954	.2866
			Roy's Greatest Root	.3207	.0147

For the second analysis, a similar canonical correlation was run after converting the X set to the difference between the percent of total credit hours completed before a course was taken, to the percent expected by the recommended course order. That is, I applied the same percent credit hours calculations to the recommended psychology map. For example, Psych 101 should be completed in first semester so has a percent credit hours value of 0. In the example above, participant S2 completed Psych 101 in his or her third semester with a percent credit hours value of 0.168. For this new analysis, this cell value will now be 0.168 – 0 = 0.168 (negative numbers means the student completed the course early, positive number means the student completed the course later than the recommended semester). The first three canonical correlations in this

revised analysis are 0.4759, 0.2693, and 0.1903, so there was very little gained in predictive value from this refining of the predictor measures.

This last canonical correlation analysis was the weakest of the four. It appears that there is very little predictive information in the order in which courses are taken, at least for this particular measure of achievement in one's major, the Major Field Test.

Discussion

The MFT is designed to measure the basic knowledge and understanding of psychology achieved by senior psychology undergraduates. This paper investigated the relationship between student performance and their performance on the MFT. Student performance was broken into four specific areas: (a) summary curricular measures (i.e. GPA), (b) individual course grades, (c) number of courses completed within course groups (i.e. clusters defined by BYU Psychology Department), and (d) order in which students completed their course work.

In testing the first research hypothesis I tried to predict MFT performance using five summary curricular measures: Total GPA, BYU GPA, BYU Graded Credit Hours, Major Credit Hours, and Major GPA—all with ACT composite score regressed out. The canonical correlation was not significant so the first hypothesis is not supported. In previous tests of summary curricular values, other researchers (Stoloff and Feeney, 2002) have found significant results. This difference can be attributed to the fact that in the present study the ACT composite scores were not included as a predictor but instead their effects were removed. This is further supported by the fact that when included in a bivariate regression, ACT can predict MFT performance by itself.

This brings up an intriguing possibility. Although the Major Field Test is touted as being a good "direct measure" of college achievement in one's major, it is possible to predict MFT

performance using pre-college data alone. The impact that the ACT composite score has may be attributed to the fact that both this and the MFT are standardized tests and may be influenced by a student's ability to take standardized tests. This may also account for the low percent of variance accounted for when using ACT as the predictor (highest *R-squared = 0.118).* From this first analysis the MFT would seem to be more a reflection of skills, habits, and knowledge acquired before college.

In order to better visualize the extent to which this view of the MFT holds up across the other hypotheses I compare the canonical correlation and regression results (see Figure 1). The leftmost bar for the results from each of these four hypotheses gives the squared canonical correlation coefficient, which represents the maximum possible *R-squared* for any multiple regression analysis. The other three bars are the three highest actually obtained multiple regression R-squared values. The most obvious result in this figure is that grades in specific courses are substantially better at predicting MFT performance. In looking at the *R-squared* value for predicting MFT performance from ACT these are a weak predictor in comparison to specific course grades.

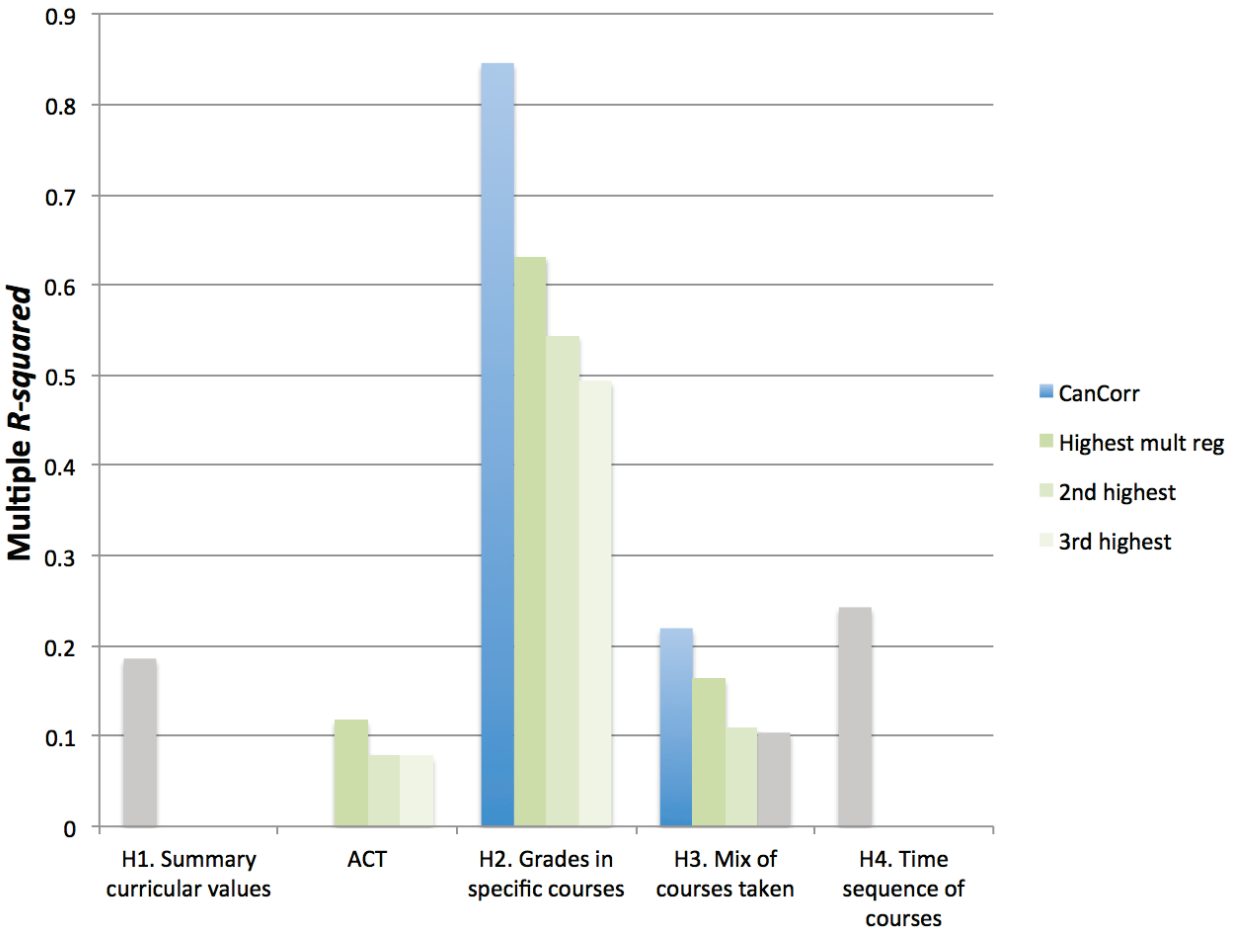


Figure 1. Five sets of bar graphs comparing the *R-squared* values for the analyses for each of the four hypotheses along with predicting MFT from ACT, with *R-squared* values from canonical correlation on the left in each group, and actual multiple regression *R-squared* values in each of the other three bars. Results that are not significant are greyed out.

The Major Field Test results on the whole are reasonably well accounted for by grades in specific psychology courses (see Figure 1). Although it is somewhat disappointing that MFT results can be predicted from high-school level variables, it speaks well for the Major Field Test that the highest predictions of MFT come from specific psychology course performance. Also, even though mix of courses taken is not a strong predictor of MFT scores, the bar for the highest multiple regression in that analysis is the one for Perception & Physiology Subscore. This is an

additional indication that the two most biologically based MFT scores best measure knowledge, skills, and achievements from college work that is also biologically based. This is further punctuated by fact that the Behavior, Brain, and Cognition course group was the strongest predictor.

By pulling together the significant predictors that accompany each of the regression analyses, I get Table 17 that serves as an executive summary. In this table I can see that hypothesis 1 has no significant predictors, aside from the ACT composite score not shown. The predictors that account for the most variance are the grades in specific course. Within this test, all the courses were found to be significant predictors. By looking at hypothesis 3 the table shows that Behavior, Brain, and Cognition is significant for the two subscores that deal with more biology based areas: Learning and Cognition Subscore and Perception and Physiology Subscore.

Table 17
Significant Predictors Associated with Each of the Regression Analyses Performed for Each of the Four Hypotheses

Research Question	Criterion Variable (Y)	Significant Predictor Variables (X)	R^2
H1 - Predictions from Summary Curricular Values	*Not Significant*	*Not Significant*	*Not Significant*
H2 - Predictions from Grades in Specific Courses	Total MFT	All Nine Course Grades	0.544
	Learning & Cognition Subscore	All Nine Course Grades	0.296
	Perception, Sensation, & Physiology Subscore	All Nine Course Grades	0.338
	Clinical, Abnormal, and Personality Subscore	All Nine Course Grades	0.632
	Developmental & Social Subscore	All Nine Course Grades	0.494
H3 - Predictions from Particular Mix of Courses Taken	Total MFT	*Not Sgnificant*	0.103
	Learning & Cognition Subscore	Behavior, Brain, & Cognition Cluster	0.104
	Perception, Sensation, & Physiology Subscore	Behavior, Brain, & Cognition Cluster	0.163
	Clinical, Abnormal, and Personality Subscore	Other Requirements (none cluster courses)	0.109
	Developmental & Social Subscore	*Not Sgnificant*	0.084
H4 - Predictions from Time Sequence of the Courses	*Not Significant*	*Not Significant*	*Not Significant*

In comparing our results to those of Stoloff and Feeney (2002), I found some similarities and differences. Their main findings were that they

> …confirmed that performance on the MFT correlated with the number of psychology courses completed. However, this correlation was weak, and many psychology courses did not contribute to score improvement on the MFT… [and] after completing this analysis we agree that an assessment that relies too much on the MFT will fall short of measuring the full range of knowledge and skills gained during an undergraduate psychology major program (p. 96).

Similarly, I found that the number of specific courses taken in the major does predict MFT but did so with significant results. I also found that specific sets of courses had higher predictive weight than others, namely two of the course sets in the third research set: Behavior, Brain, and Cognition; and also Other Requirements. Of these findings, I find most weight in the Behavior, Brain, and Cognition course set as a predictor of two subscales of the MFT. As opposed to

Stoloff and Feeney, I found very strong relationships between specific courses and MFT performance in specific domains. This may be attributed to the way I controlled for a small sample size. This further supports the idea that the MFT measures what a student has learned in an undergraduate psychology program.

In looking at the BYU psychology department's second step of the continuous improvement cycle, statistical analysis of patterns, I indeed found some interesting patterns and I looked into identifying the curricular predictors that account for strong outcomes in knowledge of the field of psychology. In looking back at the four research questions, I have found the following take away items:

1. High school summary curricular measures predict MFT performance better than college-level summary curricular measures.
2. Individual course grades are better predictors of MFT performance than any other measure used.
3. In looking at course sets, I saw stronger predictions of biologically based MFT subscales from more biologically based courses.

These patterns need additional follow-up work. Specifically, longitudinal data would allow us to better analyze these findings over time. Larger sample sizes would allow analyses across more courses and it would be possible to perform additional drill-down analyses that look into each of the course sets individually. These analyses are particularly important since 25% of 4-year psychology programs use the MFT as the main means of assessment for observed learning outcomes.

Overall, I agree with Stoloff and Feeney (2002) that the MFT by itself is lacking. Psychology programs should look into pulling in additional direct measures, such as curriculum-

embedded assessments and student portfolios and perhaps a more fine-grained unit of learning outcome, such as those used in competency-based education. Granted, the MFT does have notable strengths, including standardization of the measure across universities, the comparative context of national norms, and the connection to GRE scores. There are also clearly several serious disadvantages of the MFT. The high cost of the instrument is a large disadvantage. Also, with standardization comes inflexibility, and the somewhat awkward fit between what the instrument measures and actual achievement, knowledge, and skills achieved, as demonstrated in this study. One of the biggest disadvantages of the MFT is the standard procedure of reporting of only four subscale scores for each respondent, whereas more detailed item-level analysis (factoring items focusing on specific targeted subgroups of items, etc.) could provide substantially more information tailored to each individual program. It is possible to obtain this information from Educational Testing Service, but this makes an already very expensive process even more expensive.

Most of the MFT advantages could be achieved while using it infrequently, perhaps every five to seven years, and even then perhaps with only a relatively inexpensive moderately small and carefully randomized sample, fine-tuned to be more representative using the precision of stratified or clustered sampling methods. In our use of the MFT at the BYU Psychology Department, I have found that our standing relative to the national norms did not shift much over two administrations of the MFT. Both administrations showed us to be at or slightly above the national norms. A well-designed internally created instrument along the lines just described could be used in the intervening years, and for the majority of our students even in the MFT-testing year, to give a broader, more finely-tuned, and less expensive assessment of our students' acquired knowledge and skills within the discipline of psychology.

References

Adam, S. (2004). *Using learning outcomes: A consideration of the nature, role, application and implications for European education of employing 'Learning Outcomes' at the local, national and international levels*. Edinburgh, Scotland: Scottish Executive.

American Psychological Association, Task force on undergraduate psychology major competencies. (2002). *Undergraduate psychology major learning goals and outcomes: A report.* Washington, DC: Author. Retrieved March 7, 2007, from www.apa.org/ed/pcue/taskforcereport.pdf

Bogue, E.G. (2000). *Exploring the heritage of American higher education: The evolution of philosophy and policy*. Phoenix, AZ: Oryx Press.

Council of Independent Colleges (CIC) (2008), *Evidence of learning: Applying the collegiate learning assessment to improve teaching and learning in the liberal arts experience*. Washington DC: Council of Independent Colleges.

Dolinsky, B., & Kelley, J.M. (2010). For better or for worse: Using an objective program assessment measure to enhance an undergraduate psychology program. *Teaching of Psychology*, *37*(4), 252 – 256.

Donald, J. G. (2002). *Learning to think: Disciplinary perspectives*. San Francisco: Jossey-Bass Publishing.

Dugan, R.E., & Hernon, P. (2002). Outcomes assessment: Not synonymous with inputs and outputs. *The Journal of Academic Librarianship*, *28*, 376-380.

Ewell, P. (2010), The US National survey of student engagement (NSSE). In D. Dill & M. Beerkens (Eds.), *Public policy for academic quality: Analyses of innovative policy instruments* (pp. 87-102). New York: Springer.

Ewell, P.T., Finney, J., & Lenth, C. (1990). Filling in the mosaic: The emerging pattern of state based assessment, *AAHE Bulletin*, *58*(4), 3-5.

Frye, R. (1999). Assessment, accountability, and student learning outcomes. *Dialogue*, *2*, 1-11. Retrieved from https://www.uky.edu/IRPE/assessment/references/dialogue.pdf

Gatson, P. (1991). Measuring the marigolds: Literary studies and the opportunity of outcomes assessment. *Journal of the Midwest Modern Language Association*, *24*(1), 11-20.

Hardison, C. M., & Vilamovska, A. (2009). *The collegiate learning assessment: Setting for performance at a college or university.* Santa Monica, CA: Rand.

Harris, K-L., & James, R. (2010). The course experience questionnaire, graduate destination survey, and learning and teaching performance fund in Australian Higher Education. In D. Dill & M. Beerkens (Eds.), *Public policy for academic quality: Analyses of innovative policy instruments* (pp. 103-124). New York: Springer.

Hussey, T., Smith, P. (2008). Learning outcomes: A conceptual analysis. *Teaching in Higher Education*, *13*, 107-115.

Kuh, G.D., & Ewell, P.T. (2010). The state of learning outcomes assessment in the United States. *Higher Education Management and Policy*, *22*, 9-28.

Lewis, R. (2010). External examiner system in the United Kingdom. In D. Dill & M. Beerkens (Eds.), *Public policy for academic quality: Analyses of innovative policy instruments* (pp. 21-38). New York: Springer.

McMurtrie, B. (1999). Accreditors revamp policies to stress student learning. *The Chronicle of Higher Education*, *12,* A29.

Mentkowski, M., & Associates (Eds.). (2000). *Learning that lasts: Integrating learning, development, and performance in college and beyond*. San Francisco, CA: Jossey-Bass.

Organization for Economic Co-operation and Development (2009), *OECD feasibility study for the International Assessment of Higher Education Learning Outcomes*. Retrieved from www.oecd.org/document/22/0,3343,en_2649_35961291_40624662_1_1_1_1,00.html.

Pace, C.R. (1979). *Measuring the outcomes of college: Fifty years of findings and recommendations for the future*. San Francisco, CA: Jossey-Bass.

Pascarella, E.T., & Terenzini, P.T. (2005). *How college affects students, Vol. 2, a third decade of research*. San Francisco, CA: Jossey-Bass.

Pusateri, T. Halonen, J. S., Hill, B., & McCarthy, M. (2009). *The assessment cyberguide for learning goals and outcomes (*2nd ed.*).* Washington DC: American Psychological Association. Retrieved from http://www.apa.ord/ed/governance/bea/assessment-cyberguide-v2.pdf

Spangehl, S. (1987). The push to assess, *Change*, *19*, 35-39.

Stoloff, M.L., Feeney, K.J. (2002). The Major Field Test as an assessment tool for an undergraduate psychology program. *School of Psychology, 29*(2), 92-98.

Tartwijk (van), J., Driessen, E., Vleuten (van der), C. (2007). Factors influencing the successful introduction of portfolios, *Quality in Higher Education*, *13*(1), 69-80.

Thrash, P.A. (1990). State agency expectations for assessment in the north central region. *NCA Quarterly, 65*(2), 415-434.

U.S. Department of Education (2006). *A test of leadership: Charting the future of U.S. higher education*. Washington, DC: Author.

Weerts, D., & Vidal, J. (2005). Enhancing alumni research: European and American perspectives. *New Directions for Institutional Research*, *126*, 1-112.

Appendix A

Major Field Test Description

The content distribution is as follows:

1) Experimental or Natural Science Oriented Areas (about 40% of the questions)
   A. Learning, Cognition, and Perception (about 24% of the questions)
      1. Learning (6-8%)
         - Classical conditioning
         - Operant conditioning
         - Knowledge acquisition
         - Social learning
         - Biological constraints
         - Theories and issues
      2. Language (1-3%)
         - Structure
         - Speech perception and processing
         - Communication
         - Disorders
         - Theories and issues
      3. Memory (5-7%)
         - Levels of processing
         - Types of memory phenomena
         - Encoding strategies and failures
         - Retrieval strategies and failures
         - Semantic organization
         - Theories and issues
      4. Cognition (4-6%)
         - Representation
         - Information processing
         - Problem solving
         - Reasoning
         - Metacognition
         - Theories and issues
      5. Perception (3-5%)
         - Psychophysics, Signal detection
         - Attention
         - Perceptual symptoms and organization
         - Theories and issues
   B. Comparative and Evolutionary (about 3% of the questions)
         - Instinct, genetics, learning, adaptation
         - Aggression and Dominance
         - Attachment, Sociality, Altruism
         - Sexual behavior
         - Parenting behavior
         - Evolutionary psychology
         - Theories and issues
   C. Sensation and Physiology (about 13 % of the questions)
         - Neurons and neural communication
         - Sensory structures and functions
         - Motor structures and functions
         - Central and peripheral nervous system
         - States of consciousness
         - Psychopharmacology
         - Hormonal factors
         - Neurophysiological models (e.g., memory, motivation, arousal, emotion)
         - Theories and issues
2) Social or Social Science Oriented Areas (about 41% of the questions)
   A. Clinical and Abnormal (about 10% of the questions)

- Types of disorders
- Biological factors
- Psychological factors
- Sociocultural factors
- Diagnostic systems
- Treatment of disorders
- Prevention
- Theories and issues

B. Developmental (about 12% of the questions)

- Nature-Nurture
- Behavioral genetics
- Motor, Sensory, Perceptual
- Attention, Cognition, Memory
- Language
- Learning, Intelligence
- Social, Personality, Emotion
- Socialization influences
- Cultural influences
- Periods of development (e.g., infancy, adolescence, adulthood)
- Theories and issues

C. Personality (about 7% of the questions)

- Behavioral approaches
- Phenomenological approaches
- Psychodynamic approaches
- Social cognitive approaches
- Trait approaches
- Assessment
- Theories and issues

D. Social (about 11% of the questions)

- Social perception, cognition, attribution, beliefs
- Attitudes and behavior
- Self
- Social influence and persuasion
- Interpersonal attraction
- Group processes
- Cultural influences
- Theories and issues

3) Other Areas (about 21% of the questions)

A. Historical (about 3% of the questions)

B. Applied (about 3% of the questions)

- Industrial-Organizational and Human factors
- Educational
- Applied, Public policy
- Health psychology

C. Measurement and Methodology (about 15% of the questions)

- Measurement, scales, tests
- Research designs
- Statistics
- Interpretations of findings
- Ethics

Psychology MFT scores are reported as follows:

Total Score

Reported for each student and summarized for the group.

Subareas

Reported for each student and summarized for the group.

- Learning and Cognition (including Language, Memory, and Thinking) (27)
- Perception, Sensory, Physiology, Comparative, and Evolutionary (26)
- Clinical, Abnormal, and Personality (26)
- Developmental and Social (29)

Assessment Indicators

Reported for the group* only.

- Memory and Thinking (15)
- Sensory and Physiology (18)
- Developmental (15)
- Clinical and Abnormal (14)
- Social (14)
- Measurement and Methodology (19)

Note: Numbers in parentheses are approximate number of questions in each category.

Appendix B

Course Groupings

A) Developmental and Clinical
Psych 341
Psych 342
Psych 320
Psych 321
Psych 348
Psych 311
Psych 343
Psych 322

B) Social
Psych 350
Psych 353
Psych 311
Psych 338
Psych 354
Psych 352
Psych 356
Psych 330
Psych 358

C) Behavior, Brain, and Cognition
Psych 375
Psych 370
Psych 381
Psych 365
Psych 361
Psych 387

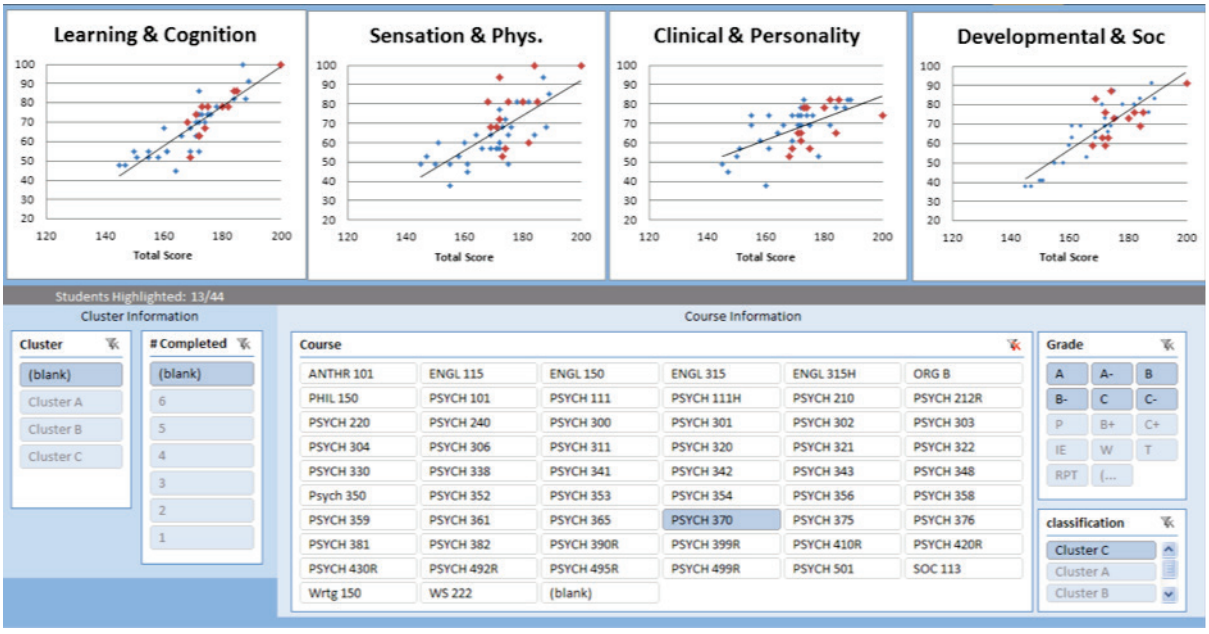
D) Methods
Psych 301
Psych 302
Psych 304

E) Other Requirements
Psych 101
Psych 111
Psych 210

Appendix C

Figures from an Initial Tool Created for this Dissertation, Illustrating Possible New Directions for Development in the Analysis of Predictors of Learning Outcomes

Initial image for excel spreadsheet containing random MFT scores. Users will have the ability to highlight students based on courses, clusters, and grades.



Note: The setup shown in this image highlights all students who took Psych 370. The red dots in the "Sensation & Phys." graph show that most of the 13 students performed higher on this MFT subarea. Further selection could specify course grade or number of Cluster C courses completed, and so forth.