International Journal of Instruction *July 2021 ● Vol.14, No.3*
e-ISSN: 1308-1470 ● www.e-iji.net *p-ISSN: 1694-609X*
pp. 629-644

Article submission code: 20200802225525
Received: 02/08/2020
Revision: 10/01/2021
Accepted: 03/02/2021
OnlineFirst: 10/06/2021

Analysing Student Performance on the Major Field Test in Business at a Canadian University

Ron Messer
School of Business, Kwantlen Polytechnic University, Canada, *ron.messer@kpu.ca*

The Major Field Test in Business (MFTB) is a nationally administered student evaluation that measures the accumulated knowledge of undergraduates enrolled in a four-year degree program. Research to date has focused primarily on understanding how different variables correlate with performance on this standardized test -- such as student grades, gender and ethnicity. The research objectives of this essay are to analyse the ambiguous results found in previous studies and to highlight how interaction effects among variables can be used to better explain test success. Using data on student performance collected over 11 semesters at a Canadian university, this essay uses a multi-variable regression model to understand the factors affecting scores on the MFTB. The model results suggest that examining the interaction between variables provides important insights and can help to better explain the ambiguity in prior studies. This study is unique in that it uses a statistical measure known as the extra sum of squares F-test to demonstrate the significance of interaction variables.

Keywords: virtual simulation, web-based application, multimedia, teaching, learning

INTRODUCTION

The Major Field Test for Business (MFTB) is a national exam that measures the accumulated knowledge of undergraduate business students enrolled in a four-year degree program. It covers several subject areas and is administered by the Educational Testing Services. There is no preparation required to complete the MFTB, other than that a student should be in their final semester of studies.

The test is comprised of 120 multiple choice questions, with approximately half requiring numeric aptitude. It is administered on-line in two, sixty-minute modules. The test components by subject area -- including their relative weighting and whether numeric skills are required -- are shown in table 1. [1]

Citation: Messer, R. (2021). Analysing student performance on the major field test in business at a Canadian University. *International Journal of Instruction*, 14(3), 629-644. https://doi.org/10.29333/iji.2021.14337a

Table 1
MFTB subject area, weighting and numeric evaluation

Subject Area	Weighting	Numeric Skills *
Accounting	15%	Yes
Economics	13%	Yes
Management	15%	No
Quantitative Analysis	11%	Yes
Information systems	10%	No
Finance	13%	Yes
Marketing	13%	No
Legal Environment	10%	No
International Issues	Overlaps other subjects	N/A
TOTAL	100%	Numeric = 52%

* Subjects that require numeric skills involve calculations and problem-solving ability.

I have been a faculty member at a Canadian university for about twelve years. Established in 1981, the university has four campuses and offers bachelor's degrees, associate degrees, diplomas and certificates in more than 120 programs.[2] The School of Business is among the largest such schools in Canada. It consists of over 150 faculty members and 3,500 full-time students.

I teach an advanced level course where students apply management accounting concepts to multi-subject cases and business projects. The class is a ‘capstone’ course, since it requires students to use ‘higher level’ cognitive skills, including those of synthesis and integration. The course can only be taken by 4th year accounting majors. As part of their course grade students are required to complete the MFTB, which provides our school with an important measure of how much graduates have absorbed during their undergraduate studies. The test provides evidence of assurance of learning.

In attempting to advance our understanding of what affects MFTB performance, researchers have generally adopted three approaches: (i) increased the sample sizes of their studies, (ii) analyzed more than a single post secondary institution and/or (iii) added more explanatory variables, to an ever-increasing list of factors that could influence test scores. This essay takes a different approach by refining the analysis of some of the more common variables examined in the past.

The research question being addressed is how to explain the ambiguous results in prior studies concerning the variables affecting student performance on the MFTB. The objective of this essay is to better understand these results by analysing the interaction effects between variables; this will be done by using a multivariable regression model, whose validity is confirmed with the extra sum of squares F test.

Literature Review

The literature on the variables affecting MFTB scores is extensive, inconclusive and sometimes contradictory; this is particularly true for test performance and its relation to grades, gender and ethnicity.

Allen & Bycio (1997) showed a positive correlation between MFTB results and students' GPA scores (including cumulative GPA, overall business GPA and business major GPA), as well as SAT-V and SAT-M assessments. This same positive correlation was found by Bycio & Allen (2007), Mason et al. (2011) and Rook & Tanyel (2009).

Some investigations have shown a correlation between gender and performance on the MFTB, where males had better results than females. Other studies have shown the opposite, where females outperformed males. Bean & Bernardi (2002) provided evidence that gender had a significant influence on student MFTB scores. This conclusion was supported by the work done by Chowdhury, et al. (2013). This positive performance relationship for males has also been shown in studies by Bagamery (2005); Bean & Bernadardi (2002); Black & Duhon (2003); Contreras et al. (2011); Mason et al. (2011); Settlage & Settlage (2011). (Refer to table 14.) However, Allen & Bycio (1997) did not find any gender-based differences in MFTB performance.

Ethnicity has been investigated in several studies and shown to have a significant relationship with MFTB scores. Mason, et al. (2011) found a relationship between ethnicity (White, Black, Hispanic and American Indian) and MFTB performance.

Some of the research has shown moderate to strong MFTB correlations with both gender and ethnicity, such as Chowdhury, et al. (2013). However, the relationship between these variables has not been consistent. Of the seventeen studies reviewed by Green et al. (2014), nine showed gender as a significant variable, while five found it insignificant. Two studies showed ethnicity as a significant factor, while five found it insignificant. [4, 5]

Based on a comprehensive review of the literature, Green, Stone & Zegeye (2014) concluded that MFTB scores are "significantly influenced by specific student characteristics. Consequently, the use of these scores for assessment required detailed analysis of these characteristics" (Green, Stone & Zegeye, 2014, p.22). This essay undertakes such a detailed analysis by examining the importance of the interaction effects between gender and ethnicity in explaining student performance on the MFTB. [6]

METHOD

Most previous studies have looked at the relationship between MFTB performance and students' grades, gender and ethnicity. [3] This research has primarily focused on the significance of these variables in isolation, but not in combination. In this study, student MFTB performance is analyzed through: (i) descriptive statistics; (ii) correlation analysis; (iii) regression models. The study shows how interactions between variables provides a better understanding of factors affecting test scores. This is done by constructing two multi-variable regressions -- where the first is a simpler version of the second -- and demonstrating, through the statistical measure know as the extra sum of squares F-test, the explanatory power of interaction effects. Demonstrating the significance of interaction variables through this statistical technique has not been done in previous studies.

Data Set and Variables

The study examines a dataset of 307 accounting students enrolled in a fourth-year capstone course who wrote the MFTB. The dataset includes students who were taught over a period of 4 academic years, comprised of 11 sections and covering 9 semesters. Tables 2, 3 and 4 show the composition of the students in this study based on course grades, gender and ethnicity.

Table 2
Dataset of students by course grade (n = 307)

Student Course Grade*	Proportion of Total
Less than 50 (F)	3%
50 to 59 (D)	8%
60 to 69 (C)	42%
70 to 79 (B)	40%
80 to 99 (A)	7%

**Out of 100 excluding MFTB results*

Table 3
Dataset of students by gender (n = 307)

Student Gender	Proportion of Total
Male	50%
Female	50%

Table 4
Dataset of students by ethnicity (n = 307)

Student Ethnicity	Proportion of Total
White	24%
Chinese	31%
Indian	35%
Other	10%

The correlations between course grade, gender, ethnicity and performance on the MFTB are shown in table 5, which summarizes the Pearson correlation co-efficient (r) for these variables.

Table 5
Variables and correlation (r) with MFTB Results (n = 307)

Variables	Correlation (r)
Grades	+.332
Gender = Male	+.330
Gender = Female	-.330
Ethnicity = White	+.306
Ethnicity = Chinese	-.210
Ethnicity = Indian	-.063
Ethnicity = Other	-.012

Note: ‘+’ indicates a positive correlation; ‘- ‘indicates a negative correlation with MFTB scores.

The results show moderate correlations (i.e., r greater than .3 but less than .6) based on grade and gender. The table also shows both moderate and weak correlations (i.e., r less than .3) based on ethnicity. [7]

Grades Variable:

Two grading measures were used:

i. The final course grade, which is based on 100 marks, exclusive of the MFTB results.

ii. The MFTB percentile rank based on the students' performance. This measure uses student (scaled) scores, which are converted to percentile rank. [8]

Table 6 shows summary statistics for the final course grades and the results of the MFTB percentile rank for the students in this study. (Note that both the final course grade and the MFTB percentile rank use a range, from 0 – 100.)

Table 6
Final course grade and MFTB percentile rank

Measure (n = 307)	Final Course Grade (max = 100)	MFTB Percentile Rank (max = 100)
Mean	68	70
Standard Deviation	9	24

A comparison of the frequency distribution of marks for the final course grade and the MFTB percentile rank shows markedly different distributions. While the final course grade is normally distributed (see figure 1) that of the MFTB appears to be somewhat bi-model and skewed (see figure 2).

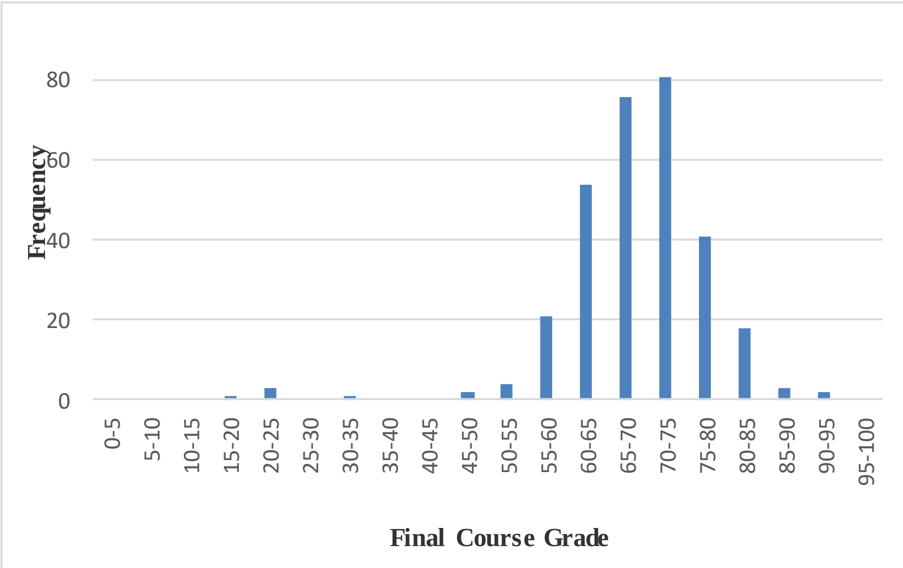

Figure 1
Course grade frequency distribution (n = 307)

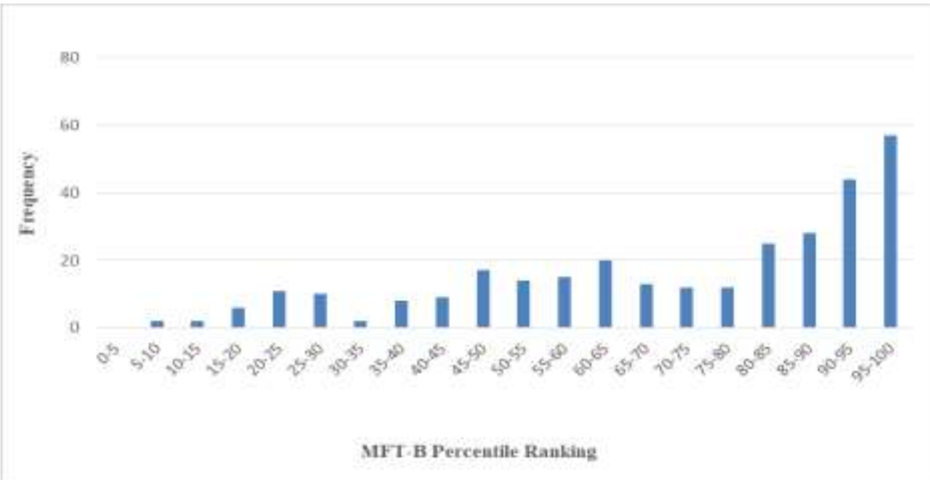

Figure 2
MFTB percentile rank frequency distribution (n = 307)

The course grade distribution has a relatively narrow range but a common pattern, with a mean of approximately 75% (representing a letter-grade of B), which is considered normal for this senior level class. The distribution of the MFTB percentile rank has two means: one approximating the 50^{th} percentile (which would be expected) and the other showing results that are skewed towards the upper end of the ranks. This unusual feature suggests two populations of students, one of average ability and the other in the superior range. (The frequency distribution, by percentile rank, for all students who wrote the MFTB between September 2013 and June 2015 shows a 'classic' normal distribution with a 50% mean for all writers. [9])

Gender Variable:

The correlation with gender in table 5 indicates that males performed better on the MFTB than females, which has been observed by other studies. The male correlation coefficient is +0.330, while the female is -0.330.

Table 7 shows the average course grades and MFTB rank, along with standard deviations (Std-dev), by gender. The table illustrates that females did not perform as well on the MFTB as males, although the final course grades for both males and females were almost identical.

Table 7
Final course grade and MFTB percentile rank by gender

Gender	Final Course Grade		MFTB Percentile Rank	
	Mean	Std-dev	Mean	Std-dev
Female	68	8	63	26
Male	69	10	79	20
Total (n = 307)	69	9	71	25

The standard deviation is more notable for the MFTB percentile rank, likely reflecting the two populations of students observed in figure 2.

Ethnicity Variable:

The only significant relationship between ethnicity and MFTB performance was for White students, who showed a positive correlation of .306. There was little correlation between MFTB performance and other ethnic groups.

The university where my course is taught has both domestic and foreign students. Among the domestic students, many are first generation offspring of recent immigrants. Domestic students originate from across Canada. The ethnic groups investigated have been classified as White, Chinese, (East) Indian and Other.

The Chinese students are primarily from mainland China and Hong Kong. Included in this group are foreign students, recent immigrants, first generation children and long-term residents. The East Indian students are primarily first-generation children of immigrant parents. The population of 'Other' students is comprised of groups from the Middle East, Latin America and the Philippines. Table 8 shows student performance by ethnicity.

Table 8
Final course grade and MFTB percentile rank by ethnicity

Ethnicity	Final Course Grade		MFTB Percentile Rank	
	Mean	Std-dev	Mean	Std-dev
Chinese	67	9	63	28
Indian	68	10	69	22
Other	70	7	71	23
White	71	9	84	18
Total (n = 307)	69	9	71	25

All ethnic groups, other than Chinese, showed somewhat better performance on the MFTB than the final course grade. There is a significant difference in performance on the MFTB between Chinese and White students. There was also less variation in MFTB scores for White compared to Chinese students.

FINDINGS AND DISCUSSION

Multi-Variable Regression Models

As noted, previous studies have been inconclusive concerning the impact of gender and ethnicity as explanatory variables for MFTB outcomes. To better understand this anomaly, two regression analyses were completed. Regression 1 (R1 model) analysed the causative relationships between MFTB results as the dependent variable and category variables for gender, ethnicity, and course grades as independent variables (see table 9). Regression 2 (R2 model) analysed the causative relationship between MFTB scores and interaction variables for gender and ethnicity, as well as course grades (see table 10).

Table 9
R1 model independent variables

Independent Variables	Category/Value*	Type
Gender	Male	Category variable
	Female	Category variable
Ethnicity	White	Category variable
	Chinese	Category variable
	Indian	Category variable
	Other	Category variable
Course Grade	0 – 100	Numerical variable

**'Male' was the reference variable used for gender and 'White' the reference variable used for ethnicity.*

Table 10
R2 model independent variables

Independent Variables	Interaction/Value	Type
White-male	Ethnicity and gender	Interaction variable
Chinese-male	Ethnicity and gender	Interaction variable
Indian-male	Ethnicity and gender	Interaction variable
Other-male	Ethnicity and gender	Interaction variable
White-female	Ethnicity and gender	Interaction variable
Chinese-female	Ethnicity and gender	Interaction variable
Indian-female	Ethnicity and gender	Interaction variable
Other-female	Ethnicity and gender	Interaction variable
Course Grade	0 – 100	Numerical variable

R1 Model Category Variables:

The R1 model was optimized [10] so that only those variables that provided additional explanatory power were included, as shown in table 11. These variables were significant at a 95% confidence level. This means that they had a high multiple R, along with a p-value < .05 (i.e., a 5% risk tolerance for a type 1 error). [11]

Table 11
R1 model results

Variables*	Coefficients	Standard Error	t Stat	P-value
Chinese	-16.8209935	3.326128579	-5.057228876	< 0.0001
Indian	-13.79893893	3.204255544	-4.306441462	< 0.0001
Other	-17.32501585	4.499417009	-3.850502368	0.0001
Female	-15.452625	2.448014803	-6.312308641	< 0.0001
Final Grade	0.741255205	0.131541734	5.635133297	< 0.0001

The R1 model shows that Chinese, Indian and Other ethnic groups did not perform as well as White students (the reference variable) on the MFTB, as indicated by the negative coefficients. In addition, females, relative to males (the reference variable), did not perform as well on the MFTB (as noted in previous studies). The ANOVA table for R1 is shown in table 12.

Table 12
R1 model ANOVA table

ANOVA	df	SS	MS	F	Significance F
Regression	5	52302.0417	10460.40834	23.66979972	< 0.0001
Residual	301	133021.1049	441.930581		

R2 Model Interaction Variables:

The R2 model was optimized, including only variables that provided additional explanatory power, as shown in table 13. These variables were significant at a 95% confidence level. This means that they had a high multiple R, along with a p-value < .05 (i.e., a 5% risk tolerance for a type 1 error).

Table 13
R2 model results

Variables	Coefficients	Standard Error	t Stat	P-value
Final Grade	0.694078915	0.128324194	5.408792318	< 0.0001
Chinese-Female	-25.16086933	3.464184019	-7.263144564	< 0.0001
Indian-Female	-11.25753739	3.518515408	-3.199513456	0.0015
Other-Female	-27.9747936	7.536512279	-3.711901815	0.0002
Indian-Male	-6.754485711	3.440800914	-1.963056242	0.0506
White-Male	11.05509156	3.958561503	2.792704258	0.0056

The R2 model shows the increased explanatory power of using interaction variables, where ethnicity combined with gender indicated negative impacts for Chinese, Indian and Other females, as well as Indian males (negative coefficients). In addition, the model showed positive effects for white males (positive coefficients). The ANOVA table for R2 is shown in table 14. Table 15 compares the results from the two regression models.

Table 14
R2 model ANOVA table

ANOVA	df	SS	MS	F	Significance F
Regression	6	59328.52146	9888.08691	23.54406841	< 0.0001
Residual	300	125994.6251	419.9820837		

Table 15
Comparison of regression results

Regression Statistics	R1: Category Variables	R2: Interaction Variables
Multiple R	0.531	0.566
R Square	0.282	0.320
Adjusted R Square	0.270	0.307
Standard Error	21.022	20.494

The statistical significance of gender and ethnicity increased when the variables were combined through interaction effects. The co-efficient of determination (r^2) increased to .320 (a change of 0.038), which was statically significant at a 95% level of confidence. Without using interaction effects, the r^2 in the regression (R1) was only 0.282.

Extra Sum of Squares F-Test

To determine whether R2 is better at explaining student performance than R1, the F-statistic is useful for comparing the two models. The extra sum-of-squares F-test compares the relative fits of R1 and R2. [12] If the R2 model is correct, the relative increase in the residual sum of squares (RSS) -- i.e., moving from the more complicated interaction variable model to the simpler category variable model -- is expected to equal the relative increase in residual degrees of freedom (RDF). In other words, if the simpler R1 model is correct it is expected that:

(Residual Sum of Squares R1 – Residual Sum of Squares R2)

÷ Residual Sum of Squares R2

is approximately equal to:

(Residual Degrees of Freedom R1 – Residual Degrees of Freedom R2)

÷ Residual Degrees of Freedom R2

(Note that the number of degrees of freedom is equal to the number of data points, less the number of variables. The more complicated model has more variables and therefore fewer degrees of freedom.)

If the R2 model is correct, it is expected that the relative increase in the residual sum-of-squares (proceeding from the R2 to the R1 model) will be greater than the relative increase in the residual degrees of freedom. In other words:

(RSSR1 – RSSR2) ÷ RSSR2 > (RDFR1 – RDFR2) ÷ RDFR2

The F ratio quantifies the relationship between the *relative* increase in the sum-of-squares and the *relative* increase in the degrees of freedom.

$$F = \frac{(RSSR1 - RSSR2) \div RSSR2}{(RDFR1 - RDFR2) \div RDFR2}$$

If the R1 model is correct, an F ratio near 1.0 is expected. If the ratio is much greater than 1.0, there are two possibilities:

i. The more complicated R2 model is correct.

ii. The R1 model is correct, but randomness led the more complicated model to fit better.

The p-value answers the question about whether R1 is correct by providing information concerning the chance that the dataset fits R2 better. If the p-value is low, we conclude that R2 is significantly better than R1. Otherwise, we conclude that there is no compelling evidence supporting R2, and so we accept the simpler model, R1.

The F-ratio for R2 using the extra sum-of-squares F-test is 16.73, which is significantly greater than 1.0 and therefore suggests a better explanation for student performance than that provided by R1. The p-value for R2 = 0.00001, which provides assurance about the validity of the model. In summary, R2 which uses interaction effects, provides a better explanation of student performance than does R1, which uses category variables for gender and ethnicity in isolation.

The results of this research are significant in two respects. *First*, the comparison of the two regression models illustrates how the ambiguity in previous studies concerning the effects of gender and ethnicity on MFTB success may be better understood by analysing the interaction between these variables. The extra sum of squares F-test validates the significance of such an approach. *Second*, in broader application, the research suggests that other interaction effects can be studied to better appreciate the variables that, in combination, may affect success on the MFTB and which can also be validated using the extra sum of squares F-test.

CONCLUSIONS AND LIMITATIONS

This study shows that an analysis of interaction effects between explanatory variables, such as gender and ethnicity, can improve our understanding of factors affecting student performance on the MFTB. It is suggested that the ambiguous results of prior research about the effects of gender and ethnicity can be explained by such interaction effects, whereby student performance on the MFTB can be accounted for based on *both* gender and ethnicity. If MFTB scores are "significantly influenced by specific student characteristics" (Green et al., 2014, p.22), then it is necessary to examine the effects of causative variables in combination and not only in isolation.

More broadly, this study suggests that increased insight into student performance on the MFTB can be obtained by examining interaction effects using other multi-variable regression models. While this analysis has focused on the interaction of gender and ethnicity, different combinations could prove useful -- for example: ethnicity and GPA; business major and gender; age and gender. The possibilities are both intriguing and warrant further exploration. Here, additional research can provide practical contributions towards identifying those interaction effects that determine student achievement on standardized tests, like the MFTB.

In this study, further investigation is necessary to determine why Chinese females did not perform well on the MFTB. Possible explanations for these observed deficiencies could include language barriers, inadequate prior preparation, motivational issues or testing bias. The significantly positive results of white males noted here as well as other studies should also be further analysed, particularly in view of the large differences noted in comparison to their peer group, primarily in U.S. schools. [13] This investigation is limited in the following ways: (i) in common with many other studies, it examined only a single course at one university; (ii) also, in common with similar research, it was limited by the sample size; (iii) some of the MFTB correlations were found to be weak; (iv) the concept of 'ethnicity' does not distinguish between domestic enrolments, permanent residents, new immigrants and foreign students.

REFERENCES

Allen, J. S., & Bycio, P. (1997). An evaluation of the educational testing service major field achievement test in business. *Journal of Accounting Education*, *15*(4), 503-514.

Bacon, D. R., & Stewart, K. A. (2017). Why assessment will never work at many business schools: A call for better utilization of pedagogical research. *Journal of Management Education*, *41*(2), 181-200.

Bagamery, B. D., Lasik, J. J., & Nixon, D. R. (2005). Determinants of success on the ETS Business Major Field Exam for students in an undergraduate multisite regional university business program. *Journal of Education for Business*, *81*(1), 55-63.

Bean, D. F., & Bernardi, R. A. (2002). Performance on the Major Field Test in Business: The explanatory power of SAT scores and gender. *Journal of Private Enterprise*, *17*(2), 172-178.

Black, H. T., & Duhon, D. L. (2003). Evaluating and improving student achievement in business programs: The effective use of standardized assessment tests. *Journal of Education for Business*, *79*(2), 90-98.

Blackford, B., & Shi, T. (2015). The relationship between business simulations in capstone management courses and standardized test scores. *The International Journal of Management Education*, *13*(1), 84-94.

Bush, H. F., Duncan, F. H., Sexton, E. A., & West, C. T. (2008). Using the Major Field Test-Business as an assessment tool and impetus for program improvement: Fifteen years of experience at Virginia Military Institute. *Journal of College Teaching & Learning (TLC)*, *5*(2).

Chowdhury, M. I., & Wheeling, B. (2013). Determinants of Major Field Test (MFT) score for graduating seniors of a business school in a small mid-western university. *Academy of Educational Leadership Journal*, *17*(1), 59-71.

Fairchild, C., & Hahn, W. (2020). Accounting and finance majors outperform other majors on the major field test in business and the Comprehensive Business Exam: An analysis of exam performance drivers. *Journal of Education for Business*, *95*(6), 345-350.

Green, J. J., Stone, C. C., & Zegeye, A. (2014). The Major Field Test in Business: A solution to the problem of assurance of learning assessment?. *Journal of Education for Business*, *89*(1), 20-26.

Ling, G. (2012). Why the major field test in business does not report subscores: Reliability and construct validity evidence. *ETS Research Report Series*, *2012*(1), i-26.

Ling, G., Bochenek, J., & Burkander, K. (2015). Using the Major Field Test for a bachelor's degree in business as a learning outcomes assessment: Evidence from a review of 20 years of institution-based research. *Journal of Education for Business*, *90*(8), 435-442.

Mason, P. M., Coleman, B. J., Steagall, J. W., Gallo, A. A., & Fabritius, M. M. (2011). The use of the ETS Major Field Test for assurance of business content learning: Assurance of waste?. *Journal of Education for Business*, *86*(2), 71-77.

McLaughlin, J. P., & White, J. T. (2007). Major Field Achievement Test In Business-Guidelines For Improved Outcome Scores-Part I. *College Teaching Methods & Styles Journal (CTMS)*, *3*(2), 11-20.

Mirchandani, D., Lynch, R., & Hamilton, D. (2001). Using the ETS major field test in business: Implications for assessment. *Journal of Education for Business*, *77*(1), 51-56.

Rook, S. P., & Tanyel, F. I. (2009). Value-added assessment using the Major Field Test in Business. *Academy of Educational Leadership Journal*, *13*(3), 87.

Settlage, D. M., & R. Wollscheid, J. (2015). Analyzing student performance in specific subject area indicators on the ETS major field test in business. *Journal of Education for Business*, *90*(6), 322-327.

Simmons, S. A., Jones Jr, W. M., & Bolt, C. E. (2015). The major field test in business: A direct measure of learning in common business disciplines. *Journal of Education for Business*, *90*(2), 57-62.

Strang, K. D. (2014). Cognitive learning strategy as a partial effect on major field test in business results. *Journal of Education for Business*, *89*(3), 142-148.

Terry, N., Mills, L., Rosa, D., & Sollosy, M. (2009). Do online students make the grade on the business major field ETS exam?. *Academy of Educational Leadership Journal*, *13*(4), 109.

Terry, N., Walker, J., & Kelley, G. (2010). The Determinants of Student Performance on the Business Major Field ETS Exam: Do Community College Transfer Students Make the Grade?. *The Journal of Human Resource and Adult Learning*, *6*(1), 1.

Terry, N., Mills, L., & Sollosy, M. (2008). Student grade motivation as a determinant of performance on the business Major Field ETS exam. *Journal of College Teaching & Learning (TLC)*, *5*(7).

Thornton, B., & Arbogast, G. (2012). Analyzing Educational Testing Service Graduate Major Field Test Results. *American Journal of Business Education (AJBE)*, *5*(5), 531-538.

Word, W. R., Rook, S. P., & Lancaster, L. M. (2011). AN ASSESSMENT NOTE ON THE UTILIZATION OF THE MAJOR FIELD TEST IN BUSINESS. *B> Quest*.

Yaman, D. (2015). Assessment with ETS Exams. *International Proceedings of Economics Development and Research*, *83*, 30.

FOOTNOTES

[1] See here the ETS, MFT website: https://www.ets.org/s/mft/pdf/mft_testdesc_business.pdf

[2] For more background and additional information, see the Kwantlen Polytechnic University website: http://www.kpu.ca.

[3] The variable 'gender' is more correctly described as 'sex', which refers to biological characteristics using binary categories for male and female. Notably, gender is a more fluid concept. The variable for 'ethnicity' is also more correctly called 'race', which refers to physical characteristics (skin colour; hair texture); whereas ethnicity is linked to cultural expressions and identification. However, because previous studies have used the terms gender and ethnicity to categorize MFTB performance for males/females and various racial groups, they will continue to be used in this study for purposes of comparability.

[4] See here Bean & Bernardi, 2002, p. 174, where they state,

"Examinations of the effect of gender on performance report mixed results. While Doran, et al. (1991) find that males outperform female students, Tyson (1989) finds that females outperform male students. Indeed, research finds that gender does not influence performance on examinations (Fogarty, et al., 1998)."

[5] See here table 3 in Green, et al. (2014), p.23 where the authors review the results of previous studies that analyse success on the MFTB using several determinants for student performance. The table illustrates the conflicting results that arise when gender and ethnicity are used as correlates for success on the MFTB. Based on differing results across studies, it is likely that the apparently contradictory correlations result from factors unique to the institutions where the research was undertaken.

[6] Some studies have examined MFTB correlations with specific groups of students, such as MBAs (Thornton & Arbogast, 2012), or disciplines within the business program, such as accounting, marketing, finance, etc. (Mclaughlin & White, 2007; Settlage & Wollscheid, 2015; Fairchild & Hahn, 2019), finding associations between these variables and MFTB performance (Allen & Bycio, 1997). Other significant correlations with MFTB results were noted for foreign students (Terry et al., 2009, 2010, 2011), age (Bagamery et al., 2005), whether the MFTB provides course credit incentives (Bielinsk-Kwapisz et al., 2012), student learning strategies (Strang, 2014), program format (Thornton & Arbogast, 2012), student quality, program curriculum and course grading policy (Word & Rook, 2012). See also Simmons et al. (2015).

[7] In comparison, Blackford & Shi (2015) found an MFTB correlation factor (i.e., r) of .601 for final course grade and an association of -0.129 based on gender.

[8] The MFTB results are recorded initially as a raw score, which is then converted to a scaled score based on a decile rank. ETS does not report on the raw scores and has indicated that the scaled scores are an appropriate and valid measure of student

performance. In this regard see Ling (2012, 2015). See also Bagamery, Lasik & Nixon (2005); Black & Duhon (2003); Mirchandani, Lynch & Hamilton (2001) regarding test validity. For other methodological issues relating to ETS standardized tests, see Yaman (2015).

[9] See the following URL for details on the frequency distribution of MFTB test results: https://www.ets.org/s/mft/pdf/acdg_business.pdf.

[10] The models were optimized by using a backward stepwise regression technique. Backward stepwise regression is an approach that begins with all of the explanatory variables and at each step gradually eliminates some of them from the regression model to find a model that best explains the data (also known as backward elimination regression). It does this by examining the change in the r^2 as each variable is dropped, to assess whether it is significant in terms of its causative effects.

[11] Stat-tools, a statistical analysis software package, was used to perform the linear regression analysis in this study for the R1 model (category variables) and R2 model (interaction variables).

[12] Regarding the extra sum-of-squares F-test, see the following URLs: https://www.graphpad.com/guides/prism/7/curve-fitting/reg_howtheftestworks.htm?toc=0&printWindow

https://en.wikipedia.org/wiki/F-test

[13] Studies analysing the MFTB have been used primarily to determine: (i) program/course effectiveness -- i.e., measure outcomes (Compare Bacon et al., 2016); (ii) MFTB correlations -- i.e., identify key factors affecting program/course outcomes, such as business GPA and program major; (iii) a call for action -- i.e., initiate program/course changes required if results indicate that this is necessary (See here Bush et al., 2008; McLaughlin & White, 2007).