International Journal of
Mathematics And its Applications
ISSN: 2347-1557

Int. J. Math. And Appl., **11(3)**(2023), 1–10
Available Online: http://ijmaa.in

Guiding Students Towards Success in Calculus via a Focused Assessment Approach

Samia Challal[1,*]

[1]*Department of Mathematics, Glendon College, York University, Toronto, Canada*

Abstract

The purpose of this study is to ascertain the efficiency of a focused assessments approach in improving student experience in calculus.

Keywords: Focussed Assessment, Parametric inference, Nonparametric inference.

2020 Mathematics Subject Classification: 97D60, 97-11, 62Fxx, 62Gxx.

1. Introduction

A Focused assessment (FA) is very important in Healthcare. It allows for nurses to quickly assess a patient by spending their time focused on the problem area. By analogy, an FA approach in teaching consists on assessment of specific skills or learning objectives. For example, a motivational test (MT) assigned to students early during a calculus course has for focus to motivate students to take seriously their review of mathematics basics. A significant improvement in calculus scores compared to previous years where an MT was not administered, would mean that an MT can be adopted as an effective strategy for supporting students. Labs (L) and activities involving modelling (M) are also explored in this study with the objectives to assess the acquired knowledge through the use of a software and the application of calculus to real-world problems. The absence of a significant improvement would mean that the said assessment(s) ought to be revised to reach the specified objectives. Often, the FA approach is adopted instinctively. The purpose here is to discover new ways of thinking about it. First, we describe how we gathered the data, then we comment each finding and relate it to research in the area of education.

2. Method

Research Model: The study uses data from the period 2014/2016 and 2018/2022. It is quantitative and allows some interpretations. The data analysis is performed in R software [3,6] using several statistical methods [7–9].

*Corresponding author (schallal@glendon.yorku.ca)

Participants: Data was collected from undergrade students' grades in first year Differential (respectively Integral) Calculus course in the same University and taught by the same instructor. This ensures less variability in the gradings through different samples. Students were not informed about the intent of the study since the grades were part of their assessment. A third party anonymized the data.

The data received ethics review and approval by the York University's Human Participants Review Sub-Committee (Certificate #: 2022-020).

The data includes columns of scores related to specific Focussed Assessments "FA". Five types of scores are targeted: W: score on concepts, MT: motivational test score, L: score on Labs that require the use of a software, M: score on modelling, E: average of exams' scores. Four periods are distinguished depending on the FAs involved in each period. More precisely, we will have the periods: (W, E) (period 1), (W, MT, E) (period 2), (W, MT, L, E) (period 3), and (W, MT, L, M, E) (period 4) as there was a progression in assigning the different types of assessments.

3. Period W-E

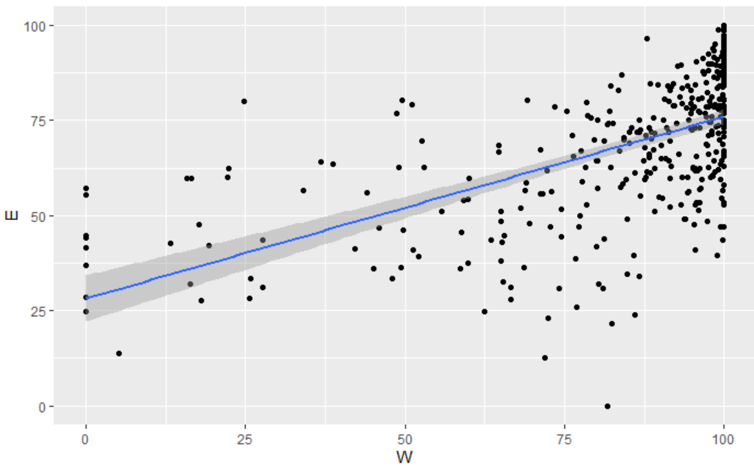

Figure 1: W-E correlation

"W" activities are formulated as quizzes from WebAssign; an online learning platform built by educators based on the book [11]. For each section, diverse questions are selected from checking a definition, applying a theorem, guided calculations, and independent solving problems. Students benefit from Webassign support by practicing other quizzes, watching videos, and reviewing the class material. They were supported also during office hours and via chat, forum, and Discord platform.

The "W" quizzes are intended to enhance students' content "C" knowledge; a terminology adopted by the authors in [10]. This content knowledge cannot be seen in isolation. It goes with the pedagogy "P" adopted. The zone of intersection "PCK represents the blending of content and pedagogy into an understanding of how particular subject matter are organized, adapted, and represented for instruction" [10]. Therefore, how successful the W activities in improving students experience will depend on other pedagogical factors. These factors are not assessed in our study.

In order to find out how the scores W and E are related, first a linear relationship is explored on individual samples (W_i, E_i) from period $i = 1, 2, 3, 4$ and then the samples regrouped into (W_{1234}, E_{1234}) where $W_{1234} = (W_1, \ldots, W_4)$ and $E_{1234} = (E_1, \ldots, E_4)$. A similar pattern is observed in each period. Figure 1 illustrates the W- E relationship in the last sample.

Table 1 shows that the W scores contributes by at least 25% in the linear model. The small p-value in the F-test shows a strong dependence between the two variables.

Linear model	Sample size	Slope	R^2	F test p-value
$E_1 \sim W_1$	140	0.59661	0.25	$3.191\ 10^{-10}$
$E_2 \sim W_2$	56	0.1784	0.3865	$3.161\ 10^{-07}$
$E_3 \sim W_3$	59	0.07209	0.4351	$1.331\ 10^{-08}$
$E_4 \sim W_4$	125	0.04462	0.3689	$5.925\ 10^{-14}$
$E_{1234} \sim W_{1234}$	380	0.03534	0.3257	$< 2.2\ 10^{-16}$

Table 1: Linear regression and Correlation

Next, we perform the Pearson's chi-square test of association, to explore more the contribution of the W scores to the exams' scores (Table 2). We reduce the data to a 2×2 contingency table by redefining each W_i as being either $'' \geqslant a_W''$ or $'' < a_W''$ and each E_i either $'' \geqslant a_E''$ or $'' < a_E''$. With the sample (W_{1234}, E_{1234}) and $a_W = a_E = 60$, the restriction stated to maintain the integrity on the χ^2 approximation holds. The residuals in the cell "$E < 60$"/"$W \geqslant 60$" (Table 3) (resp. "$E \geqslant 60$"/"$W < 60$") is negative which shows that the proportion of E counts that are less (resp. greater) than 60 deviates from the expected count. There is a strong positive association between the column "$W < 60$" (resp. "$W \geqslant 60$") and the row "$E < 60$" (resp. "$E \geqslant 60$"). On the other hand, there is evidence that the rows and the columns of the contingency table are dependent since the p-value of the test is < 0.05.

Sample size $n = 380$	$W < 60$	$W \geqslant 60$
$E < 60$	37 (14.34737) [0.038]	79 (101.6526) [0.268]
$E \geqslant 60$	10 (32.65263) [0.086]	254 (231.3474) [0.609]

Table 2: Observed Counts (Expected numbers) [Probability]- p-value=6.606e-14

Sample $n = 380$	$W < 60$	$W \geqslant 60$
$E < 60$	5.980432 (63.66%)	-2.246774 (8.98%)
$E \geqslant 60$	-3.964237 (27.97%)	1.489314 (3.94%)

Table 3: Observed Residuals (Contributions in $\chi^2 = 56.182$)

4. Period W-MT-E

The purpose of introducing the motivational test MT as part of the assessment is to guide students in making the right choice between Precalculus and Differential Calculus courses during the first two weeks of the semester. This dilemma has been a concern for many institutions. The literature makes it clear that no single set of practices are effective for every college or every student [1,2]. Remedial courses and mandatory Placement tests were adopted as solutions to the problem. However their implementation raised other concerns such as equity and inaccuracy. We refer the reader for the description of MT's implementation in [4]. MT is supported by the inhouse review package MS; (Math Support) a self-contained multimedia course covering the Precalculus content.

Table 4 shows that more than 50% ($R^2 > 50\%$) of the variation of E is explained by the linear model involving W and MT. Each of the individual t-statistics are highly significant and the F statistic as well. These are evidence of a strong correlation between the variables. In the last sample, the coefficient of MT is 0.4376, which means that if we have two students with equal W, the one with higher MT will have a higher E score and it will be higher by 0.4376 times of the higher MT score.

$E \sim c + \alpha W + \beta MT$	Sample size	p-value-α	p-value-β	R^2	F test-p-value
$E_2 \sim -10.8644 + 0.4939W_2 + 0.5757MT_2$	56	0.00931	$5.3\ 10^{-06}$	0.5865	$6.848\ 10^{-11}$
$E_3 \sim 19.44547 + 0.21358W_3 + 0.45278MT_3$	59	0.008128	$2.31\ 10^{-06}$	0.6221	$1.464\ 10^{-12}$
$E_4 \sim 21.05286 + 0.20940W_4 + 0.39645MT_4$	125	$1.90\ 10^{-05}$	$3.65\ 10^{-09}$	0.5261	$< 2.2\ 10^{-16}$
$E_{234} \sim 17.57441 + 0.23698W_{234} + 0.43760MT_{234}$	240	$2.23\ 10^{-09}$	$< 2\ 10^{-16}$	0.5332	$< 2.2\ 10^{-16}$

Table 4: Multilinear Regression and Correlation

Next, we separate the data into two groups of exams scores with no MT; the sample E_1, and with MT; the sample $E_{234} = (E_2, E_3, E_4)$, then form the groups $E < l = 50$ and $E \geqslant 50$ (Table 5). The question of independence of the two methods (Calculus scores with "MT" or without a Motivational test "nMT") is investigated using the Chi-Square test of independence on the resulting 2×2 contingency table. The residuals (Table 6) in two cells are negative, which shows that the proportion of E counts deviates from the expected count. On the other hand, we see, in the cell $(E \geqslant 50; MT)$, a larger E count than the expected count. The residuals and contribution of each cell are visualized using corrplot package [12] (Figure 2). The probability of the cell $(E \geqslant 50; MT)$ is the highest and shows that the MT implementation has a positive influence on E scores.

Sample $n = 380$	nMT	MT
$E < 50$	15 (24.32)	51 (41.68)
$E \geqslant 50$	125 (115.68)	189 (198.32)

Table 5: Observed Counts (Expected numbers)- p-value=0.0008811

Sample $n = 380$	nMT	MT
$E < 50$	-1.889 (0.112)	1.443 (0.193)
$E \geqslant 50$	0.866 (0.256)	-0.662 (0.439)

Table 6: Observed Residuals (Probabilities) - $\chi^2 = 4.0706$

(a) Residuals (b) Contributions to χ^2

Figure 2

Finally, by letting the limit "l" varying from 40 to 80, the variation of the p-values is visualized in Figure 3. In particular, the MT seems affecting the range of E scores less than 55.

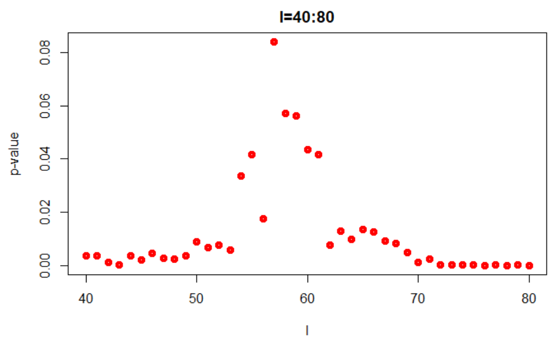

Figure 3: Variation of p-values

5. Period W-MT-L-E

$E \sim c + \alpha W + \beta MT + \gamma L$	Sample size	coefficient (p-value)	R^2	F test-p-value
$E_3 \sim L_3$	59	L: 0.14901 (0.0109)	0.1085	0.01086
$E_4 \sim L_4$	125	L: 0.22425 ($4.47\ 10^{-08}$)	0.2168	$4.469\ 10^{-08}$
$E_3 \sim W_3 + MT_3 + L_3$	59	$W: 0.217783$ (0.009322) $MT: 0.456137$ ($3.28\ 10^{-06}$) $L: -0.009157$ (**0.828007**)	0.6225	$1.112\ 10^{-11}$
$E_4 \sim W_4 + MT_4 + L_4$	125	$W: 0.18855$ (0.000397) $MT: 0.38283$ ($2.21\ 10^{-08}$) $L: 0.03624$ (**0.334785**)	0.5297	$< 2.2\ 10^{-16}$
$E_{34} \sim W_{34} + MT_{34} + +L_{34}$	184	$W: 0.21046$ ($1.09\ 10^{-06}$) $MT: 0.39158$ ($4.24\ 10^{-13}$) $L: 0.01858$ (**0.515**)	0.5492	$< 2.2\ 10^{-16}$
$E_{34} \setminus \{L = 0\} \sim W_{34} + MT_{34} + +L_{34}$	131	$W: 0.31394$ ($2.64\ 10^{-06}$) $MT: 0.37163$ ($2.14\ 10^{-07}$) $L: 0.10092$ (**0.0333**)	0.5319	$< 2.2\ 10^{-16}$

Table 7: Linear Model and Correlation

The L activity involves the use of Maple software as a pedagogical tool in a twofold manner:

- simulate some math concepts; locate local extreme points and inflections points, guess a limit behavior through a plot, locate a root, etc.

- develop critical aptitudes when using a computer. In reading the graphs and the outputs, students have to be careful. For example, by varying the size of the interval, a graph changes. Some aspects are hidden because the interval is too large or too small. The outputs are misleading because of the accumulation of errors.

Most students are not comfortable with this activity because it requires additional learning challenges. To lessen these challenges, I made available a Maple-Toolbox where Maple-codes and their outputs are classified by topics. I also planned Lab sessions to work together.
This discomfort is still perceived in Table 7. Three p-values for testing the coefficients of L, in a linear model, are high. We don't have evidence to support the influence of L scores on E scores. However, while eliminating the L null scores (where students didn't participate in the activity), the p-value shows that the L coefficient is significant at 5% level. This suggests to look for ways to better support students' engagement towards this activity. Indeed, this activity is a challenge for students for several reasons. It involves:

- an extra learning with Maple software,

- technical difficulties for accessing the software via MyApps; University's one-stop portal for students, Faculty and staff to remotely access the software they need.

- readiness with learning the concepts, and

- applications in other disciplines.

These elements interact with one another involving technology, pedagogy, and content knowledge. The combination of these three elements form the new knowledge "TPCK" that is brought to play while teaching with a technology [10]. "The addition of a new technology is not the same as adding content and pedagogy..." [10]. The technology is well integrated when the three elements work in an equilibrium together. I think that students may have felt a certain disconnection between the three parts, and working in developing the "TPCK" may lead to some improvements.

6. Period W-MT-L-M-E

The M activities are formulated as quizzes from Webassign where questions involving applications like economics, biology are selected. The questions include the necessary definitions and notations about the subject. More precisely, this is an attempt to increase student's awareness of the math concepts' applicability. It is not in our purpose to integrate another discipline as some institutions took the lead through "coordinated" courses, or "fully integrated" ones [5].

Linear Model	Sample size	coefficient (p-value)	R^2	F test-p-value
$E_4 \sim M_4$	125	M: 0.30477 (0.0109)	0.07056	0.002753
$E_4 \sim W_4 + MT_4 + L_4 + M_4$	125	W : 0.03318 (**0.20246**) MT : 0.39979 (5.78 10^{-09}) L : 0.05216 (**0.14365**) M : 0.12828 (0.00329)	0.5379	$< 2.2\ 10^{-16}$
$E_4 \setminus \{L = 0\} \sim W_4 + MT_4 + L_4 + M_4$	86	W : 0.15440 (**0.112632**) MT : 0.35387 (0.000378) L : 0.23693 (0.000786) M : 0.06156 (**0.262617**)	0.5239	1.973 10^{-12}
$E_4 \setminus \{M = 0\} \sim W_4 + MT_4 + L_4 + M_4$	110	W : 0.04596 (**0.09677**) MT : 0.49321 (2.46 10^{-09}) L : 0.05216 (**0.15170**) M : 0.10475 (0.03971)	0.4967	5.989 10^{-15}
$E_4 \setminus \{L = 0 \text{ or } M = 0\} \sim W_4 + MT_4 + +L_4 + M_4$	84	W : 0.12687 (**0.226703**) MT : 0.41221 (7.84 10^{-05}) L : 0.24451 (0.000551) M : 0.03737 (**0.501887**)	0.5208	5.178 10^{-12}

Table 8: Linear Model and Correlation

The analysis of the scores show that the four variables contribute to the multilinear model by at least 50%. However, some of these variables contribute less (see Table 8). The linear model may not be the best for these variables. The M activities are related significantly to E scores in the simple linear model with a small contribution. Like the Lab activity, the modelling activity is less engaging for students because it requires extra understanding of new definitions and making connections with the abstract math concepts. Following the framework suggested in [10], one can imagine the element modelling M added to Technology T, Pedagogy P, and Content C. To improve the M integration, one should develop the zone MTPCK and look simultaneously at the other possible intersections of the four elements.
In the following section, in order to get more reading from our data, we explore whether there is a difference in means in the E scores between the different periods, as well as the difference of variances.

7. Exploring Difference of Means and Variances

We perform z-tests for difference of means from the populations of E_1, E_2, E_3, E_4 at 5% level of significance. Table 9 gives the p-values of the tests and the lower "LCI" and upper "UCI" confidence intervals bounds for the difference of means parameter.

Comparison	Observed z-statistic	p-value	LCI	UCI
$E_1 - E_2$	2.2792862	2.265006 10^{-02}	1.056585	14.026986
$E_1 - E_3$	3.2799345	1.038312 10^{-03}	3.541138	14.057288
$E_1 - E_4$	4.2896821	1.789290 10^{-05}	5.307120	14.236754
$E_2 - E_3$	0.3431043	**7.315200 10^{-01}**	−5.925556	8.440411
$E_2 - E_4$	0.6598372	**5.093583 10^{-01}**	−4.394235	8.854538
$E_3 - E_4$	0.3511825	**7.254514 10^{-01}**	−4.456089	6.401537

Table 9: Comparing means

One can observe that:

- There is sufficient evidence to claim that there is a difference in the means between the sample with W activities only and the samples including MT, then $MT + L$, and then $MT + L + M$. The difference observed expresses that the sample mean of the population with W only is larger than the mean of the population involving MT, $MT + L$, $MT + L + M$.

- We cannot claim that there is a difference of means between populations involving more than W activities.

This again confirms our previous observations on the effect of the MT, L and M activities. We notice also that the difference of means is reasonable and doesn't affect students's success. Hence, the valuable experience that students will gain with the FAs is a worthy effort. Finally, in comparing variances, Table 10 and Figure 4 show that the value 1 is contained in each of the confidence intervals. This infer that it is possible that there is not an important difference in variance between the populations.

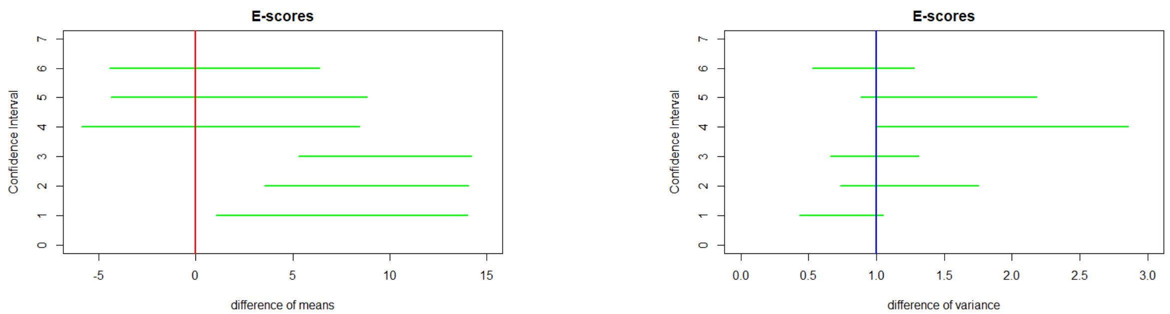

Figure 4: Comparing means and variances

Comparison	Observed F statistic	p-value	LCI	UCI
$E_1 - E_2$	0.68592	0.08126	0.430054	1.047645
$E_1 - E_3$	1.1557	0.5382	0.7321105	1.7534089
$E_1 - E_4$	0.93216	0.6855	0.6597819	1.3124966
$E_2 - E_3$	1.6848	0.05152	0.9965356	2.8591072
$E_2 - E_4$	1.359	0.1649	0.881524	2.180359
$E_3 - E_4$	0.80661	0.3616	0.5266265	1.2809625

Table 10: Comparing Variances

This data analysis was conducted also on the grades of Integral Calculus; a sequel of Differential Calculus course. The MT activity is not included in the FAs. We had similar observations. The only difference appears in the confidence intervals for the difference of means where we observe an improvement of the mean scores with the FAs. Most of the students in this course have taken the first course "Differential calculus" and were exposed to similar FAs. This may had an effect in the increase of the average and confirms the possibility of improvement while supporting better students.

Periods	Observed z-statistic	p-value	LCI	UCI
$(W,E)-(W,L,E)$	-3.192378	0.001411065	-17.436084	-4.1706229
$(W,E)-(W,L,M,E)$	-2.116565	0.034296772	-11.240342	-0.4318013
$(W,L,E)-(W,L,M,E)$	1.407145	0.159384358	-1.951474	11.8860385

Table 11: Comparing means - Integral Calculus

Periods	Observed F-statistic	p-value	LCI	UCI
$(W,E)-(W,L,E)$	1.5776	0.09221	0.9259788	2.4962657
$(W,E)-(W,L,M,E)$	1.591	0.02411	1.063940	2.315782
$(W,L,E)-(W,L,M,E)$	1.0085	0.9529	0.5969776	1.7855406

Table 12: Comparing Variances - Integral Calculus

8. Conclusion

It is worth noting that our FA selection follows Bloom's Taxonomy in learning. Indeed, one can locate: MT at the level of "recalling basic concepts", W at the level of "understanding and practicing" concepts, and L & M at the level of "focus on applying and analyzing". Thus, it guides students in their learning progressively.

The Focused Assessment approach appears as a practical way in planning an assessment based on selected learning objectives and skills to achieve. It offers flexibility in evaluating a particular FA's success/failure, implementing a new one, or trying new modifications. Our statistical analysis reveals that, with appropriate design when introducing the FAs, students may improve on their performance and benefit in gaining new skills while learning Calculus courses.

Ethics Declaration

The data used in this study has received ethics review and approval by the York University's Human Participants Review Sub-Committee (Certificate #: 2022-020).

References

[1] C. R. Belfield, *Improving assessment and placement at your college: A tool for institutional researchers*, Teachers College, Community College Research Center, Columbia University, New York, (2014).

[2] Chris Brown, *Placement testing and remedial mathematics for post-secondary students: Prescription for success?*, https://www.mun.ca/educ/faculty/mwatch/vol1/cbrown.html, (1999).

[3] M. J. Crawley, *The R Book*, 2nd Edition, John Wiley& Sons, (2013).

[4] S. Challal, *A Just-in-Time Remediation for First year Students in Calculus*, FYMSiC Newsletter 9th Edition, (2022).

[5] Jason W. Dunn and Julius Barbanel, *One model for an integrated math/physics course focusing on electricity and magnetism and related calculus topics*, Am. J. Phys., 68(8)(2000).

[6] M. Gardener, *Beginning R. The Statistical Programming Language*, John Wiley& Sons, (2012).

[7] H. J. Larson, *Introduction to probability theory and statistical inference*, John Wiley & Sons, (1982).

[8] R. J. Larsen and M. L. Marx, *An introduction to mathematical statistics and its applications*, 3rd Edition, Prentice Hall, (2001).

[9] William Mendenhall, Robert J. Beaver and Barbara M. Beaver, *Introduction to probability and statistics*, 15th Edition, Cengage and Learning, (2018).

[10] Punya Mishra and Matthew J. Koehler, *Technological Pedagogical Content Knowledge: A Framework for Teacher knowledge*, Teachers College Record, 108(6)(2006), 1017-1054.

[11] J. Stewart, *PKG Essential Calculus with EWA software*, 2nd Edition, Cengage Learning, (2012).

[12] Statistical tools for high-throughput data analysis, http://www.sthda.com/english/wiki/chi-square-test-of-independence-in-r