UNIÓN
REVISTA IBEROAMERICANA DE EDUCACIÓN MATEMÁTICA

ISSN: 1815-0640
Número 56. Agosto 2019
Páginas 26-46

www.fisem.org/web/union
http://www.revistaunion.org

Identifying academically at-risk incoming freshmen at a private university in Uruguay: Psychometric evaluation of a mathematics diagnostic test

Wilson González-Espada, Eduardo Lacués, Gabriela Otheguy, Magdalena Pagano, Alejandra Pollio Rosina Pérez, Marcos Sarasola

Fecha de recepción: 13/12/2018
Fecha de aceptación: 27/08/2019

Resumen	Este estudio determinó hasta qué grado la Prueba Diagnóstica de Matemáticas (PDM) que se ofrece en la Universidad Católica de Uruguay (UCU) es psicométricamente apropiada. Además, se exploró hasta qué grado la PDM se correlaciona con el éxito académico. Se halló que cinco de las preguntas originales de la PDM, de un total de 30, fallaron las pruebas de validez y confiabilidad, y se eliminaron. Las puntuaciones del resto de las preguntas tuvieron una alta correlación con la cantidad de cursos de matemáticas requeridos al final del primer año, confirmando que los estudiantes con una puntuación baja podrían necesitar apoyo adicional para permanecer en la carrera de ingeniería. **Palabras clave**: Teoría de respuesta al ítem, pruebas diagnósticas, evaluación, matemáticas.
Abstract	This study determined to what extent the mathematics diagnostic test (MDT) used at the Catholic University of Uruguay (CUU) was psychometrically appropriate. Also, after removing “red-flagged” items, the study measured to what extent MDT scores correlated with academic success. It was found that five MDT items (out of 30) did not meet the guidelines and were discarded. The score on the remaining items showed the highest correlation with the number of mathematics courses completed, confirming that students with low MDT-Revised scores might need additional academic support to remain in the engineering program. **Keywords**: Item response theory, diagnostic testing, assessment, mathematics.
Resumo	Este estudo tem dois objetivos principais. O primeiro é determinar até que ponto o exame diagnóstico de matemática (EDM) empregado na Universidade Católica de Uruguay (UCU) era psicométricamente apropriado. O segundo, depois de eliminar todas as perguntas não apropriadas, é determinar até que ponto o EDM poderia ser usado para identificar, a priori, estudantes de engenharia que poderiam ter dificuldade em completar seu primeiro ano de estudos universitários na UCU. Achou-se que 5 perguntas do EDM não correspondiam às indicações, e foram eliminadas. A pontuação nas perguntas restantes

mostrava a correlação mais alta com o número de aulas de matemática completadas, o que confirma que os estudantes poderiam precisar de apoio académico adicional para poder permanecer no programa de engenharia.
Palavras-chave: Teoria de resposta a item, exames diagnósticos, avaliação, matemática.

1. Higher Education: Value and Admission Models

In many countries, a college education is highly valued as a public good. A recent UNESCO report indicated that the global expansion of higher education in the last 50 years has been an unprecedented "academic revolution" (Altbach, Reisberg, & Rumbley, 2009). The report noted the chronology of higher education massification, starting with the United States, Canada, Western Europe, and Japan, from 1960-1980, followed by countries in East Asia, Latin America, China, India, and many other developing nations soon afterward. Currently there are over 200 million students enrolled in higher education institutions worldwide (UNESCO, 2015).

A college education is perceived not only as a medium for training and certifying professionals in most disciplines, but also functions to obtain a better paying job with better working conditions and benefits (Thompson, 2019; Villegas, 2019). In the United States, for example, a person with a bachelor's degree diploma earns a median weekly income that is about 65% higher compared with a person with only a high school diploma (U.S. Department of Labor, 2015). Other personal benefits include lower unemployment and better health outcomes (Baum & Payea, 2004; Institute for Higher Education Policy, 2005). From a broader viewpoint, a college education is perceived as a way for societies to develop knowledge economies based on "human capital" economic investments (Gillies, 2011). Other social benefits of higher education include higher volunteer and civic participation, and lower crime (Lochner & Moretti, 2004; Trostel, 2010).

Two basic models of student admissions in higher education are commonly used: open and selective admissions (Mullin, 2017; Tsao, 2005). Institutions that practice open admission policies undergo a non-competitive selection process where all students who want to complete a college education can do so, regardless of previous academic credentials or experiences. In contrast, selective institutions offer admission to only a fraction of the students who apply, mainly the top tier of students based on criteria like high school grades or standardized college entrance examinations.

Those that support open admissions assume that most students, with the proper support, can be successful in college (Ingram & Morrissey, 2009). Other researchers argued that open admissions is essential for assuring access to education and better opportunities for as many people as possible (Vaughan, 1985), producing well-informed members of a democratic society (Roueche & Baker, 1987) and reducing racial, ethnic and economic inequality (Ferreyra, et al., 2017; Mullin, 2012a, 2012b). Further, Shannon, & Smith (2006) argue that open admission policies influence admissions and enrollment processes, curricular structures, faculty hiring, the relationships between community colleges and four-year institutions, advising and counseling activities, and institutional responses to the needs of the schools and employers.

Other higher education institutions favor a selective approach to college admission (Scherer & Anson, 2014). Common arguments in favor of limiting enrollments to academically pre-screened student include allocating scarce money and resources to help students with the greatest probability of academic success and college completion (Bissett, 1995), maintaining high academic standards, providing accelerated academic programs and a rigorous sequence of courses (Astin, 1990), offering additional economic returns for graduates compared with non-selective institutions (Bastedo & Jaquette, 2011), and increasing prestige and improved quality for selective institutions (Dale & Krueger, 2002). Some researchers have argued that the more selective universities are, the less likely it is for poor or minority students to apply, even those with stellar grades, excluding themselves out of the process. This is known as “undermatching”, when there is a significant pool of poor students who are attending colleges that are less selective than the ones they could have attended based on their academic preparation (Bowen, Chingos, & McPherson, 2009). In the United States, most university and a few community colleges follow the selective admissions models.

2. Student Admission and Attrition in Uruguay

Regardless of whether a university has open or selective enrollment, the unfortunate fact is that not all students that are accepted into college reach graduation (Voigt & Hundrieser, 2008). The transition from high school to the first year of college seems to be particularly problematic (González-Espada & Napoleoni-Milán, 2006) Therefore, universities invest resources to reducing attrition (departure from college) and increasing retention (persistence until degree completion).

Research has uncovered many factors that affect student attrition. Raisman (2013) lists four main ones, which he summarized as “college does not care”, “poor service and treatment”, “not worth it”, and “scheduling” issues. The students’ intellectual ability in meeting the demands of academic programs is commonly mentioned as an important factor in college attrition (Byrd & MacDonald, 2005; Laskey & Hetzel, 2011). Tinto (1993) proposed that cultural assimilation, including academic and social integration, is what leads students to remain enrolled in college. Presumably, the more integrated to the college experience students become, the higher the commitment to the process and the less likely they are to leave. Conversely, Student attrition is related to students’ lack of academic integration through differing academic values, and students’ lack of social integration with faculty and peers.

Some entering freshmen might lack "soft skills" needed to succeed in an academic environment, such as attending class, maintaining concentration, using effective study techniques, applying metacognitive tools to recognize to what extent they are understanding the material, asking questions to faculty or peers, recognizing when to seek academic assistance, among others (Laskey & Hetzel, 2011). Additional factors that might contribute to college attrition include the students' personality traits (extraversion, agreeableness, conscientiousness, neuroticism, openness/intellect) and even the students' support and quality of their social network (Eckles & Stradley, 2012; Eggens, van der Werf, & Bosker, 2008; Ridgell & Lounsbury, 2004).

The postsecondary education system in Uruguay was 100% public from 1849 to 1985. Historically, the only option for a college education was Universidad de la República (UdelaR), which was fully supported by the government and where students paid no tuition. Still to this day, UdelaR is available at no cost. After 1985, private universities were allowed to become established, including Catholic University of Uruguay, Montevideo University and ORT University (Obchestvo Remeslenogo Truda; Association for the Promotion of Skilled Trades). These and similar institutions are 100% supported by students' tuition and receive no support from the government (Santalices, 2010).

In Uruguay, the last 15 years have seen slight increases in the number of college students. MEC (2013) noted that the enrollment at UdelaR, about 110,000 students, is about six times larger than the enrollment of all other private universities combined (about 20,000 students). A recent poll reported that 64% of Uruguayans have a positive view of public higher education and 45% have a positive view of private higher education. About 9% and 4%, respectively, have a negative view of higher education. The perception was similar for both residents and non-residents of the capital, Montevideo (Botinelli, 2012).

Not only do all Uruguayan universities have open-admissions policies; students can also attend the public university at no cost. The philosophical underpinnings of UdelaR, including autonomy from political influences, management by shared governance, admission without entrance examinations, no tuition costs, and academic freedom (Universidad de la República, 1958) go back to the republic's founding and were strongly influenced by the Enlightenment and the French Revolution (Karamán-Chaparenco, 2010). An open admission to higher education, Boado (2011) argued, serves an essential democratic, social and civic formation role, even if not all students obtain a college degree.

Interestingly, some scholars have questioned whether the open-admissions policy of Uruguayan universities had truly achieved the goal of equality in access to all citizens (Bruner, 2011). Enrollment data suggests that most students come from medium and high socioeconomic levels and that relatively few students from low socioeconomic levels enter higher education or complete their undergraduate degree. Currently, there are no plans for the government to change the open-admissions policy of public and private universities, or to modify their current economic support structure (Bruner, 2011).

The data associated with college student retention and attrition in Uruguay is compiled and reported by the Ministry of Education and Culture. In their latest report (Ministerio de Educación y Cultura, 2013), it was reported that 23,657 students were admitted to the public university system and 3,785 students were admitted to private universities. In the same year, 6,290 and 1,744 students obtained an undergraduate degree from the public and private university systems, respectively. Assuming that the enrollment has remained more or less similar over the last 5-6 years, that would result in a retention of about 27% and 46% for public and private universities, respectively. Note that, unlike the United States, Uruguay does not calculate retention rates using a first-time, full-time, freshmen 6-year cohort.

Boado (2011) analyzed hundreds of in-depth interviews of Uruguayan college dropouts to report several factors that caused students to stop attending college. These include: (a) Time-management; reduction of free time, difficulty studying and working simultaneously, including transportation issues and class scheduling that overlap work hours; (b) Career disillusionment; a perception dissonance between their perception and the reality of their chosen career; (c) Low academic performance; students obtained low grades in several subjects, and could not improve them even after repeating classes; (d) Curriculum; emphasis on theory instead of practical, hands-on involvement in a career as underclassmen, (e) Length of the program; it was perceived that it took too long to complete all the required coursework; and (f) Institutional climate; students reported an excessively competitive and hostile college environment. Complementary emplanations can be found by considering prior content knowledge (Bourel, Kahan, Miguez, & Stalker, 2017) and reasoning abilities (Lacues, Díaz, & Huertas, 2018) that first-year students bring.

More recently, a group of Uruguay researchers have examined data associated with attrition from local engineering programs and the role of inadequate mathematics preparation (Artigue, Flores, Lacués, & Messano, 2017). Their study concluded that students who have showed repeated academic failure in mathematics classes do not overcome them just by repeating courses until they pass them, and many of them decide lo leave engineering altogether. The researchers argued for reorganizing the curriculum and tutoring in one-on-one and small group settings as interventions that have shown a measure of success in keeping students in the career.

In an open admission university, like the ones in Uruguay, a way to identify incoming freshmen background content knowledge is through diagnostic testing of skill proficiency. This strategy can give academic advisors an idea of who might be more likely to succeed in a specific course and who might be more likely to struggle academically (Legg, Legg, & Greenbowe, 2001; OCDE, 2006). It also gives students an idea of whether or not they might be ready for specific courses they are enrolled in. Diagnostic testing is particularly common in mathematics-related courses, where “math anxiety” is commonplace (Núñez-Peña, Suárez-Pellicioni, & Bono, 2013; Perry, 2004), and many college courses build upon content supposedly learned in high school mathematics courses. Diagnostic testing allows faculty to not only determine the students' mathematical knowledge on entry, but to provide an early alert for students who are most likely to need additional academic assistance (Fhloinn, Bhaird, & Nolan, 2014).

Of course, a diagnostic test is most useful when it has been properly and psychometrically validated. Large-scale, standardized tests items are validated through the piloting of “trial” items during test administrations (these “trial” items are not counted in students’ scores). Psychometric validation is particularly difficult with locally-made tests and low sample sizes (Brown, 2000; Koretz, 2008).

Although many of the more complex calculations are not applicable to locally-made diagnostic tests and low sample size applications, others like item difficulty (when a test item is correctly answered by almost all or almost no test-takers) and item discrimination (when below-average scorers performed better on a test item than above-average scorers) can be adapted to this situation (Crocker & Algina, 1986). These are based on Classical Test Theory (de Klerk, 2014; DeVellis, 2006) and Item Response Theory or IRT (Erdodi, 2012; Yang, 2014). Examples of applying IRT ideas to classroom-made tests include González-Espada (2009, 2008) and Knell, Wilhoite, Fugate, & González-Espada (2015).

3. Research Questions and Rationale

The purpose of this exploratory study was to validate a locally-made mathematics diagnostic test (MDT) using IRT ideas. This MDT was designed by the mathematics faculty at the Catholic University of Uruguay (CUU) and was completed by incoming freshmen entering the engineering programs. After removing psychometrically dubious test items, the revised MDT (MDT-R) was used as an independent variable to determine correlations with the dependent variable, which is success in the participants' first year of college. Two research questions guided this study: (a) What MDT items were inconsistent with IRT guidelines and why? and (b) Are there significant correlations between MDT-R scores and academic success after two semesters of college?

In this study, academic success was defined in three different ways. The first was the number of successfully completed mathematics courses during freshmen year. Academic success was also defined in terms of the number of science-related courses successfully completed during freshmen year. Finally, success was defined as the overall students' grade-point average (GPA). It was hypothesized that a higher score in MDT-R is positively correlated with (a) a higher number of completed math courses within the participants' respective career sequence, (b) a higher number of completed science courses within the participants' respective career sequence and (c) a higher overall GPA.

This study is the culmination of several years of scholarly work related to the articulation of secondary and postsecondary mathematics content, including the cognitive and metacognitive mathematics strategies of freshmen students (Álvarez, Czerwonogora, Lacués, Leymonié, & Pagano, 2007; Álvarez, Lacués, & Pagano, 2002; Lacués, Díaz, & Huertas, 2018).) and the creation of validated tools to predict academic achievement based on high school mathematics courses completed (Álvarez, Lacués, & Pagano, 2004; Bourel, Kahan, Miguez, & Stalker, 2017).

So far, the Department of Mathematics has updated, revised and reorganized first year math classes, like Linear Algebra I & II and Calculus I & II, to allow additional time to cover the required content. For example, these classes now have two parts, Part A and Part B, designed in such a way that student can easily retry a half-class they failed, instead of the whole class. Also, class schedules were condensed from four 2.5-months classes to five 2-months classes, and additional summer classes were scheduled, providing students additional flexibility for retaking a class without getting behind in their engineering program. In addition, a number of different remedial interventions, such as tutoring and supplementary readings, have been adopted to assist those students who have received one failing grade in different courses, or who have repeatedly received failing grades in the same course (Artigue, Flores, Lacués, & Messano, 2017). Finally, elective mathematics classes were added as possible pre-requisites to bring students up to date in the mathematical skills needed to tackle the prescribed math sequence.

Preliminary data suggest that previous efforts have resulted in a 10% reduction in the number of engineering students who are falling behind in their engineering program due to struggling in math classes. Using a data-driven validated diagnostic test is essential to quantifying the relative success of the retention strategies that have been implemented at CUU, better identifying students at risk of dropping out of the engineering programs, and helping students ease these programs in a more efficient way.

4. Methodology

4.1 Context

CUU's School of Engineering and Technologies (SET) have been preparing engineers for more than 25 years, and currently includes seven careers: Informatics, Electric, Industrial, Power Systems, Telecommunications, Audiovisual and Food Systems. Each engineering career has its own specific math requirements and a completion timeline. Table 1 briefly describes CUU' mathematics courses. Each mathematics course is divided into two, 8-week long parts, Part A and Part B. Students must approve Part A prior to completing Part B. Table 2 summarizes the required first-year math classes.

Course	Description
Linear Algebra IA, IB	This course presents the systems of equations and elements of the matrix theory in a unified way and uses these elements to represent geometric situations. Elementary cases are introduced; concepts will be developed abstractly in Linear Algebra II.
Linear Algebra IIA, IIB	This course deals with abstractions, using axiomatic theories motivated by and based upon the issues presented in previous courses. Applications to other branches of mathematics or relating professional training are presented to help the understanding processes of construction of models, and to find the role of mathematics in engineering practices.
Introduction to Calculus	This course introduces basic concepts and procedures in calculus, to provide them with tools wich are necessary to succeed in subsequent courses of Mathematics.
Calculus IA, IB	This course introduces the concepts of calculus, limits, derivatives, integrals, are the basis of any application of mathematics to engineering. Maintaining a balance between rigor and intuition, the

	course shows the need of each of these elements in the mathematical work.
Calculus IIA, IIB	This course extends notions of Calculus I to the study of functions of several variables. Differential of a function is defined using tools of Linear Algebra. Construction of applications is emphasized, basing them on the interpretations made of the notions of partial derivatives, differential or integral.

Table 1. Mathematics course descriptions.

Engineering Career	Required First-year Mathematics Courses
Informatics	Introduction to Calculus A, B Calculus IA, IB
Electrical; Power Systems; Industrial; Telecommunications	Linear Algebra IA, IB, IIA, IIB Calculus IA, IB, IIA, IIB
Food Systems	Linear Algebra IA, IB, IIA, IIB Calculus IA, IB
Audiovisual	Linear Algebra IA, IB Introduction to Calculus A, B Calculus IA, IB

Table 2. Required mathematics courses for each SET engineering major.

4.2 Participants

The research sample consisted of 149 freshmen (69.3% male students and 30.6% female students) entering SET who voluntarily completed the MDT two weeks before the Spring semesters of 2015 and 2016 started. Approximately 70% and 30% of the participants graduated from private and public high schools, respectively. Nearly 72% of the participants came from Montevideo and the remaining 28% came from other Departments in Uruguay. The most popular engineering majors among the participants were Informatics (n = 65, 43.6%), Industrial (26, 17.4%) and Food Systems (n = 24, 16.1%). The rest of the participants were planning to major in audiovisual (n = 12, 8.1%), Power Systems (n = 10, 6.7%), Electrical (n = 8, 5.4%), and Telecommunications (n = 4, 2.7%).

Of 149 students who competed the MDT, 134 finished their first year of college. For these students, course transcripts were requested to the Office of the Registrar. These data allowed the researchers to identify the number of mathematics and science courses successfully completed. Grades included S “outstanding” (6.00-5.50), MB “very good” (5.49-4.50), BMB “good very good” (4.49-3.50), and B “good” (3.49-2.50). Failing grades include R “regular” (2.49-1.50) and D “deficient” (1.49-0.00). Also, overall grade point average (GPA) were noted.

4.3 Instrument

For this exploratory, quantitative study, the MDT was the focus of the analysis. The MDT originally consisted of 30 multiple-choice mathematics questions (1 correct option and 3 distractors) covering topics in arithmetic, probability, statistics, algebra, geometry, analytical geometry and calculus, as described in national curriculum guidelines for secondary education. Items were developed using the mathematics competencies described by Niss (2003) as a theoretical framework. These competencies were: (a) thinking mathematically; (b) posing and solving mathematical problems; (c) modelling mathematically; (d) reasoning mathematically; (e) representing mathematical entities; (f) handling mathematical symbols and formalisms; (g) communicating in, with, and about mathematics; and (h) making use of aids and tools. In addition, three basic content and process standards were integrated within MDT included: (a) executing algorithms; (b) using operations in the context of symbolic, formal, and technical; and (c) applying concepts to solve math problems.

Each item of the MDT was examined by calculating two psychometric parameters. The first one was item difficulty, DF, defined as the number of correct answers divided by the sample size. Following Morales (2012), questions were flagged and discarded from secondary analyses if they were deemed too easy (DF > 0.90) or too difficult (DF < 0.10). Questions were retained if their difficulty was moderate (0.90 > DF > 0.10).

The second parameter was item discrimination, DS, and it measures, for a specific discipline, to what extent each question differentiates between those students who know more and those who know less. DS is calculated by subtracting the number of correct answers for students in the overall top 27 test percentile and the number of correct answers for students in the overall lower 27 test percentile. This difference is then divided by either the number of students in the top or lower 27 percentile, whatever group has more students. (Backhoff, Larrazolo, Rosas, 2000). Following Ebel and Frisbie (1991), questions were flagged and discarded from secondary analyses if DS < 0.15. Questions were retained if DS > 0.16, representing average to very good discrimination.

The revised test, MDT-R, was used to calculate descriptive statistics and Pearson correlations with the proportion of mathematics and science courses completed for each engineering program. Another correlation was calculated using the overall GPA, including math courses, science courses and courses in disciplines like social sciences and humanities. Given the exploratory nature of this study, the limited sample size available, and the importance of balancing the possibilities of Type I and II errors, the researchers set the significance level for all statistical tests at 0.05, The analysis resulted in a set of data-driven equations that can be applied to future diagnostic test takers so that students below a certain threshold, could be identified to receive additional academic support.

MDT-R results were not shared with all mathematics faculty members who taught first-year mathematics courses. The rationale behind this decision was to prevent instructors from developing a priori misperceptions about the students' abilities in math that could potentially affect faculty-student interactions and grading. The fact that the expectations an-instructor sets for individual students can cause him/her to exhibit behavior differentiation and can significantly affect the students' performance has been well documented in the literature (Rosenthal & Jacobson, 1968; Workman, 2012).

5. Findings and Discussion

5.1 Test Validation

Table 3 shows the separate 2015 and 2016 DF and DS indexes for the revised MDT, as well as the combined average indexes. Items that were too easy, too difficult, or that showed low discrimination are color-coded.

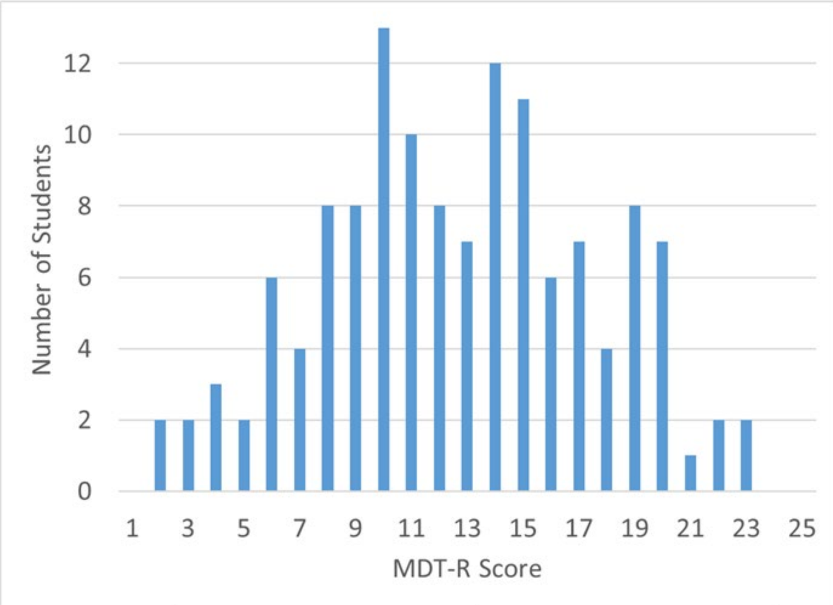
Item	2015 DF	2015 DS	2016 DF	2016 DS	Mean DF	Mean DS	Item	2015 DF	2015 DS	2016 DF	2016 DS	Mean DF	Mean DS
1	0.58	0.68	0.43	0.57	0.505	0.63	16	0.3	0.26	0.16	0.19	0.23	0.23
2	0.63	0.68	0.49	0.57	0.56	0.63	17	0.59	0.74	0.47	0.86	0.53	0.80
3	0.31	0.53	0.34	0.62	0.325	0.57	18	0.74	0.37	0.62	0.57	0.68	0.47
4	0.33	0.32	0.48	0.48	0.405	0.40	19	0.17	-0.11	0.23	0.24	0.2	0.06
5	0.54	0.47	0.47	0.38	0.505	0.43	20	0.63	0.53	0.64	0.48	0.635	0.50
6	0.67	0.53	0.71	0.43	0.69	0.48	21	0.91	0.26	0.92	0.14	0.915	0.20
7	0.55	0.42	0.64	0.43	0.595	0.42	22	0.76	0.16	0.69	0.24	0.725	0.20
8	0.75	0.16	0.57	0.57	0.66	0.37	23	0.62	0.47	0.37	0.57	0.495	0.52
9	0.66	0.58	0.68	0.43	0.67	0.50	24	0.49	0.37	0.51	0.62	0.5	0.49
10	0.24	0.42	0.27	0.19	0.255	0.31	25	0.49	0.53	0.43	0.62	0.46	0.57
11	0.78	0.37	0.67	0.67	0.725	0.52	26	0.44	0.53	0.32	0.76	0.38	0.65
12	0.51	0.79	0.51	0.71	0.51	0.75	27	0.03	0	0.08	0.19	0.055	0.10
13	0.59	0.42	0.55	0.43	0.57	0.42	28	0.43	0.32	0.41	0.24	0.42	0.28
14	0.52	0.21	0.44	0.00	0.48	0.11	29	0.86	0.16	0.76	0.19	0.81	0.18
15	0.41	0.58	0.51	0.62	0.46	0.60	30	0.45	0.05	0.47	0.14	0.46	0.10

Too easy, DF > 0.90 | Too hard, DF < 0.10 | Low discrimination, DS < 0.15

Table 3. MDT-R difficulty and discrimination values.

Items 14, 19 and 30 either have combined or separate item discrimination indexes below the threshold suggested by the literature and were removed from the diagnostic test. Item 21 has combined and separate difficulty values that are too high, that is, most students answered the item correctly, regardless of overall test score. Item 27 has combined and separate difficulty and discrimination indices that are too low; its removal was also recommended by psychometricians. The rest of the items complied with IRT guidelines. For the remaining 25 items, Cronbach's alpha, a measure of internal consistency or reliability, was calculated. The result was an alpha value of 0.794, which means that MDT-R has very good internal consistency (Ebel & Frisbie, 1991).

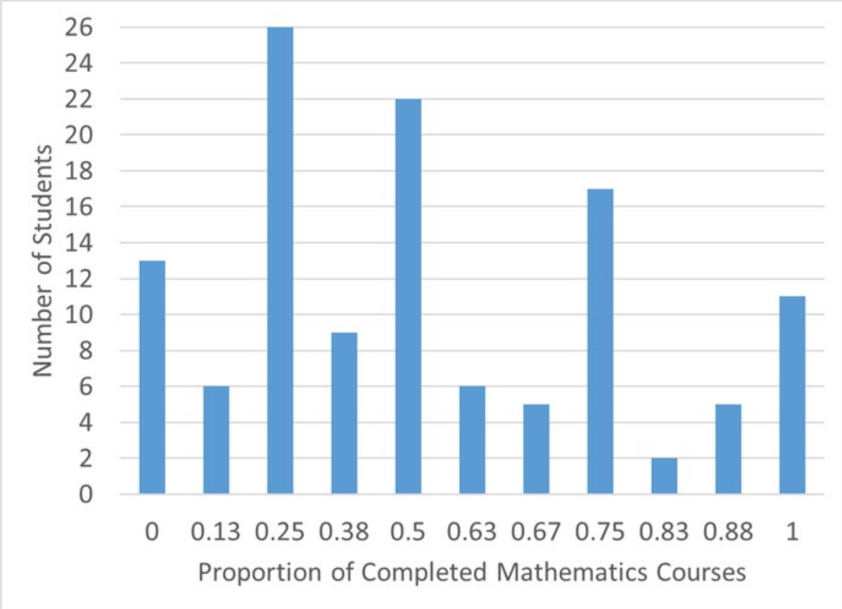
Graph 1 shows how many participants contained a specific score in the MDT-R. The average and median score were 12.65 and 13.00 points, respectively. The standard deviation was 4.85 points.



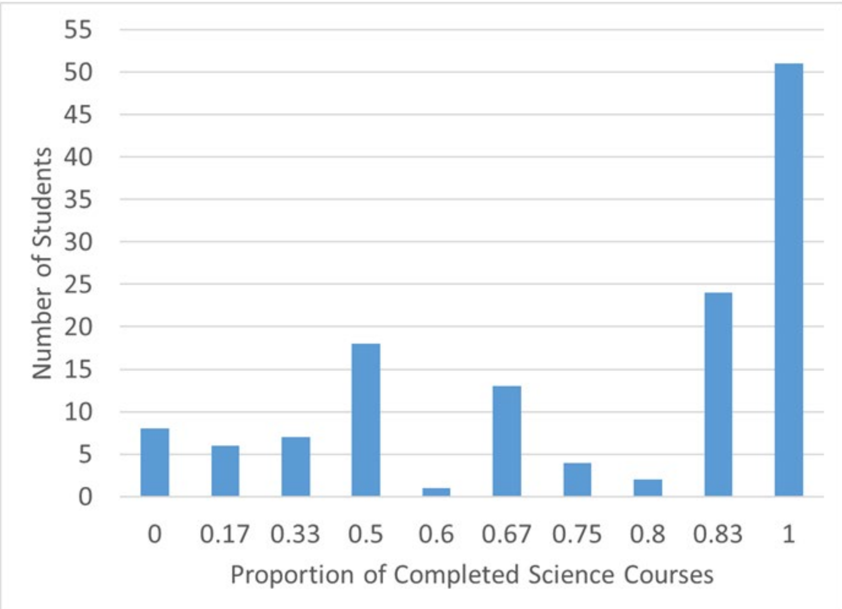
Graph 1. MDT-R Histogram

5.2 Completed coursework in math and science, overall GPA

An analysis of college transcript data revealed that the median proportion of completed mathematics courses was 0.5. In fact, 49.3% of the students was not able to finish half of the required mathematics courses by the end of the second semester. However, required science courses have a much higher completion rates. Only 15.7% of the students was not able to finish half of the required science courses by the end of the second semester. These data were converted into proportions to improve the comparisons because not all majors require the same number of math or science courses. Graphs 2 and 3 summarize these results.

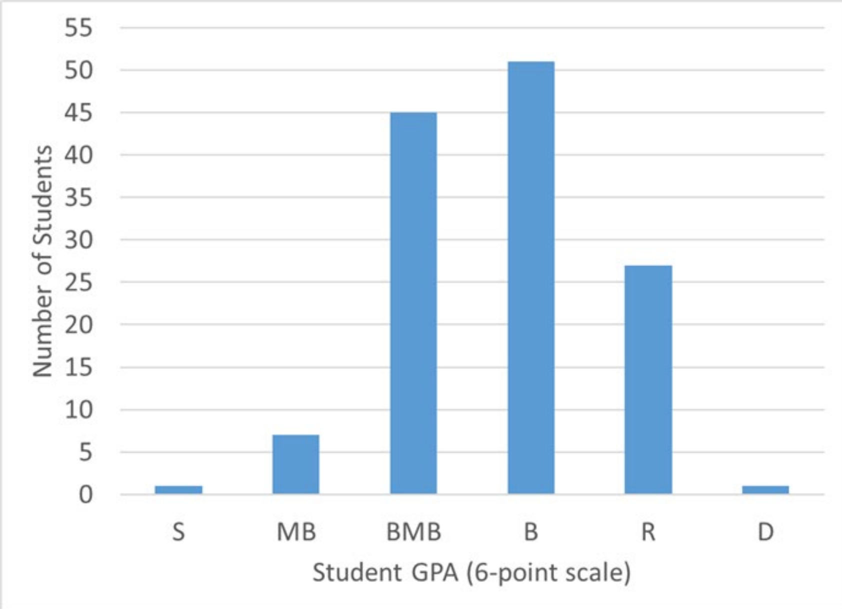


Graph 2. Number of participants who completed a specific proportion of their Year-1 math courses.



Graph 3. Number of participants who completed a specific proportion of their Year-1 science courses.

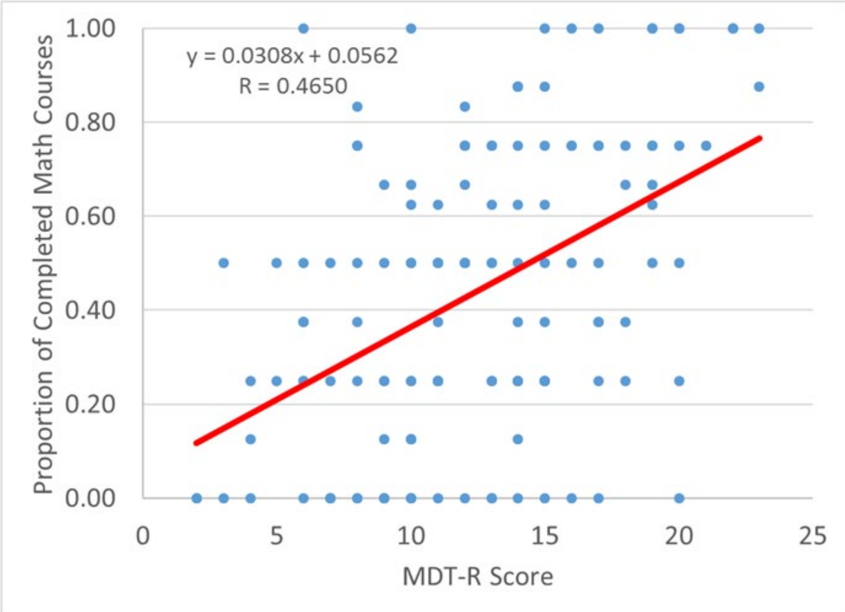
The analysis of college transcripts also discovered that the participants' overall GPA after their second semester had a median of 3.78 (on 6-point scale), equivalent to "good very good". At CUU, letter grades correspond to "outstanding" (Sobresaliente, S; 6.00-5.50), "very good" (Muy Bueno, MB; 5.49-4.50), "good very good" (Bueno Muy Bueno, BMB; 4.49-3.50), "good" (Bueno, B; 3.49-2.50), "regular" (R; 2.49-1.50) and "deficient" (Deficiente, D; 1.49-0.00). Graph 4 summarizes these findings.



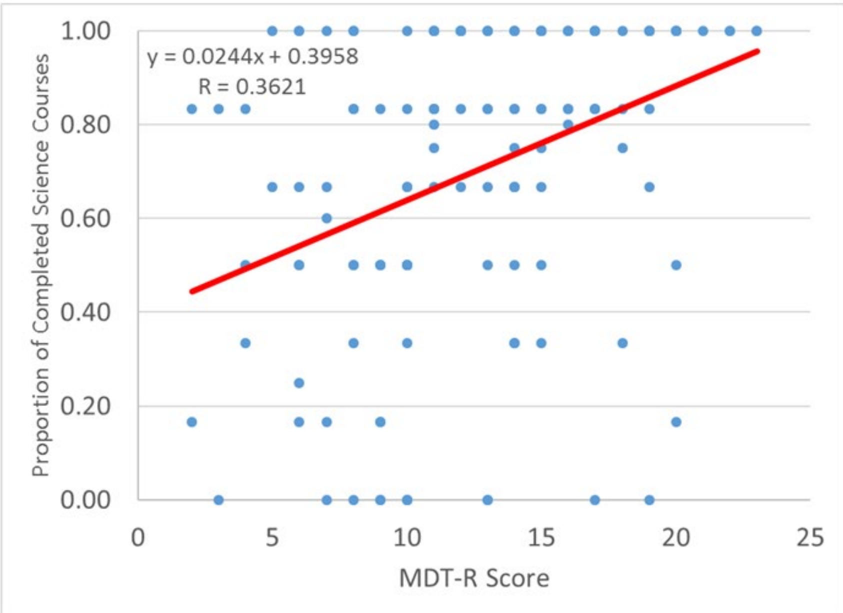
Graph 4. Number of students and their overall grade point average (GPA).

5.3 Correlations between MDT-R and academic success

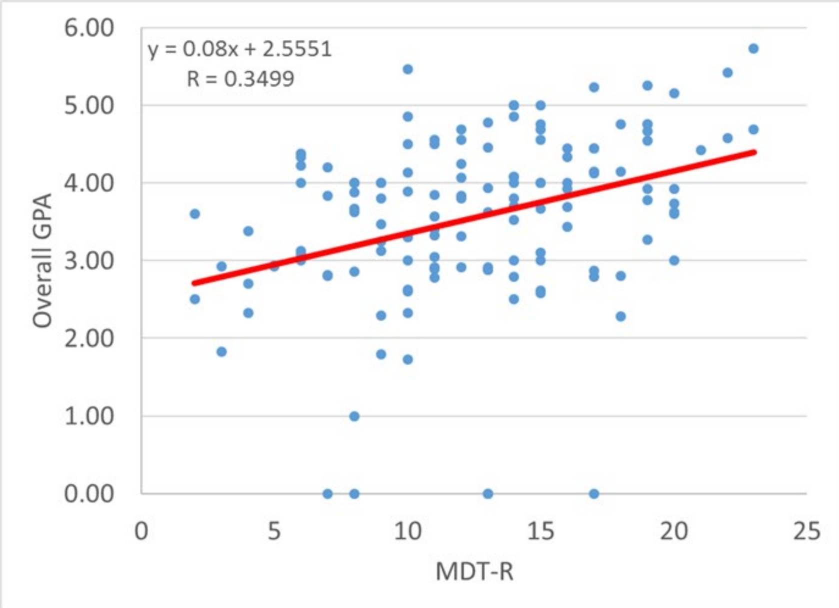
The data from the independent variable (MDT-R score) and three versions of the dependent variable "academic success" were used to calculate three linear regression equations and three Pearson correlation coefficients. In all three cases, significant correlations between MDT-R and the number of required mathematics courses, MDT-R and the number of required science courses, and MDT-R and overall GPA, were found. The strength of the correlations decreased as the comparison moved from math ($R = 0.4650$, $p < 0.0001$), to science, to all courses. This is consistent with research that has established that mathematical ability strongly correlates with success in college, regardless of the academic major selected. Graphs 5-7 summarized these findings.



Graph 5. Proportion of mathematics courses completed by the end of the participants' first year as a function of their MDT-R score.



Graph 6. Proportion of science courses completed by the end of the participants' first year as a function of their MDT-R score.



Graph 7. Participants' first year GPA as a function of their MDT-R score.

6. Conclusion and Limitations

The purpose of this study was to validate a locally-made mathematics assessment, and to determine to what extent it could help identify students who are less likely to succeed in the engineering programs offered at CUU. The first research question was answered in the affirmative; using IRT ideas, several items that were not psychometrically appropriated were identified and removed from the revised version of the diagnostic test. This finding is consistent with other studies that suggest that instructors who create locally-made tests, even those with deep knowledge of a discipline, may inadvertently include a few questions that have wording, context or content that are perceived differently by students (González-Espada, 2009, 2008; Knell, Wilhoite, Fugate, & González-Espada, 2015). These questions might not even stand out as possibly invalid during the grading process until post-hoc IRT analyses are performed (Brown, 2000; de Klerk, 2014; Koretz, 2008).

The second research question was also answered in the affirmative. It was found that students with better scores in the diagnostic test were more likely to complete the science and math courses on time, and were more likely to have a higher overall GPA. The strength of the relationship was not identical for all three measures of academic success after two semesters of college. This finding is in line with prior research that has examined student attrition from engineering careers, particularly those that occur during the high school to college transition (Daempfle, 2003; Haag, Hubele, Garcia, & McBeath, 2007; Geisinger & Raman, 2013; Suresh, 2006). This body of literature has identified inadequate high school math preparation and conceptual difficulties with core college courses in mathematics as two main factors that can explain why students leave engineering programs.

Confirming that about 50% of the participant were not able to reach the midpoint in their mathematics course requirements within two semesters has important implications for revising course offerings and considering differentiated tracks that can ease the transition into engineering careers for at-risk students who are motivated to succeed. Future research can also explore connections between academic success and other factors, such as affective, motivational, deductive structures, and linguistic reasoning.

Despite limitations due to sample size, the length of the diagnostic test, and self-selections of the participants (the diagnostic test was not compulsory and not all incoming freshmen completed it), it is clear that the Department of Mathematics at CUU has a new, validated, quantitative tool to identifying possible students at-risk of falling behind in their engineering program or dropping out altogether. By providing academic intervention strategies to these students, long-term, it is expected that many more of them will be able to finish their engineering degree on time.

7. Acknowledgments

The authors would like to thank the United States Department of State's Fulbright Fellowship Program, Morehead State University and the Catholic University of Uruguay for their support.

8. References

Altbach, P. G., Resisberg, L., & Rumbley, L. E. (2009). *Trends in global higher education: Tracking an academic revolution*. Paris, France: UNESCO, 2009.

Alvarez, W., Lacués, E., & Pagano, M. (2002). Determinación del perfil de los ingresantes a la universidad en relación con las estructuras lógicas que maneja, la capacidad que poseen en el uso del lenguaje simbólico y los conocimientos previos que tienen de cálculo diferencial. *Acta Latinoamericana de Matemática Educativa* (RELME 15 Buenos Aires), 15(2), 663-668.

Alvarez, W., Lacués, E., & Pagano, M. (2004). Diseño y validación de un instrumento predictor del éxito académico de alumnos ingresantes a la universidad. *Acta Latinoamericana de Matemática Educativa* (RELME 17 Santiago de Chile), 17(1), 116-129.

Alvarez, W., Czerwonogora, A., Lacués, E., Leymonie, J., & Pagano, M. (2007). La matemática al ingreso a la universidad. Un estudio comparativo de cuatro facultades en el Uruguay. *Revista Iberoamericana de Educación*, 42(4), 1-9.

Artigue, V., Flores, J., Lacués, E., & Messano, C. (2017). Buscando medidas de apoyo para superar el fracaso académico. *Pensamiento Matemático*, 7(2), 27-42.

Astin, A. W. (1990). Educational assessment and educational equity. *American Journal of Education*, 98(4), 458-478.

Backhoff, E., Larrazolo, N., & Rosas, M. (2016). Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos. *Revista* *electrónica de Investigación Educativa*, 2(1). Retrieved from <http://redie.uabc.mx/vol2no1/contenido-backhoff.html>.

Bastedo, M. N. & Jaquette, O. (2011). Running in place: Low-income students and the dynamics of higher education. *Educational Evaluation and Policy Analysis*, 33(3), 318-339.

Baum, S. & Payea, K. (2004). *Education pays: The benefits of higher education for individuals and society*. Washington, D. C.: College Board.

Bissett, H. G. (1995). Selective admissions in community college nursing programs: Ethical considerations. *Community College Review*, 22(4), 35-46.

Boado, M. (2011). *La deserción estudiantil universitaria en la UdelaR y en Uruguay entre 1997 y 2006*. Montevideo: Ediciones Universitarias.

Botinelli, E. (2012). *Juicio de la población sobre la educación universitaria y ranking de instituciones educativas*. Montevideo, Uruguay: Instituto Factum. Retrieved from: <http://www.factum.uy/pdf/articulos/2012/opv120920.pdf>.

Bourel, M., Kahan, S., Miguez, M., & Stalker, D. (2017). *Informe Herramienta Diagnóstica al Ingreso Componente Matemática*. Montevideo: Facultad de Ingeniería (FING). Retrieved from <https://www.fing.edu.uy/sites/default/files/claustro_citaciones/2017/distribuido/31436/56%20%282016-2018%29%20Informe_HDI_Octubre2017_final-2.pdf>.

Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line: Completing college at America's public universities*. Princeton, New Jersey: Princeton University Press.

Brown, J. (2000). What is construct validity? *Shiken JALT Testing & Evaluation SIG Newsletter*, 4(2), 8-12. Retrieved from: < http://hosted.jalt.org/test/bro_8.htm>.

Brunner, J. J. (2011). *Educación superior en Iberoamérica*. Santiago, Chile: Universia. Retrieved from: <http://200.6.99.248/~bru487cl/files/Brunner_Bogota_16102012.pdf>.

Byrd, K. L., & MacDonald, G. (2005). Defining college readiness from the inside out: First-generation college student perspectives. *Community College Review*, 33(1), 22-30.

Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth, Texas: Harcourt Brace Jovanovich Publishers.

Daempfle, P. A. (2003). An analysis of the high attrition rates among first year college science, math, and engineering majors. *Journal of College Student Retention*, 5(1), 37-52.

Dale, S. B. & Krueger, A. B. (2002). Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *The Quarterly Journal of Economics*, 117(4), 1491-1527.

De Klerk, G. (2014). Classical test theory. In: M. Born, C. Foxtrot, & R. Butter (Eds.). *Online Readings in Testing and Assessment, International Test Commission*. Retrieved from: <http://archive.is/Jji1R>.

Devellis, R. F. (2006). Classical test theory: Measurement in a multi-ethnic society. *Medical Care*, 44(11), S50-S59.

Ebel, R. L. & Frisbie, D. A. (1991). *Essentials of Educational Measurement, 5th Ed*. Englewood Cliffs, New Jersey: Prentice-Hall.

Eckles, J. E. & Stradley, E. G. (2012). A social network analysis of student retention using archival data. *Social Psychology of Education*, 15(2), 165-180.

Eggens, L., Van der Werf, M. P. C., & Bosker, R. J. (2008). The influence of personal networks and social support on study attainment of students in university education. *Higher Education*, 55(5), 553-573.

Erdoli, L. A. (2012). What makes a test difficult? Exploring the effect of item. *Journal of Instructional Psychology*, 39(3-4), 171-176.

Ferreyra, M. M., Abitabile, C., Alvarez, J. B., Paz, F. H., & Urzúa, S. (2017). *At a Crossroads: Higher Education in Latin America and the Caribbean. Directions in Development*. Washington, DC: The World Bank.

Fhloinn, E. N., Bhaird, C. M., & Nolan, B. (2014). University students' perspectives on diagnostic testing in mathematics. *International Journal of Mathematical Education in Science and Technology*, 45(1), 58-74.

Gillies, D. (2011). State education as high-yield investment: Human capital theory in European policy discourse. *Journal of Pedagogy*, 2(2), 224-245.

González-Espada, W. J. (2009). Detecting gender bias through test item analysis. *The Physics Teacher*, 47(3), 175-179.

González-Espada, W. J. (2008). Physical science lab quizzes: Results from test item analysis. *Journal of Science Education/REC*, 9(2), 81-85.

González-Espada, W. J. & Napoleoni-Milan, R. L. (2006). The impact of the freshman year experience on science students. In J. J. Mintzes & W. H. J. Leonard (Eds.). *Handbook of college science teaching*. Arlington, VA: National Science Teachers Association, 351-358.

Geisinger, B. N. & Raman, D. R. (2013). Why they leave: Understanding student attrition from engineering majors. *International Journal of Engineering Education*, 29(4), 914-925.

Haag, S., Hubele, N., Garcia, A., & McBeath, K. (2007). Engineering undergraduate attrition and contributing factors. *International Journal of Engineering Education*, 23(5), 929–940.

Ingram, W. G. & Morrisey, S. E. (2009). Ethical dimensions of the open-door admissions policy. *New Directions for Community Colleges*, 2009(148), 31-38.

Institute for Higher Education Policy (2005). *The investment payoff: A 50-state analysis of the private and public benefits of higher education*. Washington, District of Columbia: Institute for Higher Education Policy.

Karaman-Chaparenco, J. O. (2010). *De la República de las Letras a la República Oriental del Uruguay: El neoclasicismo en la formación del estado y el sujeto nacionales 1811-1837*. 194pp. Unpublished thesis, Doctor of Philosophy in Hispanic Studies, The University of British Columbia, Vancouver, Canada.

Knell, J. L., Wilhoite, A. P., Fugate, J. Z., and González-Espada, W. J. (2015). Using Item Response Theory to improve locally-constructed multiple choice tests: Measuring knowledge gains and curricular effectiveness. *Electronic Journal of Science Education*, 19(7). Retrieved from: <http://ejse.southwestern.edu/article/view/15452/ 9982>.

Koretz, D. M. (2008). *Measuring up: What educational testing really tells us.* Cambridge, Massachusetts: Harvard University Press.

Lacués, E., Díaz, L., & Huertas, J. (2018). ¿Qué estructuras deductivas usan alumnos ingresantes a la universidad? Bolema: Boletim de Educação Matemática, 32(62), 802-824. doi: http://dx.doi.org/10.1590/1980-4415v32n62a03.

Laskey, M. L. & Hetzel, C. J. (2011). Investigating factors related to retention of at-risk college students. *The Learning Assistance Review*, 16(1), 31-43.

Legg, M. J., Legg, J. C., & Greenbowe, T. J. (2011). Analysis of success in general chemistry based on diagnostic testing using logistic regression. *Journal of Chemical Education*, 78(8), 1117-1121.

Lochner, L. & Moretti, E. (2014). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review*, 94(1), 155-189.

Ministerio de Educacion y Cultura (2013). *Anuario estadístico de educación*. Montevideo, Uruguay: Ministerio de Educación y Cultura. Retrieved from: <http://www.mec.gub.uy/innovaportal/file/927/1/anuario_2013.pdf>.

Morales, P. (2012). *Análisis de ítems en las pruebas objetivas*. Madrid, Spain: Universidad Pontificia Comillas, Retrieved from: <https://educrea.cl/wp-content/uploads/2014/11/19-nov-analisis-de-items-en-las-pruebas-objetivas.pdf>.

Mullin, C. M. (2017). *When less is more: Prioritizing open access*. Washington, DC: American Association of Community Colleges. Retrieved from <https://files.eric.ed.gov/fulltext/ED579759.pdf>.

Mullin, C. M. (2012a). *It's a matter of time: Low-income students and community colleges*. *Policy Brief 2012-02PBL*. Washington, District of Columbia: American Association of Community Colleges.

Mullin, C. M. (2012b). *Why access matters: The community college student body*. *Policy Brief 2012-01PBL*. Washington, District of Columbia: American Association of Community Colleges.

Niss, M. (2003). Mathematical competencies and the learning of mathematics: The Danish KOM project. In A. Gagatses & S. G. Papastavridis (Eds.), *3rd Mediterranean Conference on Mathematical Education* Athens, Greece: Hellenic Mathematical Society and Cyprus Mathematical Society, pp. 115-124.

Nuñez-Peña, M. I., Suarez-Pellicioni, M., & Bono, R. (2013). Effects of math anxiety on student success in higher education. *International Journal of Educational Research*, 58(1), 36-43.

OCDE (2006). *PISA marco de la evaluación: Conocimientos y habilidades en ciencias, matemáticas y lectura*. España: Santillana.

Perry, A. B. (2004). Decreasing math anxiety in college students. *College Student Journal*, 38(2), 321-324.

Raisman, N. (2013). *The cost of college attrition at four-year colleges and universities*. Virginia Beach, Virginia: Educational Policy Institute.

Ridgell, S. D. & Lounsbury, J. W. (2004). Predicting academic success: General intelligence, "Big Five" personality traits, and work drive. *College Student Journal*, 38(1), 607-612.

Rosenthal, R. & Jacobson, L. (1968). *Pygmalion in the Classroom*. New York, NY: Holt, Rinehart & Winston.

Roueche, J. E. & Baker, G. A. (1987). *Access and excellence: The open-door college*. Washington, District of Columbia: Community College Press.

Santalices, B. (2010). *El rol de las universidades en el desarrollo científico y tecnológico: Informe de la Educación Superior en Iberoamérica*. Santiago, Chile: Centro Interuniversitario de Desarrollo (CINDA), Universia. Retrieved from: < http://www.fundacionuniversia.net/fichero?id=1516>.

Scherer, J. L. & Anson, M. L. (2014). *Community college and the access effect: Why open admissions suppresses achievement*. New York, New York: Palgrave Macmillan.

Shannon, H. D. & Smith, R. C. (2006). A case for the community college's open access mission. *New Directions for Community Colleges*, 2006(136), 15-21.

Suresh, R. (2006). The relationship between barrier courses and persistence in engineering. *Journal of College Student Retention*, 8(2), 215–239.

Thompson, R. (2019). *Education, inequality and social class: Expansion and stratification in educational opportunity*. New York, NY: Routledge/ Taylor & Francis Group.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition, 2nd Ed*. Chicago, Illinois: University of Chicago Press.

Trostel, P. A. (2010). The fiscal impacts of college attainment. *Research in Higher Education*, 51(3), 220-247.

Tsao, T. M. (2005). Open admissions, controversies, and CUNY: Digging into social history through a first-year composition course. *The History Teacher*, 38(4), 469-482.

UNESCO (2015). *Draft preliminary report concerning the preparation of a global convention on the recognition of higher education qualifications*. Paris, France: UNESCO.

Universidad de la Republica (2016). *Ley Orgánica de la Universidad de la República, Ley número 12.549*. Montevideo, Uruguay: Universidad de la República. Retrieved from: <http://dgjuridica.udelar.edu.uy/ley-organica/>.

U.S. Department of Labor (2015). *Current population survey*. Washington, District of Columbia: Bureau of Labor Statistics. Retrieved from: <https://www.bls.gov/emp/chart-unemployment-earnings-education.htm>.

Vaughan, G. B. (1985). Maintaining open access and comprehensiveness. *New Directions for Community Colleges*, 1985(52), 17-28.

Villegas, C. G. (2019) *The youth unemployment crisis: A reference handbook*. Santa Barbara, CA: ABC-CLIO LLC.

Voigt, L. & Hundrieser, J. (2008). *Student success, retention, and graduation: Definitions, theories, practices, patterns, trends*. Iowa City, Iowa: Noel-Levitz.

Workman, E. (2012). Teacher expectations of students: A self-fulfilling prophecy? *The Progress of Education Reform*, 13(6). Retrieved from <http://www.ecs.org/clearinghouse/01/05/51/10551.pdf>.

Yang, F. M. (2014). Item response theory for measurement validity. *Shanghai Archives of Psychiatry*, 26(3), 171-177.

Autores:

*Dr. Wilson González-Espada
(w.gonzález espada@moreheadstate.edu)

Department of Mathematics and Physics, College of Science,
Morehead State University, Morehead, Kentucky, USA.
Phone: (1) 606-783-2927
Fax: (1) 606-783-5002

Dr. Eduardo Lacués (elacués@ucu.edu.uy)
Prof. Gabriela Otheguy (gabriela.otheguy@gmail.com)
Prof. Magdalena Pagano (mapagano@ucu.edu.uy)
Prof. Alejandra Pollio (apollio@ucu.edu.uy)

Department of Mathematics, College of Natural Sciences,
Catholic University of Uruguay, Montevideo, Uruguay.
Phone: (+598) 2487-2717

Dr. Rosina Pérez (rosina.perez@ucu.edu.uy)
Dr. Marcos Sarasola (marsaras@ucu.edu.uy)

Department of Education, College of Human Sciences,
Catholic University of Uruguay, Montevideo, Uruguay.
Phone: (+598) 2487-2717

* Corresponding author