International Journal of Innovation and Research in Educational Sciences
Volume 5, Issue 6, ISSN (Online) : 2349–5219

Computerized Adaptive Test to Guide Curricular Trajectories in Mathematics

Pilar Rodriguez Morales[1]
[1] Eastern Regional University Center, University of Republic, Uruguay.
Corresponding author email id: pilarod@gmail.com

Date of publication (dd/mm/yyyy): 06/12/2018

Abstract – **The computerized adaptive test (CAT) developed aims to obtain reliable and accurate information on the level of mathematical skills of incoming students, in order to guide them in the Mathematics courses to which it suits them to register according to their performance reached. Therefore, the objective of the test, in addition to classifying them at different levels of performance, is to provide guidance on the course that was recommended to take. Alternative curricular trajectories were designed according to the performance achieved in the test. To develop this CAT, a bank of items with parameters estimated by Item Response Theory; and a procedure to progressively select the best previously developed items and a method for estimating trait levels, were used. CAT provided a personalized report of each student, which allowed to improve the academic orientation on the part of the tutors' teachers. In addition, it was observed that the students used the test to make decisions about their trajectories autonomously.**

Keywords – **Adaptive Computerized Test, Curricular Orientation, Mathematical Test.**

I. Introduction

The evaluation of diagnostic type in the educational field has its origin in the need to characterize students and obtain information that can be interpreted by decision makers (Rupp, Templin and Henson, 2010). The objective of the diagnostic tests developed and applied by our university is to determine the degree to which students exceed the construct defined as competence in the area of mathematics.

Computerized adaptive tests (TAI) are tests constructed to assess psychological or cognitive constructs and are answered through a computer device such as a computer, notebook, tablet or smart phone. They have as main characteristic that they are adapting to the level of progressive performance that is manifesting the person, that is, as the user provides an answer, the test advances to a higher level of performance, if the answer is correct, and in case opposite, towards a lower level (Van der Linden & Glas, 2000). The fundamental elements to develop a CAT are the following: a bank of items with known psychometric properties, that is, parameters estimated by Item Response Theory (IRT); a procedure to progressively select the best items; and a method for estimating trait levels (Olea & Ponsoda, 2003).

The process of developing the first CAT to guide trajectories is described in Rodriguez, Perez & Luzardo (2017). To do this, we had a bank of items calibrated by TRI whose psychometric properties are presented in Rodríguez (2017), the procedure to select the items was developed by Luzardo, Padula & Forteza (2012) and the estimation method trait's was that of García et al. (2013).

The paper shows how, based on the personalized reports provided by the CAT, teachers and students make decisions about the academic trajectories in the first year of university.

II. Background

The antecedents of the difficulties in mathematics at the beginning of the university (Rodriguez, Diaz & Correa, 2014) raises the need to guide the trajectories of the entrance students. With the aim of remedying the shortcomings brought by the students, various alternative courses have been created (differentiated trajectories), some of them with propaedeutic character, others with a different timing than traditional courses. Teachers identify as a difficulty to orient the trajectories of students efficiently and also require to have, for mathematics, an instrument suitable for this purpose, since the fixed tests that were used until then were useful to characterize the skills of the students, groups or cohorts of students, but they did not have the same degree of precision to estimate the skill trait of each students (Rodriguez et al., 2015), which come from different formations, since they can enter students with any type of high school, and therefore, possess a wide variety of mathematical knowledge.

The objective of the test, in addition to classifying them at different performance levels, was to provide guidance on the Mathematics courses that would be recommended to take. Students who did not get enough in the test were recommended to enroll in 'Introduction to Mathematical Methods' and those who achieved sufficient score were suggested to enroll in 'Mathematics I'. The course 'Introduction to Mathematical Methods' arises as a proposal of an academic group of mathematics, to provide support in mathematics to the entrance students. It is an innovative proposal in terms of the proposed content. The objective of this course is to teach mathematics as a coherent and closely related system of ideas and procedures that are logically linked with the purpose of accelerating and deepening students' learning and creating a willingness to reason. The pedagogical perspective proposed considers that the deductive, historical and philosophical aspects of the discipline are fundamental for the cognitive development in mathematics.

In this way, the results of the CAT are used to know the performance level of students upon admission at university, and also, to guide them in curricular trajectories in mathematics. The CAT provides useful information for making curricular decisions.

III. Methodology

The mathematics TAI was designed on the platform of the Learnings Assessment System (SEA) of the National P-

-ublic Education Administration of Uruguay and is supported by the technological infrastructure installed by the Ceibal Plan in the educational centers. The tests are applied with a computer. This online format for the application of the tests allows to obtain results immediately.

The methodology used for the development of the tests involved the creation and establishment of content standards through expert groups. Difficulty and discrimination indices of each item were calculated as well as the distribution of the answers. The Cronbach alpha of the tests was found to know its reliability. To study the validity, the unidimensionality of the tests was analyzed. A principal component analysis with varimax rotation is performed to determine if the a priori dimensions are one-dimensional. The items characteristic curves and the difficulty and discrimination parameters of each Mathematics booklet are calculated. To calibrate the tests we relied on the IRT and the 2-parameter logistic model was used. The two Mathematics booklets were equated by anchoring items using the mean and deviation method. These methodological aspects of the creation of the test are documented in Rodríguez (2017). The test has good psychometric properties that account for its validity and reliability as an instrument for measuring competences in mathematics. In order to establish the cut-off point, the method proposed by Garcia et al. (2013). Students were classified according to their ability in three levels of performance (insufficient, sufficient and advanced).

The calibrated items were entered into the Learnings Assessment System platform to be available. The bank consists of 76 items of multiple options that admit a single correct answer. The procedure to select the best items adapted to the skill or trait of each student is an algorithm that in Uruguay only has this platform and was developed by Luzardo, Padula & Forteza (2012). The estimation of the trait is necessary for the functioning of the adaptive algorithm and is performed by maximum likelihood when there is variability in the answers and when it is correct or error is made by Bayesian estimation (Cheng & Liou, 2000). In this way, automatically, the platform estimates the trait of each student immediately after the response to an item and displays another item adapted to their level of trait. The estimates of the level of trait obtained in the different tests will be comparable through the common scale obtained by IRT (Olea, Abad & Barrada, 2010). There are as many iterations as are necessary to estimate the overall trait, and therefore, be able to calculate the student's score in the test.

CAT was much shorter than the fixed test applied in previous years, the extension of CAT varied between 19 and 28 items. This helps to avoid cognitive fatigue due to the lenght test which is a factor that can affect the performance, such us the results that find in a similar population Ackerman and Kanfer (2009).

The teacher who applies the test can monitor the operation and can finish it by verifying that all students have finished. Automatically the results are generated, so the teacher can see the level of performance of each of the students according to the CAT predetermined levels, which in our case were cut-off points obtained by the method of Garcia et al. (2013). The results are presented on a scale with a base of 500. Table 1 shows the cut-off points of each performance level transforming the scores on a scale of 500 points.

Table 1. Cut-off points at each performance level of CAT.

Level	Score
Insufficient	0-562
Enough	>562-706
Advanced	>706

The Mathematics TAI application was made in two cohorts of entrance students (2017 and 2018) before the start of the classes, in the period in which the student is enrolling in courses. The target population was the students who had to take mathematics in their curricular trajectory. It was applied in the computing rooms, performing different tours during the day.

IV. Results

An individualized report was prepared for each one of the students, such as the one presented in Fig. 1. It contained the student's personal information (name, personal identification number, career in which he / she registered, where he / she is studying) and also, a description of the competences according to their performance in the test. These reports were sent to the tutors teachers in charge of guiding the students' trajectory one day after they took the test. The tutor gave this report to the students, since they are the ones who make the decisions regarding the registration to courses.

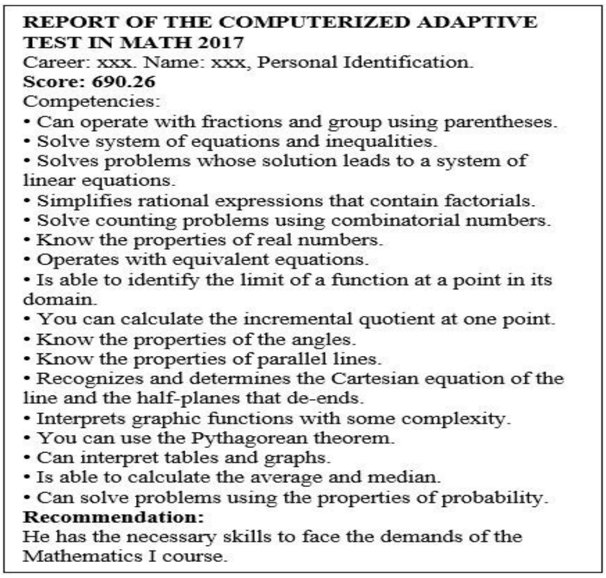

REPORT OF THE COMPUTERIZED ADAPTIVE TEST IN MATH 2017
Career: xxx. Name: xxx, Personal Identification.
Score: 690.26
Competencies:
• Can operate with fractions and group using parentheses.
• Solve system of equations and inequalities.
• Solves problems whose solution leads to a system of linear equations.
• Simplifies rational expressions that contain factorials.
• Solve counting problems using combinatorial numbers.
• Know the properties of real numbers.
• Operates with equivalent equations.
• Is able to identify the limit of a function at a point in its domain.
• You can calculate the incremental quotient at one point.
• Know the properties of the angles.
• Know the properties of parallel lines.
• Recognizes and determines the Cartesian equation of the line and the half-planes that de-ends.
• Interprets graphic functions with some complexity.
• You can use the Pythagorean theorem.
• Can interpret tables and graphs.
• Is able to calculate the average and median.
• Can solve problems using the properties of probability.
Recommendation:
He has the necessary skills to face the demands of the Mathematics I course.

Fig. 1. Example of a personalized report containing the description of the mathematical competences of a student assessed by the CAT.

This experience was carried out two years (2017 and 2018) and in both implementations the CAT fulfilled its objective, provided information on the level of performance of students in mathematics in the different dimensions that the test assess (numbers, operations, geometry and statistics).

It was expected that the CAT would be used by both teachers and students to guide their trajectories in mathematics upon admission at University.

The difficulties identified in the decision-making process were related to the interaction between students and tutors. Also, the final decisions made by the students were not expected in all cases. On some occasions the students did not go to the teacher tutor in search of guidance, although there were few cases, students who made decisions in relation to their career in mathematics without sufficient information were identified. Among the students who attended with the tutors teachers, cases were identified in which they did not take into account the orientation of the tutors, making decisions in the opposite direction to what was recommended. This behavior, although it is not the expected one, is part of the competences for which the tutorship teaches, that is, students have to make evident how they plan for themselves and take decisions (Colomer et al., 2013).

Fig. 2 shows the different decisions made by students based on the level of performance achieved in the test. They took three different types of decisions, followed the tutor's recommendation, did not follow their recommendation or did not attend the mentoring instance.

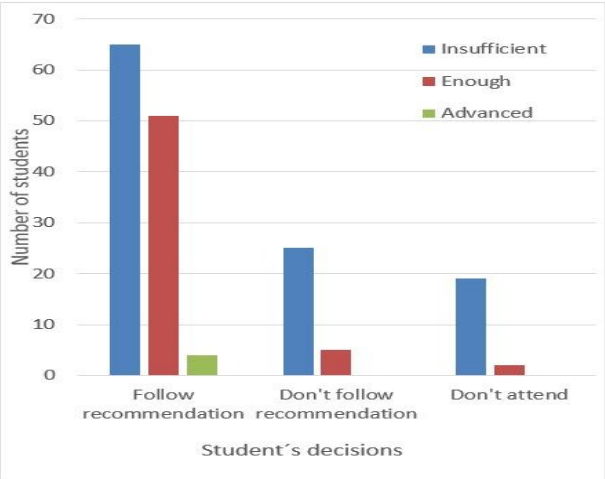


Fig. 2. Decisions of the students on the use of the information provided by the CAT according to the level of performance in the test.

This first study aimed to know the decisions made by students and make a first approximation to the reasons why they made those decisions. The vast majority of students followed the recommendation provided by CAT. 12 of these students were interviewed who did not follow the CAT recommendation and found that, having obtained an insufficient level of performance in the CAT, they enrolled in Mathematics I and conversely, students that having obtained a sufficient amount in the CAT, they chose to enroll in Introduction to Mathematical Methods. The first ones argued that if they did not take Mathematics I in the first semester they would delay the completion of their career, and the seconds justified their decision arguing that their score in the test was very close to the lower limit of the cutoff, so they believed they could also have difficulties. This seems to be a correct decision because it has been found in other groups of students a relationship between the performance in diagnostic tests of mathematics and the results in the subjects they take later (Shim, Shakawi & Azizan, 2017). In the latter case, an autonomous use of the CAT by the student is identified, which is desirable to promote, since it is one of the characteristics sought by university students, the autonomy in decision-making. Some of the problems described in this paper about the use of the CAT were already reported by Betts, Hahn & Zau (2011).

V. Conclusion

With the introduction of a TAI prior to the beginning of the courses, it was possible to determine the level of performance of the entrance students. From this, the teachers could provide academic guidance based on a valid and reliable instrument. This resulted in the improvement of the quality of this orientation, personalizing it and adapting trajectories to the needs of each student, in order to achieve an improvement in their subsequent performance in mathematics.

The standardized evaluation has been used, in general, to guide large-scale educational policies. However, in this case a standardized evaluation tool was used to guide the trajectories in a personalized way through the intervention of the tutors' teachers. There was also an independent use of the test, when students who interpreted the results and the information provided by the evaluation report were identified and opted to choose a path independently. On the other hand, we must emphasize that this experience put in articulation different actors and institutions, where each one of them contributed their capacities, experiences and resources for the achievement of the objective.

Problems were identified in the role some students assign to tutoring, because they do not adequately assess the possibility of having guidance in making decisions about their curricular trajectory in mathematics.

It remains for future research to follow up on the students evaluated and the results obtained in the mathematics courses in order to study if the CAT is a good predictor of their performance in the initial courses of mathematics at the University.

References

[1] A.A. Rupp, J. Templin, and R.A. Henson, *Methodology in the social sciences. Diagnostic measurement: Theory, methods, and applications.* New York, NY, US: Guilford Press, 2010.

[2] W.J. Van der Linden and C.A. Glas, (Eds.) *Computerized adaptive testing: Theory and practice*. Dordrecht: Kluwer Academic, 2000.

[3] J. Olea and V. Ponsoda, *Tests Adaptativos Informatizados*. [Computerized Adaptive Tests] Madrid: UNED, 2003.

[4] P. Rodríguez, G. Pérez, and M. Luzardo, *Desarrollo y aplicación del primer test adaptativo informatizado (TAI) de matemática para orientar trayectorias en la Universidad*. [Development and application of the first computerized adaptive test (CAT) of mathematics to orient paths at University]. In Jornadas de Investigación en Educación Superior, 2017. Available: https:// eventos.cse.udelar.edu.uy/indico/event/0/session/37/contribution/92.

[5] P. Rodríguez, Creacion, desarrollo y resultados de la aplicación de pruebas de evaluación basadas en estándares para diagnosticar competencias en Matemática y Lectura al ingreso a la Universidad. [Creation, Development and Assessment Tests Standards-based Application Results to Diagnose Math and Reading Skills to University Entrance] *Revista Iberoamericana de Evaluación Educativa*, 2017, 10(1), pp. 89-107. https: // doi.org/10.15366/riee2017.10.1.005

[6] M. Luzardo, D. Padula, and D. Forteza, *Validación y documentación del algoritmo del TAI de Ciencias y Matemática*. [Validation and documentation of the Science and Mathematics CAT algorithm]. Technical Report, 2012, Montevideo: DIEE-ANEP.

[7] P.E. García, F.J. Abad; J. Olea, and D. Aguado, A new IRT-based standard setting method: Application to elath-Listening. *Psicothema*, 2013, 25 (2), pp. 238-244.

[8] P. Rodríguez, M. Díaz, and A. Correa, Los aprendizajes al ingreso en un Centro Universitario Regional. [Learning at admission in a Regional University Center]. *Intercambios*, 2014, 2(1), pp. 91-100.

[9] P. Rodríguez, G. Carreño, T. Fernández, V. Figueroa, and N. Lorda, *Evaluación diagnóstica 2015 en Matemática y Lectura. Primer informe de resultados*. [Diagnostic assessment 2015 in Mathematics and Reading. First results report], 2015, Montevideo: CCI.

[10] P.E. Cheng, and M. Liou, Estimation of trait level in computerized adaptive testing. *Applied Psychological Measurement*, 2000, *24*(3), pp. 257-265.

[11] J. Olea, F. Abad, J. Barrada, Tests informatizados y otros nuevos tipos de tests. [Computerized tests and other new types of tests]. *Papeles del Psicologo*, 2010, 31 (1), pp. 97-107.

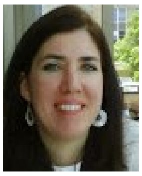
[12] P.L. Ackerman and R. Kanfer, Test length and cognitive fatigue: An empirical examination of effects on performance and test-taker reactions, 2009, *J Exp Psychol*, 15, pp. 163–181.

[13] J. Colomer; X. Vila; V. Saladó and R. M. Casellas, Tutoring as evidence of refelective practice: a case study. *Procedia-Social and Behavioral Sciences*, 2013, 93, pp. 356-363.

[14] G.T.G. Shim, A.M.H.A. Shakawi, and F. L. Azizan, Relationship between Students' Diagnostic Assessment and Achievement in a Pre-University Mathematics Course. *Journal of Education and Learning*, 2017, 6(4), 364-371.

[15] J. Betts, R.Y. Hahn, and A.C. Zau. *Does Diagnostic Math Testing Improve Student Learning*. San Francisco, CA: Public Policy Institute of California, 2011.

AUTHOR'S PROFILE



Pilar Rodriguez Morales has a Philosophy Doctor in Methodologies for the Design, Assessment and Improvement of Plans, Programs and Educational Projects by the National University of Distance Education, Spain. She is Professor and the Coordinator of the Teaching Support Unit of the Eastern Regional University Center, University of Republic of Uruguay. Her research area is the quality assessment of education, the performance assessment, the centers and educational programs assessment. She belongs to the National Research System of the National Research and Innovation Agency of Uruguay.