

RAFAEL VIDAL URIBE

MEASUREMENT OF LEARNING OUTCOMES IN HIGHER EDUCATION

The Case of Ceneval in Mexico

INTRODUCTION

In mid-2010, the National Center for the Evaluation of Higher Education in Mexico (Ceneval) received an invitation from the German Federal Ministry of Education and Research (through the Johannes Gutenberg University at Mainz and Humboldt University) to deliver a keynote lecture about the operation of the Mexican Higher Education Exit Assessments Tests (EGEL) at a conference entitled “Modeling and Measurement of Competencies in Higher Education” that was held in Berlin in early 2011. The following lines present the highlights of that presentation.

CENEVAL

Ceneval is a non-profit institution which was founded in 1994 by the National Association of Universities and Higher Education Institutions (ANUIES). Ceneval’s mission is the design and administration of tests and assessments for academic purposes. Ceneval’s tests target mainly students from high schools and universities. Since its inception, Ceneval has delivered a number of assessment reports, results and benchmarks to academic institutions, education officials and students. All reports produced by Ceneval within this project focus on the improvement of education. Ceneval does not perform institutional or program evaluations, nor peer-review evaluations, but serves hundreds of public and private academic institutions and several governmental agencies and administers its instruments all over the country, and in the U.S.A. and Ecuador.

Ceneval operates more than 240 different instruments, which are continuously maintained by more than 80 expert groups. Thousands of teachers and experts from all over the country are in charge of writing and reviewing the items. More than 3 million people take Ceneval tests every year. Around 1,800 testing events relating to Ceneval tests take place every year, and around 40,000 people take online tests using proprietary online software. More than 500 full-time staff work at Ceneval.

Ceneval’s instruments can be grouped into four broad categories:

- Admissions or entrance tests (EXANI);
- EGEL;
- Generic knowledge and skills instruments (EXDIAL);
- Instruments developed for other agencies (on demand).

In the following, we shall focus on categories 2 and 3.

HIGHER EDUCATION EXIT ASSESSMENTS (EGEL)

EGEL are tests for the evaluation of higher education (HE) learning outcomes. In the Spanish language, these instruments are called *Examen General Para el Egreso de la Licenciatura*, hence the acronym “EGEL”.

EGEL tests are standardized instruments which target students who are about to finish their higher education (a bachelor’s degree in North America or a Licenciatura in Mexico). These higher education tests have been developed for 33 different subjects or university programs. [Table 1](#) shows the 33 EGEL subjects.

Table 1. Complete list of EGEL tests in operation (as of April 2011)

<i>Life Sciences and Behavioral sciences</i> (11)		<i>Social Sciences and Humanities</i> (10)		<i>Engineering and Technology</i> (12)	
1	Agricultural Sciences	12	Accounting Education	22	Chemical Engineering
2	Biology	13	Business Administration	23	Civil Engineering
3	Chemistry	14	Communication	24	Computer Engineering
4	Clinical Chemistry	15	Economics	25	Computer Science
5	Dentistry	16	International Business Administration	26	Electrical Engineering
6	Medicine	17	Law	27	Electronic Engineering
7	Nursing	18	Marketing	28	Industrial Engineering
8	Nutrition	19	Pedagogy/ Education	29	Information Systems
9	Pharmacy	20	Social Work	30	Mechanical Engineering
10	Psychology	21	Tourism	31	Mechanical Electrical Engineering
11	Veterinary Medicine			32	Mechatronics Engineering
				33	Software Engineering

As previously mentioned, there are only 33 EGEL subjects, and although higher education institutions (HEIs) use to have many more different programs (some have up to 50 or 60 bachelor's degree programs or Licenciaturas), with EGEL tests a HEI could cover (assess) a very large part of the graduation pool. This is because EGEL tests target the higher education programs with the highest enrolment rate, and also because there are programs that, under different denominations, actually teach almost the same domains. For instance, in the case of the Business Administration EGEL Test (EGEL-ADMÓN), we found in Mexico up to 94 different higher education programs with different names that are closely related in purpose and subject; therefore, students from those 94 programs could all be evaluated with the EGEL-ADMÓN Test.

The 33 EGEL tests cover up to 75% of the country's graduation pool for one year. Up to 104,806 students from 490 HEIs in Mexico took EGEL tests during 2010. The EGEL Tests Program has been in operation since the creation of Ceneval in 1994. Over the past 17 years (1994–2010), 739,733 higher education students have taken EGEL tests.

The Nature of EGEL

The purpose of EGEL tests is to identify whether or not graduating students have the minimum knowledge, skills and competencies to go into professional practice. An EGEL test is a test of minimums (minimum knowledge required). It is the "floor"; therefore, the test is not an in-depth evaluation of each individual student, but instead provides a rough idea of what a student knows and is able to do in order to start his or her professional life.

EGEL tests are made to assess only the basics, and not everything which is taught in schools. EGEL tests offer students and HEIs an indicator of students' knowledge and skills.

EGEL tests are not a requirement for students, unless the HEI decides to make it so. However, on many occasions, EGEL tests have been administered to everyone who is graduating from a particular program. When every student in the graduation pool takes an EGEL test, the HEI gains a good indication of the performance of the program in question within the national context. As EGEL tests are not mandatory according to the law, HEIs are free to ask for the administration of these tests to their students. In general, HEIs pay for the administration of EGEL tests.

In a Nutshell, What is The Purpose of EGEL Tests?

- EGEL tests allow students to know whether or not they have reached a certain national standard set by a group of experts. If they achieve a satisfactory or an outstanding result, they also receive a Ceneval Diploma that can be used for job-seeking purposes.
- EGEL tests help principals, deans, program administrators and advisors to benchmark their higher education programs.

- For assessment and accreditation agencies, EGEL tests provide ways to improve the delivery of their recommendations and accreditations.

EGEL Characteristics and Development Processes

All EGEL tests are multiple-choice, nation-wide, domain-specific and not mandatory. As diagnostic tests, EGEL tests are all criterion-referenced assessments (Shrock & Coscarelli, 2007); this means that a student's performance is judged or measured against a standard or criterion. Moreover, this also means that the overall results are expressed not on a numerical scale, but on a categorical scale (outstanding, satisfactory or not yet satisfactory). The reports will be discussed in greater depth later in this article.

All items on the tests are oriented to practical situations. Items ask students to pronounce over situations that a young or fresh professional is likely to encounter.

In guiding the process of defining the construct domain, producing the test specification, building the item bank, creating and administering the actual tests, scoring the results and delivering reports, Ceneval follows the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999).

Each EGEL test is always supervised by a technical expert group (TEG) which is made up of university professors, industry experts and professionals from disciplines which are associated with the test subject. TEG members must reach an agreement on the minimum level that students must attain in order to enter the job market; consequently, nothing below that minimum should be acceptable. The TEG must endorse all of the activities and decisions that affect the test, and also in order to permanently enhance the test validity and reliability TEG must review statistical reports and evidence collected after every test administration.

Like many well-known tests (ACT, 2007; General Education Development, 2002, and so on), each EGEL test has a Technical Manual; within it, there is a complete description of the purpose of the test, its objectives, target population, the conditions required in order to take the test, the psychometric techniques needed to verify the quality of items and testing forms, algorithms used to obtain test scores, and any information needed to interpret the individual and institutional reports.

Trained specialists write EGEL items. Norms and specifications for writing the items are set by Ceneval and by every EGEL TEG (Osterlind, 1998; Downing, 2006; Haladyna, Downing, & Rodriguez, 2002). All items are developed in a proprietary bank platform (*e-BRAE*).

The qualitative review process is time-consuming; however, no EGEL items are ready for an operative test until they have been field-tested with representative student samples. Each item must meet the minimum quality criteria in terms of the difficulty and discrimination indices of classical test theory. In addition, tests with more than 300 examinees are analyzed using the two-parameter item response theory model (Crocker & Algina, 1986; DeVellis, 2010; Embreston & Reise, 2000; Raykov & Marcoulides, 2010).

As soon as an EGEL test produces data, Cronbach's alpha is calculated as a reliability measure (Nunnally & Bernstein, 1994). In 2010, 16 out of 33 tests reached alpha coefficients above 0.90, and the rest were above 0.80, indicating that the EGEL subjects have good levels of internal consistency reliability.

EGEL Administration and Scoring

Nowadays, almost all EGEL tests are administrated using paper and pencil forms, with only a few delivered online; however, all tests will be available in full online in 2012.

Ceneval organizes four national testing events every year in more than 60 sites all over the country. The printed test forms are administrated in two four-hour sessions in the same day (morning and evening). EGEL tests are usually between 200 and 250 items in length.

Proctors trained by Ceneval administer the tests. In order to control variables that could affect the test results, instructions for proctoring the tests are standardized across the administration events (McCallin, 2006).

As previously mentioned, all EGEL tests are criterion-referenced, meaning that the students' attainment is compared against a criterion which has been previously defined by the TEG. EGEL tests may have three, four or five subscales, and it is the decision of the TEG which combination of results from the subscales produces the overall result. For EGEL tests, there are only three possible overall results: *Outstanding*, *Satisfactory* and *Not yet satisfactory*. Student and institutional reports also show the students' performance on each test subscale. Test subscale scores are expressed in categorical and numerical terms. The numerical scale is called the Ceneval Index and has limits of 700 and 1,300 (600 points). Table 2 shows the equivalence of the Ceneval Index to performance levels for the subscale reports.

Table 2. Equivalence of Ceneval Index to performance levels

<i>Ceneval Index range</i>	<i>Performance level</i>
700–999	Not yet satisfactory
1,000–1,149	Satisfactory
1,150–1,300	Outstanding

EGEL Reports

After each test administration event, Ceneval prepares two types of report: individual (student) and institutional reports.

Individual report. Individual reports show the student's overall results and his or her results for each subscale of the test. The reverse of the individual report shows a description of the student's performance levels on each subscale. Figure 1 shows an individual report for the Medicine EGEL Test (front side). In addition, all students that achieve a satisfactory or outstanding overall result receive a Ceneval Diploma called a *Testimonio*.

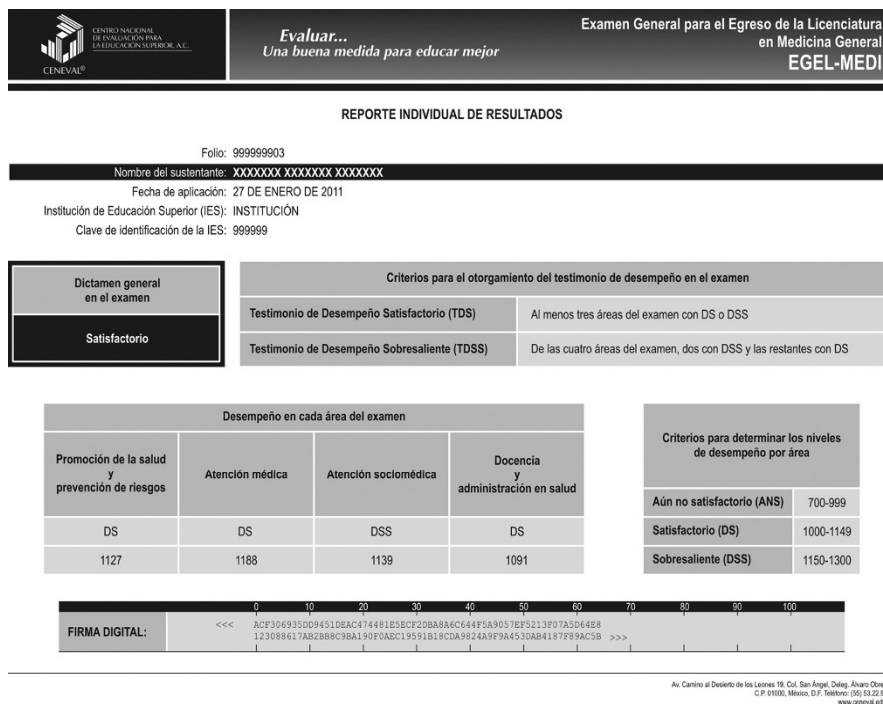


Figure 1. EGEL test individual report.

Institutional report. Institutional reports are very helpful tools for HEIs. Ceneval provides HEIs with reports of the performance of their students in the EGEL tests. Of course, information about the results of an individual student is sent only to her or his HEI. Every institutional report is a performance résumé for all of the students of one HEI. Institutional reports are sent to the program head or academic dean a few days after a testing event. No-one else receives the institutional report, as they are confidential.

Comparative institutional report. In addition, once a year, a complete comparative institutional report (CIR) with the aggregate results of all of the students who have taken EGEL tests during the past calendar year, is sent to every HEI involved in EGEL. This CIR shows the results of EGEL tests for all students and HEIs which have participated in the EGEL tests during that calendar year (more than 104,000 students and more than 490 HEIs went into the aggregated CIR for 2010); however, the names of all of the institutions are coded. The idea behind these CIRs is to provide a rough idea of the performance of a program in a specific HEI in relation to similar programs in other institutions.

Figure 2 shows an excerpt from a CIR (for medicine only).

MEASUREMENT OF LEARNING OUTCOMES IN MEXICO

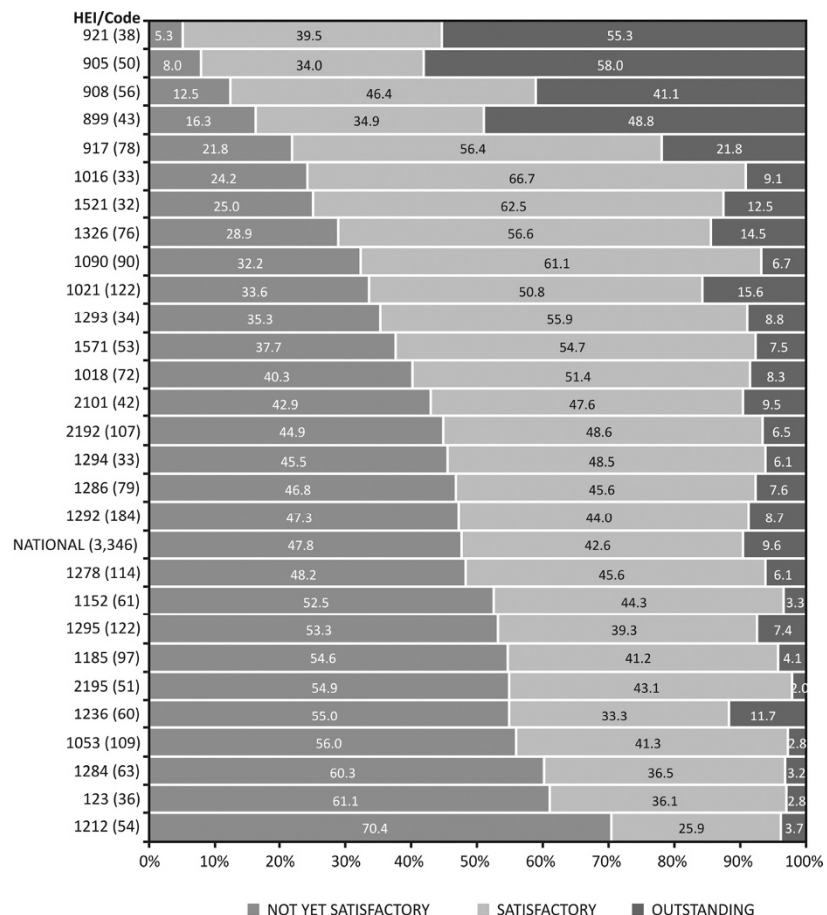


Figure 2. EGEL CIR.

GENERIC KNOWLEDGE AND SKILLS INSTRUMENTS

Ceneval also provides HEIs with a set of four generic knowledge and skills instruments; these testing instruments do not target students from a specific higher education program, but the students of several or all higher education programs. These tests are not necessarily taken at the end of the undergraduate program. These generic instruments are as follows:

Basic Science for Engineering Programs Test (EXIL-CBI)

This test was developed for students who are finishing the basic science subjects in the engineering faculties (it is usually taken at the end of the second academic year). Table 3 shows the content of the EXIL-CBI Test.

Table 3. Scales and subscales of the EXIL-CBI Test

<i>Content</i>	
1. Mathematics	1. Algebra 2. Calculus 3. Differential equations 4. Probability and statistics
2. Physics	1. Mechanics 2. Thermodynamics 3. Electromagnetism
3. General chemistry	1. Pure substances and mixtures 2. Chemical reactions

Statistics Test (EXTRA-ES)

The EXTRA-ES Test was developed in order to reach all students with statistical topics in their higher education programs. As not all students taking statistical courses take statistics at the same level, the EXTRA-ES test is organized with a common core recommended for all statistics students and three optional modules which are included at the discretion of the program director or student advisor. [Table 4](#) shows the content of the EXTRA-ES Test.

Table 4. Scales and subscales of the EXTRA-ES Test

<i>Common core</i>
Fundamentals of statistical thinking and data structure generation
Description, organization and interpretation of data
Notions of inference
<i>Optional modules</i>
Inferential methods and modeling
Sampling
Experimental statistics

Communication and Critical Thinking Test (ECCyPEC)

Communication and critical thinking skills are necessary for every student finishing an undergraduate program. Ceneval has developed a test that aims to measure these skills in a broad and general way. This instrument targets any undergraduate student towards the end of his or her years in higher education. [Table 5](#) shows the content (scales and subscales) of the ECCyPEC.

Table 5. Scales and subscales of the ECCyPEC Test

<i>Content</i>
Reading comprehension
Knowledge of written expression
Critical thinking

Written Expression Test (EEE-II)

The EEE-II is an essay test that aims to measure actual writing skills. It is not an objective test. Higher education students are asked to write an essay of two to three pages on a dilemmatic topic. Each EEE-II essay is scored by at least two expert Spanish teachers with the help of a rubric. [Table 6](#) shows the content of the EEE-II Test.

Table 6. Content of the EEE-II Test

<i>Content</i>
Conventions of language
Syntactic knowledge
Lexical variety
Thematic progression of the text
Global consistency
Planned speech
Information sources
Creativity

For more detailed information (in Spanish) about EGEL and the generic skills instruments from Ceneval, please go to <http://www.ceneval.mx/>.

REFERENCES

- ACT (2007). *The ACT technical manual*. Retrieved from http://www.act.org/aap/pdf/ACT_Technical_Manual.pdf.
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Crocker, L., & Algina, J. (1986). *Introduction to classical & modern test theory*. Belmont, CA: Wadsworth.
- DeVellis, R. F. (2010). *Scale development: Theory and applications* (3rd ed.). Applied social research methods series. Thousand Oaks, California: SAGE Publications, Inc.
- Downing, S. M. (2006). Twelve steps for effective test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 3–26). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Embreston, S., & Reise, S. P. (2000). *Item response theory for psychologists*. Multivariate application series. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- General Educational Development. (2002). *Technical Manual: 2002 series GED tests*. Washington, DC: American Council on Education. Retrieved from http://www.acenet.edu/Content/NavigationMenu/ged/pubs/TechnicalManual_2002SeriesGEDTests.pdf.
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), 309–334.
- McCallin, R. C. (2006). Test administration. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 625–652). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. New York: McGraw-Hill.
- Osterlind, S. J. (1998). *Constructing test items: Multiple-choice, constructed-response, performance, and other formats* (2nd ed.). Evaluation in education and human services. Boston, MA: Kluwer Academic Publishers.
- Raykov, T., & Marcoulides, G. A. (2010). *Introduction to psychometric theory*. New York: Taylor and Francis Group, LLC.
- Shrock, S. A., & Coscarelli, W. C. (2007). *Criterion-referenced test development: Technical and legal guidelines for corporate training and certification*. (3rd ed.). San Francisco, CA: John Wiley and Sons.

Rafael Vidal Uribe
National Center for the Evaluation of Higher Education (Ceneval),
Mexico