La evaluación mediante pruebas de gran escala en México

Evaluation by large scale tests in Mexico

Edgar Jesús Bautista Sánchez
Instituto de Ciencias de la Educación del Estado de México
flobasatoluca@hotmail.com

Resumen

El presente artículo es resultado del trabajo desarrollado al cursar los estudios de Doctorado en Ciencias de la Educación, siendo fundamental para integrar la investigación desarrollada sobre las tensiones que se generan en el docente de Educación Media Superior a partir de evaluaciones mediante pruebas de gran escala, que se han venido convirtiendo en el común denominador del sistema educativo de México en los últimos años.

El estudio desarrollado fue de tipo cualitativo y se realizó a partir de la obtención de narrativas a docentes, orientadores y directivos de un Centro de Bachillerato Tecnológico en el Estado de México, donde a partir del análisis desarrollado en el trabajo de campo se observa una tendencia por parte del docente a priorizar resultados de los exámenes de gran escala sobre los conocimientos del alumno, reflejándose en efectos colaterales de su práctica docente.

Palabras clave: Examen, evaluación de la educación, prueba de respuesta múltiple, maestro.

Abstract

This article is the result of the work developed while studying the PhD in Educational Sciences, being essential to integrate the research carried out about the tensions that are generated in the teaching of Higher Education based on the evaluations using tests of large

scale, which have been the common denominator of the education system in Mexico in recent years.

The developed study was qualitative and was performed on the obtaining of narratives from teachers, counselors and principals from a Technical High School Centre in the State of Mexico, where from the analysis developed in the field, there is a tendency by the teacher to prioritize test results of large scale over the knowledge of the student, reflecting a collateral side effects on their teaching practice.

Key words: Examination, evaluation of education, test of multiple-choice, master.

Introductión

This article is derived from the research work entitled "Tensions generated by teachers in Higher Education derived from the ENLACE standardized test", which was developed during my PhD studies at the Higher Educational Sciences Institute of Mexico State between 2012 and 2014. This article is comprised of three sections, the first considers the origin and development of the evaluation, the second the elements of the evaluation, the third concerns subjects of evaluation, the fourth concerns the examination and tests of large scale national and international, and finally some considerations as conclusions are presented.

Origin and development of the evaluation

One of the processes that relate the theme of education immediately is the assessment, which represents a polysemic term which has been transforming and adapting to circumstances over the years, favored by the public policies prevailing at that time.

The origin of the evaluation according to Izquierdo (2008), has been subject to intense debate held by many reviewers as Guba and Lincoln (1989), who placed the birth of

evaluation in North America after the first World War, linked to studies of education and public health, while Patton (1990) proposed as an area of research in the decades of the 1950s and 1960s.

The concepts as well as the methods of assessing have been changing as different paradigms of science have done. One of the initial evaluation models that predominated was the classic model, where Tyler stood out and its evaluation by objectives, for who, according to Izquierdo (2008), "knowledge of the objectives is what will guide the choice of the dependent variables and the criteria used for evaluation" (p.120). Carrion (2001), points out that Tyler calls it educational evaluation, referring to it as the degree of fulfillment of the objectives that were previously specified and applied. However, is now observed that the current assessment is no longer based on objectives, seeking thereby the fulfillment over different competencies.

In the Decade of the fifties, Carrion (2001) pointed out that it is still considered that evaluation in education was associated exclusively with the student's learning, reason why Benjamin Bloom proposed to "complicate" the idea of the educational objectives, incorporating psychology into human development to make them more comprehensive, which gave rise to three types of objectives with direct correspondence to three domains of personality: cognitive, affective and psychomotor. Each of them is present in Latin American educational systems, educational competences pretended by their educational systems.

In the sixties, Carrión (2001) reports that from a lot of proposals on evaluation models arises what has been called evaluative research, which proposes that the assessment is a systematic process of making judgments based on a description of the differences between the performances of students and teachers on learning standards.

Carrión (2001) reports that in the seventies Daniel L. Stufflebeam defined educational evaluation from its relation to planning, as an activity that provides information for decision-making such as structuring processes, execution and recycling, to give birth to their evaluation model considered integrate the context, inputs, processes, and ultimately products. Of these, only the context assessment was considered continuous, while others

depended on decision processes. Here the context is beginning to take a leading role in the assessment to consider factors that influence favorably or unfavorably.

Alicia de Alba (1991) noted that in the seventies "begins to emphasize program evaluation, to consider it as an important area within the field. It is with the rise of program evaluation that this practice is beginning to take root in different sectors and levels of society, so that programs of institutional, sectoral and national "(p.80) Evaluation begin to develop. This began evaluating everything related to education, involving both students, teachers, programs, curricula, institutions, etc., all under the pretense of educational technology, allowing validate and legitimize decisions made on evaluation, but hid behind his speech a control problem, with implications not only in education but also in social matters which were beginning to glimpse impact on the discipline and the perception of individuals when faced with various problems social.

With respect to the decade of the eighties, Alicia de Alba (1991) reports that the Nile was a consideration in stating that saw evaluation as a "delicate or very sharp tool (axiological)" (p.79), indicating the main problem specifically for Latin American countries was that not read your reality. Meanwhile, if he could perceive this reality because he saw from outside as belonging to one of the countries of the hegemonic-dominant group of the OAS.

The control problem appears again in this decade. It is assumed that the evaluators in Latin America, the filmmakers repositories should apply and report, provided technical evaluation functions. From a series of preset criteria and requirements aim to show the state of things as natural. Alicia de Alba (1991) puts it promptly when he says "if the decision is in the hands of some sectors or groups, it is they who exercise control; the evaluator builds on a number of criteria and requirements already built, and his work is reduced to the operationalization "(p.88).Desde los años noventa en México:

Multiple forms of assessment are promoted in education in terms of institutions, academia, research programs and teaching and learning in school. It applies to all educational levels (from kindergarten to graduate), and has different manifestations such as accreditation, certification, accountability, performance, entrance exams and graduation ... and from different approaches such as quality, effectiveness or suitability. (Glazman, 2005, p.24)

These models have increased regarding the evaluation, trend of global education policy, where the countries of the hegemonic-dominant group indicate the strategies to follow, the parameters to consider as well as assessment methods and their short-, medium- and long term.

For the new millennium, the evaluation continues with individual treatments lacking communication between analysts, which has caused further evaluation in response to particular demands based on specific needs. It has improved educational practices, as Glazman (2005) points out, "in terms of definitions of the subjects of evaluation, problems encountered, clarification of criteria and the reports themselves" (p. 15), that meet these particular demands and needs specific, thereby leaving relegated the formation of specialized evaluators who can team up with their peers to generate significant progress, but in the part of the evaluation are isolated and continue spraying knowledge.

Evaluation items

From different positions of evaluators in recent decades, the management of the evaluation has several uses, reaching even reduced to something technical to prioritize the quantitative part during the process, a situation that has been playing in recent years with influence of behavioral-experimental psychology in technical and administrative fields of society, or in other words, attached to a behaviorist, positivist approach.

However, there is another position on the assessment that it is necessary to recover the comprehensive approach of prioritizing the qualitative part of that evaluation process. This lets focus on results and prioritizes processes; recognized from the outset that the evaluated is a person, not an object, why the assessment is made explicit from the notion that handles Alicia de Alba (1991). Hence the notion of evaluation is constituted by two elements: the theoretical-conceptual understanding and axiological assessment.

a) Theoretical-conceptual understanding

Alicia de Alba (1991) notes that evaluation in education is a "complex process of reflection, critical analysis and evaluative conceptual synthesis, from which it knows, understands and values" (p.94), this concept seems distanced fashion as has been driving assessment in

recent decades by the Mexican educational system, which was minimized or left out the human-social, self-interpretation of the Aristotelian tradition (or verstehen-understanding) and it has given itself increasingly to the functional and mechanistic importance of the natural and exact sciences Galilean tradition (or erklären).

Consider the theoretical element of evaluation is important. Who conducts the assessment should have a strong conceptual foundation of what is to be evaluated, ie understanding, a critical analysis from a theoretical position to understanding what evaluated as lacking this fall in the prosecution and not assessment, a situation that would lead to arbitrary and authoritarian decision-making.

If you consider the theoretical element, the person doing the evaluation has the minimum elements necessary to recognize each of the parts that make up the appearance evaluated. Therefore, if the evaluator evaluated on large scale tests must have sufficient theoretical and conceptual understanding in regard to large-scale tests to let you know what is being evaluated.

b) Axiological rating.

The second element is intended to clarify the specifics of the evaluation, and where intervening values such cultural, social, political, economic and ideological, which eventually form the judgment with which the evaluation is performed, which is linked with tradition Aristotle, where the task of evaluation is to try to give answers to questions like why evaluate evaluate ?, Who ?, to whom be evaluated? It refers to a set of values from which the assessing and commitment subject-object that the evaluating body acquires the subject is discussed, condition or process that is under evaluation, because eventually the evaluator is a political subject that through this work establishes commitments to those who evaluate and those who have entrusted the.

The axiological nature of evaluation consists of various dimensions that help integrate a judgment about what is assessed, which are: cultural, social, political, economic and ideological, since the position of Alicia de Alba (1991), would be understood by as follows:

• Cultural dimension: Refers to the various cultural formations on which the valuations are constructed to avoid setting cultural contradictions between the evaluators and the evaluated instances.

• Social dimension: This refers to the (international, national, regional) wide social structure and social organization of the various groups, sectors and institutions.

• Political dimension: Consider the valuation or rejection by the evaluator instance of political-social project that will be developed as well as the specific way in which affect the process.

• Economic Dimension: Refers to the controversy between structure and superstructure, between determination, mediation and relative autonomy, being one of the arguments used to justify changes in educational systems because in them the role played by the underdeveloped countries is defined.

• Ideological dimension: it is given the task of explaining the different ways of conceiving ideology as a set of ideas about the world and that meet the interests of any social formation involving the cohesion of a group guiding their behavior towards those ideal.

While all of these five dimensions of axiological nature of the evaluation manage a set of core values that have interference at the time of the evaluation, note that the most important thing they bring to it is the importance they give to contextualization as a part closely linked to the assessment process, as evaluators and evaluated are people first, like it or not linked to a reality governed by these dimensions.

Treatment areas evaluation

There are two main areas of treatment evaluation, pedagogic and administrative, from the work done by Raquel Glazman (2005), are perceived as follows:

Pedagogic.

Ancestral origins, acquired momentum with behaviorism. In this area, the evaluation is intended to allow students to learn and it is based on theoretical and conceptual elements of education, being a process of clarification and search data that allow precise and weigh

situations or originated early analysis of learning, later to correct errors, review goals and change where appropriate methodological proposals, with the idea of improving the conditions of education.

This is performed if the following series of steps considered its support is true:

Set an object to be evaluated.

Make a judgment against a standard.

Collect information.

Analyze the information.

Systematize information.

Evaluate the reality evaluated.

To express this value.

And from the above steps the purpose intended by the pedagogical nature of the assessment is satisfied.

Administrative

Becomes important in the eighties, where assessment is associated with accountability, accreditation and certification, and aims to establish hierarchies, propose comparisons and make decisions regarding the inclusion and exclusion of subjects, based on these reports applied evaluations situation to be observed in detail in large-scale tests.

Within this position papers related to educational issues highlights the socio-economic approach of education systems, operating under economic criteria of production, efficiency and control, to allow preparation skilled labor under the idea of raising the rates of development of the country question.

This position, according Glazman (2005), is related to the economist character that has taken Mexico's education policy and developing countries, where reduction of government

resources for education is observed on the pretext of the crisis the corresponding education system.

The examination and testing of large national and international scale

There are multiple ways to assess and each needs at least an instrument to carry it out. In our case we refer to the assessment tool called specifically to review and scale tests.

When thinking of a test, people will relate to educational activities and specifically with the evaluation process, you might even think that the same relationship has been part of the history of the examination, however, Angel Diaz Barriga (1993) demonstrates the falsity of this claim on three points:

First, the test was an instrument created by the Chinese bureaucracy to elect members of lower castes. Secondly, even before the Middle Ages there was no evidence of an examination system linked to educational practice. Third, the note (grades) is a legacy that the twentieth century left to pedagogy, which brought many problems still suffer. (Díaz Barriga, 1993, p.13)

Diaz Barriga (1993) notes that "the test is a space where a multitude of investment in social relations and teaching are done" (p.14). Specifically there are three investments: one that makes social problems in teaching, one that turns methodological problems problems one exam and a final which reduces the theoretical problems of education to the technical scope of the evaluation.

In terms of Foucault (2003), this implies that the review is a space that reverses the relations of knowledge and power. Knowledge relationships really are power relations. This causes the assessment lacks neutrality and then exams conducive to contemplate another difficulty that the ratings that generates hardly be considered as objective and fair for its large number of personal subjectivities, academic or institutional. This makes it difficult for this instrument reflects all the qualities of learning, making that qualification only serve to justify to society and assigned authorities.

Foucault (29003) concerns the impact of power and embedded control in the evaluation process, in its "Discipline and Punish" where referred to as a set of relations of submission,

objectification and standardization from the school, saying that "examination is the technique by which power, instead of making the signs of its power, rather than imposing their brand to their subject, keeps these in a mechanism of objectification "(p. 174). This relates as Foucault (2003) considers that the review combines two types of techniques, the first is the hierarchy, used to monitor, and the second is the penalty, which normalizes, allowing therefrom qualify , classify and punish. Despite this being a study by Foucault on examinations of the seventeenth and eighteenth centuries, can be extrapolated and attached to the current reality of the Mexican educational system.

With regard to large-scale tests, these were developed in the United States in the mid-twentieth century, are type tests used to assess learning achievement in conceptual and procedural part to a large number of students. Have been recommended for use in the education systems of many countries, international agencies, particularly the Organization of Trade and Economic Development (OECD) and the World Bank, where the purpose of these tests was to support countries in improving their school system, looking for a better preparation of young people for entry into changing, globalized world. As Master (2006) points out, this allows comparisons of the results for "assess the effectiveness and performance of the adopted educational policies, analysis of the evolution of the system and its comparison with other educational realities" (p. 315).

Among the evidence of large-scale international type is the Program for International Student Assessment (PISA), in which Mexico, since he has participated, has remained at the bottom. Only in 2012, in mathematics is ranked 53rd among 65 countries, while in Spanish it is ranked 54th among 65 countries.

Therefore, it was originally thought to reviews as a tool supporting education, is now entering a phase in which the tests are prioritized, especially exams scale, implementing actions to raise their results in the short medium term, although this does not really reflect increased knowledge of students, but the indicators of these emerge, as these results are linked in some way by education authorities and media with the performance of teachers and institutions where they work.

The original idea of using large-scale testing has been perverted as it relates Padilla (2009): "exam with their massive application has ceased to be an instrument of teaching to become an instrument against teachers" (p. 44). You can see how consciously and unconsciously have been implementing similar activities as resolution tests, looking for familiarization to them by students, rehearsals of themes addressed in previous examinations in times of class, and so on. This priority is given to results linked with economic incentives and the labeling of teachers to authorities, public and private organizations with interference in education. For this reason, recent work by Glazman (2005) it is assumed that large-scale tests reflect and reproduce the economic positioning and globalizing ideology of hegemonic-dominant groups, when they consider globalization "in its socio-political and economic dimension acquires specific ideological connotations that determine courses of action with special effects on education "(p.19).

Two of the most frequent criticisms of this type of evidence is, first, decontextualization. It is assumed that all students have equal educational purposes, same social and academic conditions, using the same tests for students with different realities. On the other hand, is the competitive environment generated by these tests between students, teachers, schools, states and even countries, following the rankings of international tests such as PISA, or applied nationally in Mexico as the National Assessment Academic Achievement in Schools (LINK), which since 2014 suspended its application at the level of basic education, mainly because they have been linked to unethical practices during implementation by teachers He said education it related to economic incentives for teachers of students with better results.

Also unchanged at Level Superior Middle arguing that at this level is not vitiated the examination from the perspective of the authorities of the Ministry of Public Education.

While this level is not the ENLACE test relates to financial incentives for teachers, there are points of alert in the management of results by the educational authorities, who eventually to conflict the teacher in a classroom, as it prioritizes indicator scale testing against the knowledge of the course.

Conclusions from the results of large-scale tests

As a result of the implementation of large-scale tests in the assessment process and especially the Mexican educational systems, recurrently observed the following effects:

1. The test results are used for the convenience of public and private organizations.

2. The test results are prioritized over student learning.

3. familiarization of these large-scale tests on students is encouraged.

4. It has become an instrument of control and accountability.

5. The teacher's role is reduced to implement and deliver reports.

6. The standardization in testing assumes equal pupils in different realities.

Large-scale tests in education of Mexico have been distorted from its original intent, to be overwhelmed and prioritized by the demands of dominant hegemonic groups represented mainly by the World Bank and the OECD, gradually promoting educational systems failed before their parameters and criteria, causing collateral damage in countries that are members of these groups or bodies. In the pursuit of these countries continue to have privileges primarily of an economic nature with capital injections to certain educational sectors will continue giving priority to evaluating student learning.

The shortcomings identified exams scale rarely translate into proposals for improvement, because it has simply been chosen to implement remedial strategies that result in short periods of time on improving indicators, which ultimately only reflect a simulation or fantasy of an educational system that is deficient and operates solely to meet the requirements of the dominant hegemonic system.

Bibliography

Carrion, C. (2001) Values and principles for evaluating education. Mexico: Polity Press.

De Alba, A. (1991) Curriculum Evaluation. Conceptual shape of the field. Mexico: CESU-UNAM.

Diaz, Á. (1993) Examination. Texts for its history and debate. Mexico: Plaza y Valdés.

Ferrer, G. Arregui, P. (2003) International learning tests in Latin America and its impact on the quality of education: criteria to guide future applications. PREAL. No. 26, pp.1-13.

Foucault, M. (2003) Discipline and Punish. Birth of the Prison. Argentina: Twenty-First Century editors.

Glazman, R. (2005) The faces of educational evaluation. Mexico: PAIDEIA.

Guba, EG, Lincoln (1989) Fourth Generation Evaluation. London: SAGE Publications.

Maestro, M. (2006) Evaluation of the education system, Journal of Education extraordinary. Spain: Ministry of Education and Science, pp. 315-336.

Izquierdo, B. (2008) From the classical evaluation pluralistic evaluation. Criteria for classifying the different types of evaluation, to kinship magazine Methodology of Social Sciences, (16) pp. 115-134. Recovered from http://www.redalyc.org/articulo.oa?id=297124024005

Padilla, R. (2009). International and national mass screening. Dating or misunderstandings ?, ¿Profiles Education, 31 (123), pp. 44-59. Retrieved November 6, 2013, in http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-26982009000100004