EVALUATION OF NUMEROUS GROUPS OF STUDENTS IN MATHEMATICS

G. Leví and E. Ramos[1]
UNED - SPAIN

ABSTRACT In this work we present some ideas about the evaluation of a numerous group of students in the field of Mathematics. The tool used is a multiple choice test and we discuss some ways to formulate non trivial and significant questions. Also we present a computer approach to automatic formulation for these questions and to calculate the right answer as well as the distractors.

1. INTRODUCTION

The main objective of an educational institution is the formation of those students who come to their classrooms searching the knowledge that the competitive society demands from them. The measurement of the level of formation reached is one of the basic pillars of any educative model. Society demands the educative institutions precision in the diplomas they grant, therefore it is necessary to design adequate models of evaluation that guarantee the degrees that properly fulfil the quality standards.

Evaluation is, without any doubt, one of the most important problems that pedagogy faces. The type and goals of the course, the subject to evaluate and its extension, the number of students and professors, etc. etc. these are elements to consider when designing an evaluation model. Although it is known that most teaching professionals have solid knowledge on the methods used to evaluate their students, acquired in good part throughout years of experience, some considerations on the procedures and alternatives of evaluation are useful.

Several ingredients take part in the design of an evaluation model that can make a certain procedure more effective, or more feasible, or both. Specially when the group of students to evaluate is big and the number of teachers available and the of time in which the evaluation has to be done are limited. To design an evaluation method fulfilling standards of quality, reliability, suitability, homogeneity, etc. is a complex problem which becomes even more serious, when what we try to evaluate is Mathematics.

Here, we try to share with teachers of Mathematics and related disciplines, some ideas, that might serve as a starting point to design appropriate models of evaluation for numerous groups of students in these fields. These ideas emerge mainly from the experience of the teachers in charge for the past fourteen at the National University of Distance Education of Spain of Basic Mathematics and the Foundations Course for people aged 25 or older. Nearly twelve thousand students from all over the world are enrolled each year, following the methodology of the distance education.

2. - THE EVALUATION

Evaluation in pedagogy is understood as a mean to obtaining information on the nature and quality of what a student has learned. Ideally, this process should be carry out in a systematic and continuous way, although, sometimes precise and isolated information can solely be available. Evaluation is a natural demand of any educational process for many reasons. For example, the need to evaluate is helpful to make the proper arrangements to adapt knowledge to the social demands, to maintain academic standards, to asses students, to fit contents and the educational goals, etc.

The means by which evaluation is made are diverse, depending on factors such as number of students, objectives of the evaluation, etc. However, commonly, the student faces a certain test or reactive that the professor proposes the student so that he can show the obtained knowledge, aptitudes, abilities, etc.

The design of quality evaluation tests, that properly discriminate and provide trustworthy and precise information on the student is, without a doubt, one of the main challenges whereupon a teacher must face in his professional activity. A good test of evaluation must include excellent and original questions, be written up in a clear and univocal way, and display a difficulty of response proportional to the goals pursued. The search of all these qualities can make the task of designing evaluation tests an arduous one and a considerable amount time and effort has to be spent for their accomplishment.

[1] Depto. Estadística e Investigación Operativa, UNED
Calle Senda del Rey 9, 28040 Madrid, Spain
e-mail: eramos@ccia.uned.es

2.1 Evaluation tests Evaluation tests can be classified in two great groups: objective tests and essay tests. The essay tests are free answer tests, usually without space limit, although, at times, limited space is mandatory. These are suitable to raise questions concerning theoretical aspects, exercises and problems, interpretations, remarks, personal comments, summaries, etc.

Among the advantages of the essay tests we can mention that they are easy to prepare, they allow to appreciate the creativity, originality, capability of organization and expression of the student, encourages a depth study and are almost impossible to be answered by chance. Nevertheless they are not free of disadvantages. In general, they are difficult to evaluate, some questions can allow different answers and, consequently, different correctors can evaluate differently, influencing even in the evaluation the mood and fatigue of the corrector, the number of students and the mean level of the group, the handwriting and presentation, etc. Therefore they should be graded by the professor in charged. Furthermore they stand badly statistical analyses.

The objective tests outstand mainly because their evaluation does not depend on the judgment of the corrector, admitting even a mechanical evaluation, by hand or by computer. Among the common features of these objective tests is that they use a great number of questions or items, are of brief formulation, the correct answers are unquestionable and they might be answered in specific sheets to be read automatically. The objective tests commonly used are the so called multiple choice tests. In this type of tests each item has an initial proposal, a correct answer and several false alternatives to distract. In the initial proposal the necessary information is offered to find out the answer, through a question or by formulating incomplete sentences. Options are offered, one of which is correct being the rest *distractors*. In order to identify the correct alternative various strategies can be used. For example, to indicate only the correct answer, indicate the best one among several answers than can be correct, indicate the incorrect answer or the one that does not bear relation to others or even indicate several alternatives that can be correct.

The advantages of multiple choice tests are several. First of all they allow to cover large contents of the study program and to measure knowledge, applications, analysis, etc. without the influence other factors like hand writing, spelling, style, display, etc. Secondly, usually they are attractive to students and easy to answer; in addition they are easily corrected using optical readers and, therefore, the evaluation is obtained quickly. Thirdly, they stand good statistical analyses allowing to discriminate levels of competence among students. These positive aspects of the multiple choice tests have turned them into a popular instrument of evaluation. Nevertheless, it cannot be ignored that they also display important disadvantages. In the first place, it is necessary to keep in mind that some contents can be difficult to evaluate using these tests, for they do not measure creativity, style, capability of organization and synthesis, originality and other qualities of the sort. Secondly, designing the test is laborious, because if superficially prepared can cause bad study habits, inviting to learning by heart, and finally the correct answer can be obtained by chance, making it necessary to use evaluation formulas to counteract this factor.

3. DESIGNING OBJECTIVE TESTS IN MATHEMATICS

In the scope of Sciences, like Mathematics, commonly used evaluation tests are essay test. The traditional activities and problems proposed in text books correspond to the model of essay tests. Nevertheless, in Mathematics the use of objective tests is a fact. With an intelligent exposure of the questions it is possible to design an evaluation test of the objective sort which demands on students an effort to similar to the one required in the traditional one. Not only in tasks about concepts, questions of true/false type, but also to problems corresponding to classic schemes within the Mathematics. We are going to display some basic ideas extracted of our own experience professors of a course known as Basic Mathematics on the Foundations Course for people aged 25 or older.

EXAMPLE 1

A traditional exercise on rational numbers is to obtain the decimal expression to the fractional expression of a number. The classic statement usually is as follows:

Find a fraction that represents the decimal number 0.5037037037...

In order to solve this exercise, the student must know the procedure to look for a fraction whose decimal expression is a mixed-periodic number like the proposed one. The answer demands the accomplishment of certain calculations that lead to solution 68/135.
The exposure of this exercise in an objective test could be as follows:

The decimal number 0,5037037... is represented by the fraction

a) 68/135
b) 70/135
c) *62/225*

This statement has the inconvenience that the correct answer can be obtained by simply dividing each one of the possible answers indicated in each one of the alternatives. Therefore we could not be sure weather or not the student knows how to calculate the fractional expression of a mixed-periodic number which is precisely sort of knowledge to be evaluated.
A different statement from the previous question that it allows to obtain the desired objective could be the following one.

The decimal number 0,5037037... is expressed by a fraction

a) With numerator equal to 68.
b) With numerator equal to 70
c) With denominator equal to 225.

Notice that the only correct answer is alternative a) and to find it out, it the student must find fraction 68/135 through the general procedure search of the fractional expression of a mixed periodical decimal number.

3.1 Automatic processing of objective tests in Mathematics

One of the advantages of multiple choice tests is that they can be automatized, not only as far as evaluation is concerned, but also its processing, remarkably simplifying the task of designing the evaluation tests.
A first simple mechanism of automatization in designing objective tests consists in having teachers prepare a more or less broad number of questions, introduce them in a data base and by some random procedure extract as many forms of examination as needed. The random procedure can add different criteria that might lead to a balanced test. For example, a certain percentage of questions of each one of the lessons of the program might be desired. Also, it can be interesting to handle data about the frequency of successes, errors and blanks of each one of the questions in past courses to prepare a test with a predetermined level of difficulty. Autoevaluation procedures can be offered to students, complemented with reinforcements, addresses, etc. All these ideas lead to a greater quality evaluation.
Nevertheless the idea of having a data base of evaluation questions is not free of inconveniences. When the questions are fixed, there is a risk, students might learn them by heart and answer them mechanically. Therefore it could be interesting to devise mechanisms of automatic formulation of questions to avoid memorization of the questions and its answers. Many ideas can be used; in synthesis, they happen to look for a parametrically way to construct the statement of the question, the correct answer and the diverse alternatives. Variants of the question can be obtained by modifying the parameters, resorting to the random selection. Here is an example of a *parametric question* used in our courses.
EXAMPLE 2
The question of the previous example is a particular case of a parametric that we are going to describe next.
In the first place, two prime numbers are set for example 3 and 37. Then the following sets, whose elements are going to act like addends, factors and exponents to create the variants.
X = {0,1,2}; Y = {0,1}; Z = {1,2}; K = {1,2,3,...,100}; P = {0,1,2}; A = {1,2,4,5};
U = {-1,1}; V = {-2,2}
With these data we calculated the following decimal number

$$\frac{3^{x} \cdot 37^{y} \cdot (3k+z)}{999 \cdot 10^{p}}$$

where $x \in X, y \in Y, k \in K, z \in Z, p \in P$ are selected randomly. The statement of the question can then be written up as follows:

The decimal number $\dfrac{3^x \cdot 37^y \cdot (3k+z)}{999 \cdot 10^p}$ *is expressed by a fraction*:

As correct alternative any of the two following answers can be chosen

a) With numerator equal to $\dfrac{a \cdot 3^x \cdot 37^y \cdot (3k+z)}{m}$

a') With denominator equal to $\dfrac{a \cdot 999 \cdot 10^p}{m}$

being $a \in A$, randomly selected and $m = m.c.d.(3k+z, 3^{3-x} \cdot 37^{1-y} \cdot 2^p \cdot 5^p)$

The reason of introducing these correction factors is to avoid the appearance of numbers such as $999 \cdot 10^p$ that quickly would indicate the correct answer.

As distractors the following alternatives can be used:

a) With numerator equal to $\dfrac{a \cdot 3^x \cdot 37^y \cdot (3k+z)}{m} + v$

b') With numerator equal to $\dfrac{a \cdot 3^x \cdot 37^y \cdot (3k+z)}{m} + u$

c) With denominator equal to $\dfrac{\frac{a+v}{3} \cdot 999 \cdot 10^p}{m}$

c') With denominator equal to $\dfrac{\frac{a+u}{3} \cdot 999 \cdot 10^p}{m}$

where, $u \in U, v \in V$ selected randomly also cause results that do not include tips on the correct alternative.

4. CONCLUSIONS

To design quality evaluation procedures is one of the most important challenges the teaching staff must face throughout his professional life. When the group of students is numerous, teaching staff reduced, the objectivity values and the fastness in the obtaining evaluations lead inevitably to using objective tests. Although the traditional evaluation system in Mathematics usually uses essay tests, it is not difficult to reframe many of the classic exercises in objective tests without pursuing certain evaluation goals. In addition, the evaluation tests thus conceived can be obtained automatically, letting off trite remarks and repetitions that could distort the aimed objectives.

REFERENCES

García Aretio, L. (1994): *Educación a distancia hoy,* UNED, Colección Educación Permanente.

Medina Rubio, R., T. Rodríguez Neira and L. García Aretio: *Teoría de la Educación,* UNED, Colección Unidades Didácticas.