Earli | 11th Biennial Conference Earli 2005

Submitted Abstract

The Abstract is already submitted. You can view information in Read Only Mode.

Proposal Type:	Individual Paper
Domain:	Assessment and Edumetrics
SIG:	Assessment and Evaluation
Type	Submitted Paper
Equipment	Overhead projector Internet access

Paper Details

Title	Development of edumetrical indices for the analysis of the spectral quality of higher education standardized tests
Abstract	The confidence degrees technique associated to the MCQ makes it possible to bypass the "binary" character of students assessment performance (the selected proposal is either correct, or incorrect) provided that a series of methodological rules are followed called "admissible probability measurement procedures" by Shufford & al. (1966). Usually, the confidence percentages which accompany the MCQ answers are used to deliver more subtler feedback on each student's spectral performances. The innovative aspect of our approach lies in the fact that we have exploited the confidence percentages provided by the students to provide spectral information on the quality of the questions (as opposed to information on the quality of students performances). Our research thus led to the development of a series of original indices for the analysis of assessments' spectral quality. These spectral indices are intended to be used when the assessor must highlight problematic MCQ and, within those, the proposals which contain anomalies.
Summary	Our starting intuition for the construction of these new indices is as follows: logically the students who answer a question correctly should provide percentages of certainty higher than the students who answer incorrectly. Thus, for a multiple choice question which functions normally from the point of view of the certainty percentages use, we should observe among the students who choose the correct answer a tendency to answer with higher percentages of certainty and, in parallel, among the students who choose a wrong proposal a tendency to answer with lower percentages of certainty. We will then say that there is "spectral coherence". In the case where this situation does not arise, for example when students choose higher certainty percentages for one of the incorrect answers rather than for the correct answer, we are confronted with a problem of inconsistency in the use of the percentages of certainty, we will then speak of "spectral inconsistency". To measure spectral coherence we created two new types of indices starting from the classical point bi serial correlation coefficient (classical rpbis) calculation principle. The classical rpbis makes it possible to evaluate up to what extent each proposed alternatives solutions of each MCQ discriminates the students according to the criterion of the number of correct answers. Logically, one expects that the students who collect a high number of correct answers tend to choose the correct answer for a given question and that the students which collect a lower number have a tendency to choose an incorrect proposal. The two new types of spectral coherence indices of measurement are: (1) the Spectral Contrasted rpbis (rpbis SC) and (2) the Spectral Contrasted rpbis calculated after Turbo analysis (rpbis SCT). During a former research (Gilles, 1998) we developed the rpbis S in order to analyze the tendency to use higher certainty in case of correct answers than in case of incorrect answers. We use the name "rpbis SC" to indicate the rpbis S is constructed with a "Contrasted treatment" which consists in the computation of the rpbis SC of an incorrect answer by using the data of the students who choose this wrong response in contrast with the data restricted to the students who choose the correct answer excluding the data from the students who choose an other incorrect answer. The advantage lies in the elimination of the data of the

students who have chosen the other incorrect answers. This avoids introducing into the measurement of spectral coherence, the "background noise" generated by the data of the other incorrect answers.

The principle of "turbo analysis", consists in operating a selection in the data used for rpbis SC calculation on the basis of the level of realism reached by the students. We can thus increase the reliability of information related to the spectral indices by restricting the data to the students who make less errors in their self-assessments (in their use of certainty percentages). The name rpbis SCT indicates rpbis SC calculated with a Turbo analysis. The word "turbo" refers to the rise to power of the instrument in terms of the quality of information provided as one restricts oneself progressively to data from sets of students who make less and less errors in their self-assessments. We mention in the index's name the threshold of realism used to select the data. For instance rpbis SCT80 was calculated starting from the data of the students whose realism is equal to or higher than 80 (those who make between 0% and 20% of errors in their self-assessments).

The new indices rpbis SC and rpbis SCT are designed for the detection of problems located at the "alternatives level" inside the MCQ. We also adapted other spectral indices initially planned for the analysis of students performances so that these indices deliver information on the MCQ performances, at a "question level". For example, we also measured the level of spectral coherence of a question (NCSq) by taking into account the rpbis SC of the various proposals of a MCQ. The "turbo analysis" principle was also applied to the calculation of these spectral indices.

We tried to test these new spectral indices designed for detection of suspect alternatives within suspect MCQ using the data of several thousands of answers and certainty percentages collected during the ten standardized tests of the MOHICAN project (Leclercq & al., 2003). They consisted in ten tests of knowledge of the principal subjects at the level of the end of secondary education which were submitted to groups of students entering first year studies in eight of the nine university institutions of the French Community of Belgium (the number of questioned students varied between 1.392 and 3.846 according to tests). These standardized tests consisted of MCQ for which students were invited to choose one answer and systematically accompany it with a percentage of certainty.

The total number of MCQ for the ten MOHICAN tests was 173. For two MCQ among them, the 3rd and the 20th question of the general knowledge test in History and Socio-Economy, the values obtained with the spectral rpbis indicate situations of marked spectral inconsistency, the students tending to give lower percentages of certainty for the answers considered as correct and higher percentages of certainty for the incorrect answers. Studying the proposals of the two problematic MCQ by using the classical rpbis indices, we notice that the two MCQ do not function correctly from the point of view of classical discrimination. When we ask for the opinion of the experts of the contents, they confirm that the two MCQ display problematic results: for one of the questions a wrong proposal could also be regarded as being a correct answer and for the other, there is an error in the encoding of the correct answer. In the case of the MOHICAN tests, the spectral analysis thus allows to highlight two questions that a more qualitative analysis (posterior opinions of the experts) as well as a classical analysis of discrimination (classical rpbis) also indicate as questions with problems.

Keywords

Edumetrics
Metacognition
Testing

Authors

Name	Surname	Institution	Country	e-mail	EARLI Number	Presenting
Jean-Luc	Gilles	University of Li?ge	Belgium	jl.gilles@ulg.ac.be	298400824	*