

## A NOTE ABOUT A NEW INDEX TO EVALUATE TEACHING QUALITY

M. Civardi  
University of Milano Bicocca, Italy  
C. Crocetta, L. De Cesare,  
University of Foggia, Italy  
E. Zavarrone  
University of Milano Bicocca, Italy  
marisa.civardi@unimib.it

*Aim of this paper is to study a new index, called CI that has been proposed to measure the teaching quality in University courses. The questionnaire adopted by many Italian Universities to evaluate teaching quality of courses, has many items surveyed on a four points ordinal scale. This solution has many advantages and some limitations. Civardi (2002) has proposed a new index that summarizes results on each item for a single course on an interval  $[-1, 1]$ . The CI index is the algebraic sum of two indexes: an index expressing the score obtained on the half-plane of positive assessments and another corresponding to the half-plane of negative ones. In order to study the main characteristic of the CI index we have considered its sample space under different situations.*

### INTRODUCTION

Research workers in education undertake research studies that involve more aspects of teaching using questionnaires based mainly on closed questions.

Most typically, the closed question presents a proposition to which the student responds by ticking a box or writing a number indicating a scaled reaction. Typically, a five point scale is used which ranges from “Strongly agree,” through “Agree,” “Neutral,” “Disagree,” “Strongly disagree.”

Some workers prefer to use a four-point scale, which forces a student to express a positive or negative viewpoint. The “neutral” score is eliminated. Sometimes it can be replaced by categories for alternate or more specific viewpoints: “Not applicable,” “Don’t care” and so on.

Again, sometimes, it may be appropriate to employ different scales, such as “Almost Always” (AA), “Usually” (U), “Quite Frequently” (QF), “Occasionally” (O) and “Almost Never” (AN).

Another scale often used is “Excellent,” “Good,” “Satisfactory,” “Fair,” and “Poor.”

There is no necessity to use the same scale throughout the questionnaire. In fact, to mix the scales may encourage the respondent to think harder about answers, though too many different scales can confuse and waste the respondent’s time.

Further, it is a good idea to mix the questions or propositions some of which solicit a positive and some a negative expected response.

This discourages a respondent from unthinking repetition of a response, e.g., ticking the same column.

All the qualitative solutions above mentioned permit only numerical analysis of the frequencies, for this reason a quantification of responses can be very useful for large numbers of results.

### THE EVALUATION OF TEACHING IN ITALIAN UNIVERSITIES

The assessment of teaching performance in Italian universities has occupied an increasing proportion of academics’ time in recent years.

Such assessment has been a natural response to the demand for improved accountability of publicly funded organizations.

A common reaction of university staff is grudging acceptance of the new instruments of measurement which seems to have been imposed from above and which have encouraged a quantification culture in which it is not always easy to see the wood for the trees.

Few years ago, the Comitato Nazionale per la Valutazione del Sistema Universitario (CNVSU) has proposed a new questionnaire for the evaluation of teaching quality with many

items all surveyed on a four degree ordinal scale: “Yes, very much” (YVM), “Yes, so and so” (YSS), “No, little” (NLT), “No, nothing at all” (NNA).

With these ordinal scores, it is not possible to calculate a total score, either.

To analyze this kind of data usually the scale is dichotomized by considering the percentage of positive/negative assessments for each dimension, but there is very little to do more than calculating the median. The dichotomization entails a rather relevant loss of information, but the synthesis through the median is even less informative because based on a limited 4-degree scale and it has low possibility of discrimination within both subjects and courses.

In order to perform more sophisticated analysis it is useful to transform the results expressed in ordinal scale in values defined in the interval  $[-1, 1]$  by using a new index, called CI, that has been proposed to measure the teaching quality inside the Italian Universities.

This transformation can be justified by the fact that the students’ judgements can be expressed on a psychological continuous scale, whereas the answers are given on a four point ordinal scale.

### THE CI INDEX

The CI index is based on the observed distribution of the answers, given by students. It supplies a numerical score synthesizing one aspect of the courses under examination.

It is the algebraic sum of two indexes: an index expressing the score obtained on the half-plane of the positive assessments and another corresponding to the half-plane of the negative ones.

With reference to positive assessments, let us consider, for a generic course  $h$ , the distribution of the scores assigned to  $k$  items.

Let  $x_{ih}$  the percentage of positive assessments given by the formulae:

$$x_{ih} = 100[(N_{ih}(YVM) + N_{ih}(YSS) / N_{ih})]$$

and  $y_{ih}$  the percentage of very positive assessments on the overall positive assessments:

$$y_{ih} = 100[(N_{ih}(YVM) / (N_{ih}(YVM) + N_{ih}(YSS)))]$$

Then the couple  $P_i = (x_i, y_i)$  identifies, in the Cartesian plane, a point  $P_i$ , which lies in a 100 square surface (the positive assessment area).

Following the same procedure for the negative judgments let  $x_{ih}^*$  the percentage of negative assessments given by the formulae:

$$x_{ih}^* = 100[(N_{ih}(NNA) + N_{ih}(NLT) / N_{ih})]$$

and  $y_{ih}^*$  the percentage of very negative assessments on the overall positive assessments:

$$y_{ih}^* = 100[(N_{ih}(NNA) / (N_{ih}(NNA) + N_{ih}(NLT)))]$$

Then the couple  $P_i^* = (x_i^*, y_i^*)$  identifies, in the Cartesian plane, a point  $P_i^*$ , which lies in a 100 square surface (the negative assessment area).

Using the percentages above defined we can calculate:

$$CI^+ = \frac{x_i + ky_i}{\max(C_i^+)} \quad \text{and} \quad CI^- = -\frac{x_i^* + ky_i^*}{\max(C_i^+)}$$

where:  $CI^+ = 100(1+k)$  and  $k$  is an arbitrary parameter  $0 \leq k \leq 1$ .

The CI index is:

$$CI = CI^+ + CI^-$$

where  $-1 \leq CI \leq 1$ .

The structure of the CI index is very simple and it is easy to calculate. We give now some examples.

Given 2 points  $P_1 = (30,10)$  and  $P_1^* = (80,70)$  for  $k = 0$  we have:

$$CI = CI^+ + CI^- = 0.30 - 0.80 = -0.50.$$

For  $k = 1$  we have:

$$CI = CI^+ + CI^- = 0.20 - 0.75 = -0.55$$

In order to study the main properties of the CI index, using the software Matlab, we have realized a routine that calculates, for any number of respondents and for any  $0 \leq k \leq 1$ , the values of  $CI$  for all-possible models of answer we can get.

For example for 105 respondents, we have 204,156 different combinations of the four answers YVM, YSS, NLT and NNA. If we suppose that, the distribution of said answers is multinomial with constant probability 0.25 we have the following graphs of the CI index for different values of  $k$ .

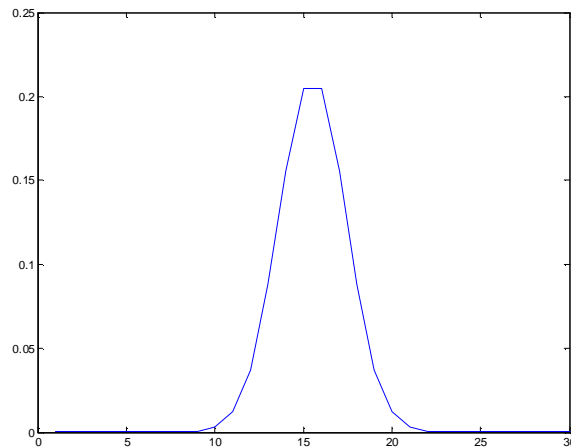


Figure 1: Distribution of the CI index for 105 respondents and  $k=0.1$

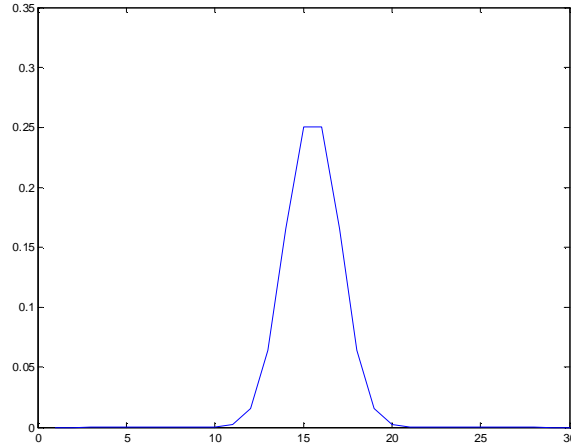


Figure 2: Distribution of the CI index for 105 respondents and  $k=0.9$

The above results can be considered as the starting point of our study because we have to study how the CI index varies for different values of  $k$  and considering an increasing number of respondents. In addition, the strong hypothesis of equi-probability of the four answers needs further studies.

#### REFERENCES

- Civardi, M. (2002). *Le Opinioni Degli Studenti*. Nucleo di valutazione, Università Milano “Bicocca”.
- Crocetta, C. and Toma, E. (2003) Un cruscotto di indicatori per la valutazione della didattica nell'Università di Foggia, L. Fabbris Editor, *LAID –OUT: Scoprire i rischi con l'analisi di segmentazione, Determinazione e previsione di rischi sociali e sanitari*, n. 3, (pp. 159-172), Cleup, Padova, February 2003
- Cuttance, P. (1987). *Modelling Variation in Effectiveness of Schooling*. Edinburgh: Centre for Education Sociology (CES).
- Gibbs, G. and Haigh, M. (1984). *Alternative Models of Course Evaluation: Examples from Oxford Polytechnic*, SCEDSIP, Standing Conference on Educational Development Services in Polytechnics.
- Slogane, P. (1993). Teaching quality assessment in economics in Scotland. *Royal Economic Society Newsletter*, 83, 12-13.