ADVANTAGES OF USING SELF-ORGANIZING MAPS TO ANALYSE STUDENT EVALUATIONS OF TEACHING

M. T. Sorrosal Forradellas[a], M. G. Barberà-Mariné[b], A. Fernández Bariviera[c], M. J. Garbajosa-Cabello[d]

Surveys to evaluate teaching performance are one of the most widely used instruments for assessing teaching capabilities and, consequently, the quality of teaching. Their success is largely due to how they are designed and the way in which they process information. The aim of this paper is to simplify the design of the student evaluations by removing the most correlated items, and to propose that Kohonen's self-organizing Kohonen maps be used to group teachers in accordance with all the characteristics surveyed. The methodology is applied to the particular case of the Rovira i Virgili University.

Keywords: *surveys to evaluate teaching performance, higher education, kohonen's self-organizing maps*

JEL Classification**:** *I21, I23, C45*

1. INTRODUCTION

Surveys to evaluate teaching performance were first used in America in the 1920s. One of the first publications based on the results of these surveys can be found in Remmers and Brandenburg (1927). They have subsequently been used in most universities in the world. Kulik (2001) notes that in 1973, 29% of American universities used student evaluations as a tool to assess teaching, while in 1993, they were used in 86%.

Student evaluations have two types of users: teachers and educational institutions. In this respect, Catano and Harvey (2011) point out that surveys are often used to provide remunerative supplements, to decide on continuity of teachers and to support faculty promotion decisions.

Theall and Franklin (2001) state that teaching staff are extremely sensitive to this issue. Some teachers, especially new ones, consider evaluations to be positive as they help improve the effectiveness of teaching. Kulik (2001), however, notes that some teachers are reluctant because they fear that students can turn them into a sort of personality contest.

[a,b,c,d] Department of Business Management, Universitat Rovira i Virgili, Spain. E-mail: mariateresa.sorrosal@urv.net, gloria.barbera@urv.net, aurelio.fernandez@urv.net, mariajose.garbajosa@urv.net

Student evaluations can also be understood as a form of accountability to society on the efficiency and effectiveness of teaching conducted by a university.

Among the existing literature, Arthur (2009) has studied the validity of using the views of students as a means of evaluating the performance of a teacher. Other studies analyze the factors that affect the evaluation. For example, Remedios and Lieberman (2007) studied the impact on outcomes of qualifications and workload during the course, and Griffin (2004) studied the effect of gender.

Although the content of evaluations is open to criticism and some items are of questionable relevance, Penny (2003) argues that universities continue to use them as a central element in assessing the effectiveness of teaching. It is therefore important that results be interpreted correctly.

This study has two objectives: first, to consider whether student evaluations of teaching can be simplified by reducing the number of items; and, second, to examine whether overall evaluations take into account all those aspects of teaching that need to be measured.

Although this study has been carried out in the particular case of the Rovira i Virgili University (URV), the methodology and the procedure can be generalized to any type of questionnaire.

The paper is organized as follows. Section 2 describes introduces the main methodological tool used, unsupervised Kohonen's self-organizing maps, a particular type of neural network. Section 3 describes the data used and shows the results obtained from the application of the Kohonen maps. Section 4 states presents the conclusions and there are some annexes at the end.

2. METHODOLOGY: KOHONEN'S SELF-ORGANIZING MAPS

To evaluate the data obtained in the survey we used a particular type of artificial neural network called a Kohonen's self-organizing map (SOM) (1989), which contains all the features of overall teacher evaluation and does not limit its analysis to the mean of the valuation (or average and standard deviation).

These networks were first developed by Dr. Teuvo Kohonen who, in 1982, began work on systems that stored information in the same way that the brain does: that is to say, similar memories are stored in areas of the brain that are close to each other, while disparate memories are stored in areas that are distant. Likewise, Kohonen maps produce a distribution of the elements analyzed (teachers) such that the proximity of teachers on the map indicates that they have similar characteristics. And the greater the distance between two teachers, the more different they are in terms of the items evaluated.

The advantage of using this type of artificial neural network to analyze the results of this or any other questionnaire is that it groups individuals in terms of the uniformity of the characteristics that define them, reducing the size of the problem to a two-dimensional map while maintaining all the information about the n features (items) valued. It should be pointed out that if attempts are to be made to improve teaching, the aspects that need to be improved must be known. The mean evaluation score dilutes the individual score on each of the items in such a way that teachers with the same mean can have a very different teaching profile.

In this paper, we compare the results of the group of teachers using the Kohonen self-organizing maps SOM produced only on the basis of the item that asks students their overall opinion of the teacher, thereby responding to the second of our goals.

Kohonen self-organizing maps SOM consist of two layers of neurons: an input layer with one unit for each of the characteristic patterns analyzed (in our case, each input neuron represents the value of one of the items of the questionnaire) and an output layer in the form of a two-dimensional map that locates teachers according to their degree of similarity (that is, two teachers who are close together on the map have similar educational characteristics). The dimension of the output layer depends on the amount of data being analyzed: the greater the volume of data, the larger the output layer.

As for the architecture of these networks, information is propagated within the system through feedforward connections (connections that have their origin in one layer and their target in a subsequent layer, in this case between the input and the output layer), lateral connections (connections between units of the same layer; in Kohonen maps SOM they are found in the output layer) and recurrent connections (connections from one unit to itself; in this case they are also found in each of the units forming the two-dimensional map). The lateral and recurrent connections of the neurons in the output layer allow the competitive process so characteristic of the unsupervised learning of these networks. After the learning process, only one neuron remains active in the output layer (the so-called winning neuron) and its position indicates where teachers are located within the map.

One of the main applications of Kohonen maps SOM is grouping (or clustering). Once the network has distributed all the elements (teachers) in the two-dimensional map on the basis of the similarity between the values of the items in the survey, the network can be told to group those elements. Of all the various possibilities for grouping, the most appropriate optimizes a dual objective function that seeks to minimize the number of highly homogeneous groups that are established.

Below we describe the process by which self-organizing Kohonen maps SOM are obtained and discuss how they can be used for the purpose of our work.

The patterns that form the input information represent each of the teachers who were surveyed at the Rovira i Virgili University in 2009–10. Each input vector component is one of the evaluations obtained by a teacher on the items in the survey: $x^p = \left(x_1^p, x_2^p, ..., x_i^p, ..., x_n^p\right)$. The superscript p indicates the teacher who is being analyzed and the subscript i = {1, 2, ... n} indicates the item that contains each component of the vector. We conducted two analyses: the first took into account the 11 characteristics assessed in the survey (but not the overall evaluation item with which we compare the results); and the second removed the correlated variables and decreased the number of features that define a teacher (n = 9).

Once all the patterns or teachers have been defined, the learning process of the network starts through the learning algorithm, which we describe briefly below:

- Each unit in the input layer is connected to k different units in the output layer (the number depends on the amount of data in the network, in our case the number of teachers from the Rovira i Virgili University surveyed). Associated with each connection are some weights $w_{k,i}$ (weight associated with the connection between neuron i in the input layer and neuron k in the output layer). These weights initially take a random value.
- When a pattern (the data for a teacher) is input into the network, the distance between the vector that contains the values of the items in the questionnaire and the weights associated with each of the units in the output layer is calculated. Although different definitions of distance can be used, the most common is the Euclidean distance, which we have also used here.

$$d = \sqrt{\sum_{i=1}^{n} \left(x_i^p - w_{ki}\right)^2}$$

- The neuron in the output layer that is at the shortest distance d from the pattern that has been introduced in the network is determined, and this neuron is the winning neuron for the input pattern.
- Modifications are made to the weights of the winning neuron and the neurons that are close to it (this is what is referred to as a learning process). To identify neurons that are close to the winning neuron a neighborhood area is established and defined in the design of the network. Taking the winning neuron as a reference, this rectangular or hexagonal area contains all the adjacent neurons within a particular radius which usually decreases as the number of iterations increase in the system. Thus, by reducing the number of neurons that change their weights, the network ensures convergence to a final result.
- New weights (corresponding to the time t +1) are calculated from the previous weights (corresponding to t) in accordance with the expression $w_{ki}(t+1) = w_{ki}(t) + \alpha(t)\left[x_i^p - w_{k^*i}(t)\right]$ where k* is the winning neuron and

$\alpha(t)$ denotes the learning coefficient, which has a value between 0 and 1 and decreases with the number of iterations to ensure convergence to equilibrium.

The process continues with other patterns being introduced into the network. When all the vectors have been input into the network, the whole process is repeated and a number of iterations are performed to stabilize the association between the different patterns and a unit in the output layer.

The resulting map shows the distribution patterns (teachers surveyed) in terms of their similarity, and is based on all the characteristics evaluated in the survey of teaching. To interpret the location of each of the patterns, the score of each of the variables in the map needs to be known. This information can also be obtained from the network in the form of maps of variables.

Finally we obtain the limits defined by the groups of teachers. The result is that the teachers are grouped in relation to the scores of all the items surveyed (and not just their mean and standard deviation). This grouping can be compared with the grouping obtained by applying the item that gives an overall assessment to determine whether this last item of the evaluation provides a full summary of the information contained by the previous items.

3. DATA AND APPLICATION

The evaluations that URV students make of teaching staff contain 12 items (see Table A1 in Appendix A). In this study we have used the outcome of the 906 surveys conducted in 2009–10.

First we obtained the correlation matrices for all the items from the overall data, to determine whether there are any high correlations between some of them. If there are, then according to the students the questions asked cover similar aspects and provide redundant information.

Applying a high correlation criterion (90%) shows that items 4 and 6, and 9 and 11 have correlations higher than 90%. We conclude that items 4 or 6, and 9 or 11 could be removed.

If we analyze each of the 12 centers of the URV in the same way (see Appendix B), we see that items 4 and 6 have correlations above 90% in all but two centers: F and E. And even in these cases the correlations are very high (84% and 85%, respectively). For items 9 and 11, however, variability is greater. Correlations are below 90% in five centers (D, E, F, H and K), although only in D is it below 85%.

In the above review we have ignored the correlations of the items with 12 ("Overall I think he/she is a good teacher"), as it can be seen that they are all highly correlated with it: in all cases the correlations are higher than 75% and in six they are above 85%. Therefore, this item can almost be considered to be a summary of the entire survey.

To summarize, two conclusions can be drawn from this first phase:

- Items 6 and 11 can be removed because they are highly correlated with 4 and 9, respectively. This shortens the evaluation and little information is lost in the sense that the results of the student opinions in the longer and shorter versions differ only very slightly. We have not made a formal assessment of the content of these questions and the information provided by their answers; we simply note that students respond very similarly to them.
- Item 12 is not included in the analysis because of its high correlation with all other questions. However, it can be used for purposes of comparison and as a global measure of the survey.

We will now move on to address our second objective: to what extent is question 12 representative of the entire survey and is this question in some way sufficient to classify teachers. We should point out that what we wish to determine is whether this last item classifies the teachers in the same way as all the other items taken together, not whether it is sufficient to make an evaluation in itself, since it is evident that a considerable amount of information on strong and weak points according to students would be lost.

As discussed in the section above, this study used SOM self-organizing Kohonen maps. Of the 12 URV centers, we discuss the analysis for three of them: centers B, C and F. These centers were chosen for two reasons: First, to check whether the conclusions we reach are independent of the degrees offered at the centers, which is why we have selected three centers from different branches of knowledge; and, second, because the three centers have a distinct correlation structure, particularly center F for which the items were considerably less correlated than for centers B and C.

The data provided by center B gave a total of 35 patterns or teachers (B1, B2, ..., B35) each with 11 features (all the survey questions except for item 12).

With this data and using the Toolbox (SOM) for Matlab to implement the SOM Kohonen maps, we obtain the distribution of teachers presented in Figure 1.

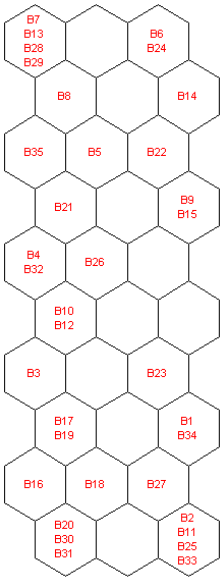

Figure 1. Distribution of teachers within the map (center B)

To interpret the characteristics of teachers from their position within the map, you must have information about the value of the variables in each of the cells. In this regard, each of the 11 images in Figure 2 shows the value of each of the 11 survey questions.

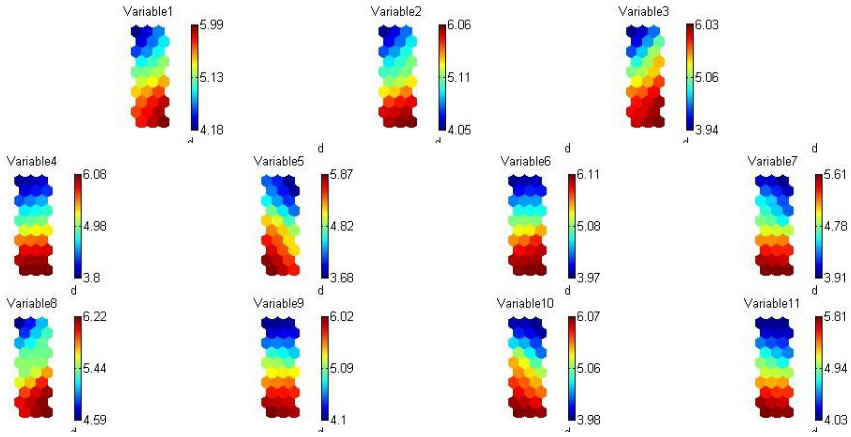

Figure 2. Value of each of the 11 survey questions (Center B)

As shown, the teachers at the top of the map are those who have obtained a low score on all items of the survey (shown in the chart by a range of cold colors), while the teachers at the bottom of the SOM Kohonen map have obtained higher scores (represented by a range of warm colors). This makes it possible to distinguish different teacher profiles visually. For example, in the central area of the map we found the teachers with average values on all items. However, there is a clear distinction between those who are in the middle right area of the map (with high scores on item 3 and lower scores on items 5 and 10) and those on the middle left (low scores on item 3 but higher scores on items 5 and 10).

However, if the default clustering of Toolbox is used to create clusters (keeping the number of groups to a minimum and maximizing homogeneity within each group), only two groups of teachers are obtained (see Figure 3). This classification is clearly of little use for designing specific strategies for improving teaching, and it is preferable to have more groups so that the specific training needs can be determined for each of them. Because center B does not have many teachers, we continue the analysis with the result obtained from this group, but for a larger center the process should be repeated with more groups so that decisions can be taken on the basis of the specificities of each group.

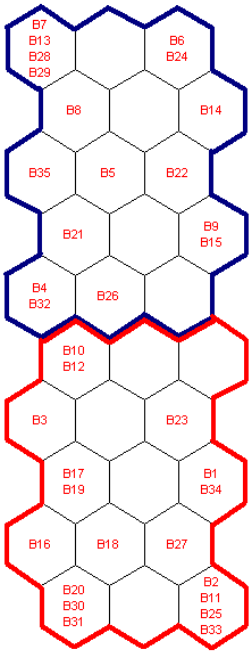

Figure 3. Groups of teachers of center B (11 items)

If the division into two groups is maintained, the teachers surveyed in center B are ordered as a function of item 12 and this order is compared to the group of teachers established by the SOM Kohonen maps, the correspondence obtained is the one displayed in Table A2 (see Appendix A). As can be seen, the relation between teachers belonging to a group and their position relative to item 12 is almost perfect. This result indicates that item 12 is a good summary of the assessment of other survey items.

However, since the items used to obtain the SOM Kohonen maps are correlated, the analysis was repeated with only 9 variables. Items 6 and 11 were removed because, as we have argued, they are highly correlated with items 4 and 9, respectively.

Using the survey results for the 9 items, the teachers from center B are clustered differently: in fact, four groups are formed (see Figure 4). For center B, then, which has a high correlation between item scores, the removal of the more correlated items allows greater insight into students' perception of the teaching staff. It may be noted, however, that the classification of these four groups is still consistent with the order that would be given if only item 12, the overall assessment of the teacher, were to be taken into account (see Table A3 in Appendix A).

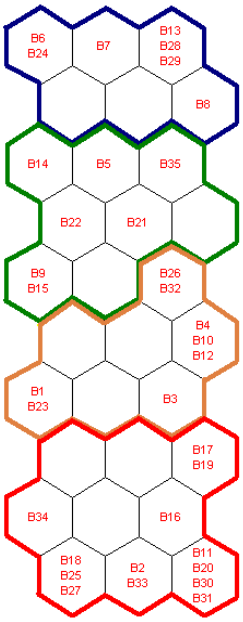

Figure 4. Groups of teachers of center B (9 items)

We conclude that, for center B, item 12 is a good approximation of the student evaluation of teachers and reflects very closely the other item scores. We would

like to point out that we are not suggesting that this item should replace all the other items; we are trying to determine whether there is any consistency between the nuances expressed by the individual items and the overall rating of the last item.

To determine whether these findings can be generalized to other centers of the URV, we performed the same analysis in centers C and F.

In center C, we found that for both 11 and 9 variables, the result of applying the SOM Kohonen maps gives exactly the same composition for the two groups into which the 37 teachers are divided (see Appendix C, Table C1). Again, this grouping is entirely consistent with how the teachers are ordered by item 12.

Finally, for center F (where the correlation between the item responses was lowest), when items 6 and 11 were removed, the SOM Kohonen map changed from two to five groups (see Appendix C, Table C2), showing that in this center there is a greater variety of teaching profiles (a total of 55 teachers were evaluated). As can be seen, the greater dispersion in the characteristics of the teachers makes the use of SOM Kohonen maps more valuable, since if they are sorted only as a function of item 12 numerous nuances are lost, many of which are fundamental features of each of the groups.

Finally, note that the number of groups does not always increase when we use 9 variables instead of 11 to analyze the teachers of a center. The increase in the number of groups will depend on both the number of teachers who work in the center, and the correlation structure of the items.

It should be pointed out that one of the most useful aspects of these artificial neural networks is the number of groups created for the analysis of evaluations. Although in this study we have allowed the SOM Kohonen network to determine the number of groups as a function of the homogeneity within the group but keeping the number to a minimum, Kohonen maps SOM can also be used to distribute teachers into k clusters where k is a number that is fixed a priori. Ask increases, the groups will be more homogeneous, which will help in taking decisions on policies of teacher training, evaluation of teachers and so on.

4. CONCLUSIONS

Student evaluations of teachers are commonly used by educational institutions, and especially universities, to obtain information about teachers and the perception that users (*i.e.* students) have of the quality of educational processes. It is important, therefore, for evaluations to be simple and short but at the same time, to contain enough information to be used subsequently as a basis for improvement.

To this end we propose the use of Kohonen's self-organizing maps, artificial neural networks that cluster teachers into categories. Each category or group is homogeneous with respect to all characteristics evaluated in the survey, not only with respect to their mean, which makes it possible to design specific actions for each group of teachers.

In this case, Kohonen's maps were applied to data from student evaluations of teachers carried out in 2009–10 at the Rovira i Virgili University. The data was provided by individual faculties or schools (centers), and we focused on those whose distribution of correlations between items was quite different, so that the goodness of the results could be confirmed in all cases. The analysis of these results not only allows the institution to define general policies but those occupying positions of responsibility to design more specific policies for its own objectives.

The methodology enables the number of groups to be set a priori, which is very useful for implementing policies to improve teaching, the ultimate goal of analyzing the results of student evaluations of teachers.

ACKNOWLEDGMENTS

This work was partially supported by Institut Ciències de l'Educació of Universitat Rovira i Virgili (Project B06/2010).

REFERENCES

ARTHUR, L. (2009). "From performativity to professionalism: lecturers' responses to student feedback". *Teaching in Higher Education*, 14, p. 441-454.

CATANO, V.M.; HARVEY, S. (2011). "Student perception of teaching effectiveness: development and validation of the Evaluation of Teaching Competencies Scale (ETCS)". *Assessment & Evaluation in Higher Education*, 36(6), p. 701-717.

COOPER, C.; KILMER, R. (2012): "Using Self Organizing Maps to Find Good Comparison Universities" *Lecture Notes in Computer Science*, 7081, p. 148-153.

GRIFFIN, B.W. (2004). "Grading leniency, grade discrepancy, and student ratings of instruction". *Contemporary educational psychology*, 29, p. 410-425.

KOHONEN, T. (1989). *Self-Organization and Associative Memory*. Springer-Verlag.

KULIK, J.A. (2001). "Student Ratings: Validity, Utility, and Controversy". *New Directions for Institutional Research*, 109, p. 9-25.

PENNY, A.R. (2003). "Changing the Agenda for Research into Students' Views about University Teaching: Four shortcomings of SRT research". *Teaching in Higher Education*, 8(3), p. 399-411.

REMEDIOS, R.; LIEBERMAN, D.A. (2007). "I liked your course because you taught me well: the influence of grades, workload, expectations and goals on students' evaluations of teaching". *British Educational Research Journal*, 34, p. 91-115.

REMMERS, H.H.; BRANDENBURG, G.C. (1927). "Experimental data on the Purdue Rating Scale for Instructors". *Educational Administration and Supervision*, 13, p. 519–527.

SPOOREN, P.; MORTELMANS, D.; THIJSSEN, P. (2012). "'Content' versus 'style': acquiescence in student evaluation of teaching?", British Educational Research Journal, 38(1), p. 3-21.

THEALL, M.; FRANKLIN, J. (2001), "Looking for Bias in All the Wrong Places: A Search for Truth or a Witch Hunt in Student Ratings of Instruction?". *New Directions for Institutional Research*, 109, p. 45–56.

APPENDIX A:

Table A1. Evaluation items for 2009–10

1	At the beginning of the subject, the teacher clearly stated the objectives, the syllabus and the evaluation criteria.
2	The teacher restricts explanations to the aims and content of the syllabus.
3	The teacher's explanations clearly show that he/she has prepared the class.
4	The teacher presents and explains concepts clearly and thoroughly.
5	The teacher tries to stimulate student interest.
6	The teacher responds satisfactorily to the questions asked.
7	The teacher uses effective educational techniques (speech, visuals, new technologies, etc.)
8	The teacher carries out his/her duties (timetable, tutorials, etc.)
9	As well as providing basic knowledge, the teacher expands on it and (if needs be) exemplifies using current issues.
10	The teacher relates and communicates well with students.
11	The teacher provides insight into the latest research and developments affecting the subject.
12	Overall I regard him/her to be a good teacher.

Table A2. Center B: comparison between the group of teachers and how they are ordered by item 12 (11 items)

List of teachers	Rating item 12
B29	3.67
B7	4.05
B24	4.07
B28	4.09
B13	4.25
B6	4.34
B14	4.47
B8	4.62
B5	4.81
B35	4.82
B15	4.83
B22	4.84
B9	5.09
B26	5.18
B21	5.22
B12	5.39
B32	5.41
B4	5.54

List of teachers	Rating item 12
B23	5.57
B10	5.58
B3	5.61
B1	5.76
B33	5.97
B19	6.06
B18	6.10
B25	6.11
B34	6.14
B2	6.14
B17	6.22
B27	6.22
B31	6.30
B16	6.33
B30	6.36
B11	6.39
B20	6.63

Table A3. Center B: comparison between the group of teachers and how they are ordered by item 12

List of teachers (groups with 11 items)	Rating item 12	List of teachers (groups with 9 items)	List of teachers (groups with 11 items)	Rating item 12	List of teachers (groups with 9 items)
29	3.67	29	4	5.54	4
7	4.05	7	23	5.57	23
24	4.07	24	10	5.58	10
28	4.09	28	3	5.61	3
13	4.25	13	1	5.76	1
6	4.34	6	33	5.97	33
14	4.47	14	19	6.06	19
8	4.62	8	18	6.10	18
5	4.81	5	25	6.11	25
35	4.82	35	34	6.14	34
15	4.83	15	2	6.14	2
22	4.84	22	17	6.22	17
9	5.09	9	27	6.22	27
26	5.18	26	31	6.30	31
21	5.22	21	16	6.33	16
12	5.39	12	30	6.36	30
32	5.41	32	11	6.39	11
			20	6.63	20

APPENDIX B: Correlation between the 12 items of the survey by center

Table B1. Center A

Col.	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.98	1										
3	0.95	0.94	1									
4	0.67	0.67	0.64	1								
5	0.76	0.81	0.77	0.67	1							
6	0.74	0.73	0.67	0.93	0.79	1						
7	0.79	0.81	0.71	0.83	0.78	0.87	1					
8	0.83	0.87	0.91	0.56	0.75	0.60	0.65	1				
9	0.88	0.88	0.88	0.71	0.90	0.80	0.84	0.82	1			
10	0.74	0.78	0.74	0.76	0.94	0.86	0.84	0.71	0.92	1		
11	0.75	0.78	0.70	0.65	0.87	0.80	0.80	0.70	0.92	0.89	1	
12	0.89	0.91	0.87	0.78	0.90	0.87	0.90	0.84	0.95	0.95	0.90	1

Table B2. Center B

Col.	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.95	1										
3	0.93	0.93	1									
4	0.88	0.94	0.88	1								
5	0.62	0.65	0.55	0.72	1							
6	0.86	0.90	0.87	0.96	0.76	1						
7	0.76	0.79	0.70	0.87	0.85	0.90	1					
8	0.78	0.84	0.80	0.78	0.50	0.80	0.67	1				
9	0.86	0.90	0.83	0.90	0.85	0.90	0.87	0.73	1			
10	0.74	0.80	0.71	0.87	0.88	0.90	0.89	0.75	0.90	1		
11	0.82	0.86	0.79	0.88	0.84	0.90	0.86	0.70	0.95	0.89	1	
12	0.86	0.92	0.86	0.95	0.82	0.97	0.90	0.81	0.95	0.94	0.94	1

Table B3. Center C

Col.	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.93	1										
3	0.83	0.84	1									
4	0.79	0.87	0.84	1								
5	0.79	0.79	0.75	0.79	1							
6	0.77	0.82	0.75	0.90	0.81	1						
7	0.85	0.83	0.80	0.84	0.85	0.88	1					
8	0.91	0.86	0.84	0.74	0.78	0.71	0.85	1				
9	0.75	0.69	0.69	0.73	0.80	0.75	0.93	0.73	1			
10	0.72	0.74	0.64	0.77	0.90	0.87	0.84	0.70	0.78	1		
11	0.77	0.73	0.75	0.80	0.81	0.76	0.93	0.77	0.96	0.76	1	
12	0.83	0.82	0.78	0.86	0.85	0.93	0.89	0.76	0.83	0.90	0.83	1

Table B4. Center D

Col.	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.86	1										
3	0.82	0.87	1									
4	0.85	0.84	0.87	1								
5	0.75	0.80	0.77	0.88	1							
6	0.84	0.91	0.90	0.92	0.86	1						
7	0.74	0.67	0.71	0.83	0.79	0.79	1					
8	0.77	0.83	0.82	0.77	0.73	0.85	0.73	1				
9	0.58	0.57	0.63	0.67	0.68	0.69	0.79	0.64	1			
10	0.68	0.78	0.74	0.80	0.86	0.86	0.70	0.76	0.68	1		
11	0.54	0.52	0.63	0.71	0.69	0.67	0.67	0.52	0.73	0.56	1	
12	0.83	0.86	0.86	0.94	0.88	0.95	0.79	0.83	0.66	0.84	0.67	1

Table B5. Center E

Col.	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.88	1										
3	0.81	0.83	1									
4	0.82	0.85	0.88	1								
5	0.72	0.78	0.75	0.79	1							
6	0.82	0.84	0.84	0.84	0.80	1						
7	0.70	0.70	0.72	0.73	0.74	0.75	1					
8	0.74	0.74	0.74	0.71	0.61	0.77	0.60	1				
9	0.65	0.69	0.71	0.73	0.77	0.72	0.73	0.57	1			
10	0.70	0.76	0.73	0.77	0.87	0.76	0.72	0.59	0.75	1		
11	0.64	0.66	0.64	0.68	0.75	0.72	0.76	0.55	0.89	0.71	1	
12	0.85	0.87	0.88	0.91	0.85	0.91	0.78	0.76	0.76	0.84	0.76	1

Table B6. Center F

Col.	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.83	1										
3	0.71	0.85	1									
4	0.62	0.83	0.88	1								
5	0.47	0.66	0.67	0.75	1							
6	0.57	0.76	0.81	0.85	0.88	1						
7	0.53	0.70	0.83	0.79	0.78	0.82	1					
8	0.68	0.72	0.70	0.62	0.42	0.56	0.61	1				
9	0.41	0.62	0.73	0.71	0.79	0.76	0.79	0.51	1			
10	0.49	0.67	0.70	0.72	0.87	0.85	0.79	0.64	0.76	1		
11	0.27	0.49	0.55	0.53	0.70	0.61	0.64	0.31	0.88	0.62	1	
12	0.66	0.84	0.90	0.90	0.79	0.89	0.81	0.74	0.78	0.83	0.62	1

Table B7. Center G

Col.	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.89	1										
3	0.81	0.89	1									
4	0.86	0.88	0.88	1								
5	0.81	0.81	0.78	0.84	1							
6	0.84	0.89	0.90	0.92	0.87	1						
7	0.73	0.74	0.72	0.78	0.78	0.76	1					
8	0.73	0.76	0.74	0.62	0.62	0.64	0.54	1				
9	0.68	0.68	0.71	0.75	0.74	0.77	0.69	0.53	1			
10	0.81	0.81	0.77	0.85	0.91	0.88	0.72	0.60	0.77	1		
11	0.70	0.69	0.69	0.74	0.74	0.76	0.70	0.53	0.94	0.75	1	
12	0.87	0.90	0.89	0.94	0.88	0.93	0.75	0.69	0.78	0.91	0.76	1

Table B8. Center H

Col.	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.86	1										
3	0.80	0.83	1									
4	0.85	0.86	0.88	1								
5	0.75	0.79	0.73	0.85	1							
6	0.85	0.85	0.85	0.92	0.85	1						
7	0.66	0.62	0.54	0.68	0.78	0.67	1					
8	0.71	0.73	0.75	0.73	0.67	0.73	0.57	1				
9	0.78	0.77	0.79	0.82	0.85	0.85	0.69	0.76	1			
10	0.81	0.81	0.77	0.81	0.89	0.86	0.68	0.73	0.86	1		
11	0.78	0.68	0.68	0.73	0.75	0.77	0.66	0.61	0.87	0.79	1	
12	0.88	0.86	0.87	0.90	0.87	0.92	0.70	0.78	0.86	0.91	0.80	1

Table B9. Center I

Col.	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.89	1										
3	0.87	0.91	1									
4	0.85	0.92	0.91	1								
5	0.82	0.86	0.88	0.90	1							
6	0.87	0.91	0.92	0.94	0.92	1						
7	0.84	0.85	0.88	0.89	0.92	0.90	1					
8	0.86	0.85	0.87	0.80	0.81	0.84	0.81	1				
9	0.81	0.82	0.88	0.86	0.93	0.89	0.90	0.83	1			
10	0.81	0.86	0.86	0.87	0.94	0.91	0.89	0.82	0.90	1		
11	0.74	0.73	0.80	0.78	0.88	0.81	0.86	0.73	0.92	0.84	1	
12	0.87	0.93	0.93	0.95	0.95	0.96	0.92	0.86	0.91	0.94	0.82	1

Table B10. Center J

Col.	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.85	1										
3	0.63	0.79	1									
4	0.66	0.81	0.82	1								
5	0.71	0.76	0.75	0.80	1							
6	0.71	0.86	0.83	0.91	0.83	1						
7	0.74	0.83	0.73	0.81	0.79	0.79	1					
8	0.72	0.84	0.74	0.70	0.73	0.69	0.80	1				
9	0.68	0.76	0.83	0.74	0.80	0.76	0.78	0.69	1			
10	0.68	0.76	0.74	0.73	0.93	0.83	0.71	0.70	0.78	1		
11	0.61	0.71	0.75	0.74	0.81	0.73	0.73	0.69	0.90	0.76	1	
12	0.74	0.87	0.86	0.87	0.88	0.92	0.83	0.79	0.82	0.88	0.80	1

Table B11. Center K

Col.	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.81	1										
3	0.69	0.85	1									
4	0.75	0.84	0.86	1								
5	0.75	0.79	0.73	0.81	1							
6	0.72	0.79	0.83	0.90	0.81	1						
7	0.66	0.70	0.65	0.68	0.70	0.69	1					
8	0.68	0.72	0.68	0.62	0.65	0.66	0.57	1				
9	0.62	0.74	0.75	0.77	0.78	0.79	0.70	0.65	1			
10	0.73	0.75	0.68	0.77	0.87	0.81	0.68	0.71	0.80	1		
11	0.56	0.71	0.72	0.73	0.72	0.74	0.67	0.60	0.85	0.71	1	
12	0.80	0.86	0.85	0.90	0.89	0.90	0.72	0.71	0.83	0.89	0.77	1

Table B12. Center L

Col.	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.89	1										
3	0.77	0.88	1									
4	0.80	0.91	0.89	1								
5	0.77	0.88	0.80	0.89	1							
6	0.82	0.89	0.81	0.92	0.92	1						
7	0.80	0.86	0.82	0.83	0.84	0.80	1					
8	0.71	0.79	0.74	0.72	0.73	0.73	0.74	1				
9	0.78	0.85	0.81	0.86	0.87	0.85	0.88	0.71	1			
10	0.74	0.83	0.73	0.82	0.91	0.92	0.78	0.72	0.82	1		
11	0.76	0.80	0.74	0.81	0.83	0.81	0.88	0.66	0.93	0.80	1	
12	0.83	0.94	0.89	0.93	0.94	0.93	0.88	0.82	0.90	0.92	0.86	1

APPENDIX C: Comparison between the group of teachers and how they are ordered by item 12

Table C1. Center C

List of teachers (groups with 11 items)	Rating Item 12	List of teachers (groups with 9 items)	List of teachers (groups with 11 items)	Rating Item 12	List of teachers (groups with 9 items)
19	2.86	19	5	5.07	5
31	3.00	31	4	5.14	4
35	3.11	35	36	5.18	36
18	3.33	18	15	5.30	15
23	3.44	23	20	5.33	20
12	3.55	12	28	5.41	28
24	3.55	24	37	5.41	37
10	3.92	10	1	5.44	1
11	4.11	11	8	5.44	8
32	4.18	32	17	5.58	17
3	4.29	3	33	5.66	33
9	4.29	9	21	5.68	21
27	4.29	27	22	5.79	22
34	4.29	34	7	5.82	7
30	4.32	30	2	5.85	2
26	4.48	26	25	5.94	25
14	4.81	14	13	6.23	13
6	5.00	6	29	6.25	29
16	5.00	16			

Table C2. Center F

List of teachers (groups with 11 items)	Rating Item 12	List of teachers (groups with 9 items)	List of teachers (groups with 11 items)	Rating Item 12	List of teachers (groups with 9 items)
35	3.00	35	42	5.20	42
20	3.25	20	52	5.33	52
31	3.38	31	24	5.42	24
11	3.58	11	54	5.42	54
5	3.67	5	8	5.43	8
33	3.67	33	16	5.46	16
6	3.83	6	43	5.50	43
10	3.83	10	48	5.56	48
36	3.83	36	55	5.63	55
28	3.92	28	9	5.67	9
38	4.00	38	19	5.67	19
41	4.00	41	21	5.75	21
12	4.11	12	22	5.85	22
3	4.25	3	13	5.88	13
27	4.33	27	1	5.92	1
7	4.44	7	30	6.00	30
49	4.44	49	34	6.00	34
45	4.60	45	44	6.00	44
17	4.72	17	40	6.11	40
23	4.79	23	46	6.17	46
32	4.83	32	29	6.17	29
50	4.86	50	37	6.20	37
51	4.92	51	39	6.25	39
25	5.00	25	47	6.33	47
15	5.12	15	2	6.43	2
26	5.14	26	53	6.50	53
18	5.20	18	14	6.52	14
			4	6.56	4