Available online at www.sciencedirect.com

ScienceDirect

Procedia - Social and Behavioral Sciences 228 (2016) 362 – 368

Procedia Social and Behavioral Sciences

2nd International Conference on Higher Education Advances, HEAd´16, 21-23 June 2016, València, Spain

A tool to translate scores across different systems

Gregoria Mateos Aparicio[a], Adolfo Hernández[a], Elena Martínez[a]*

[a]*Department of Statistics and Operational Research II (Decision Methods) Complutense University of Madrid*

Abstract

One of the main consequences of the globalization process of education is the mobility of students between different countries. It is vital for the success of mobility programs that the work done by a student could be easily accredited in other countries. Calculation of equivalences is also crucial in any competitive process in which one has to compare average scores across all candidates.

This paper presents a general method to translate scores based on the percentile distribution of grades achieved by students in the different countries. The objective is to provide a tool to assess a student´s academic record regardless of the evaluation system, avoiding imbalances in student scores from different countries. The method is illustrated with an application, where equivalences are obtained between two different grade systems. The question of whether areas of knowledge should be considered is also addressed. This method could be easily extended to any country where information about distribution of grades is available.

Peer-review under responsibility of the organizing committee of HEAd´16

Keywords: globalization; mobility; qualification system; equivalences; percentile distribution

1. Introduction

Nowadays, and even more in the future, university students may spend part of their educational journey in a university away from their home country. This situation has special relevance in Europe, with the creation in 1999 of the European Higher Education Area (EHEA) and the use of a single measurement system (European Credit

* Corresponding author. Tel.:+34-913942906; fax: +34-913942388.
E-mail address: adolfher@ucm.es

Peer-review under responsibility of the organizing committee of HEAd´16
doi:10.1016/j.sbspro.2016.07.054

Transfer System, ECTS). Since the rating systems of the countries are distinct, it is essential to make an equivalence with the objective that no changes to the values occur in the scores of students participating in mobility programs and, therefore, in the final average mark of their academic records. The fair and proper interpretation of the various grading systems is an important issue not only for students (participating in a mobility program, applying for a place or a grant in higher education institutions) but also for university staff who must assess their records. Although there is a large and varied set of formulas, they often forget the complexity of such processes. As mentioned in Haug (1997), 'foreign grades are not just numbers that can be calculated by applying a mathematical formula…. Simple mathematical formulas with their claim to universality are nothing but a fallacious over-simplification of a reality they fail to capture.'

This paper presents a statistical method for the calculation of equivalent scores in different assessment systems, based on the comparison of the percentage distribution of grades. This idea follows the theory developed in Linn (1993) and in Petersen, Kolen and Hoover (1993) related to the process generally known as equating, and more specifically the idea of equipercentile equating (Angoff 1984). As described in Kolen (1984) this is a two-stage process where relative cumulative frequency distributions are tabulated, and then scores with identical relative cumulative frequencies are equated. Although the problem to solve is complicated, this idea is simple and the application of our methods only requires the knowledge of the frequency distribution of grades in each country considered. It works with all types of scales, not only numerical, but also qualitative although obviously, one thing they all have in common is their ordinal nature. Our objective is to show an easy way to implement these techniques and to encourage administrators and teaching staff to make use of these tools in their institutions. To illustrate this method, an example is developed, in which equivalences are obtained using as a reference a numerical scale from 0 to 10, with one decimal place. We have taken this system since it is a very common scale.

Finally, the question of whether we should consider the areas of knowledge is considered. We conclude that there are differences between the records of students in terms of the scientific area and, therefore, this variable should be taken into account in the method.

2. Methodology

Suppose we have a rating system A (source system), this could be the rating system used in any country, with any type of scale; on the other, we have a rating system B (reference system) which uses a numerical scale. The information available on these two rating systems can be summarized through the distribution of frequencies of the different points of the scales. These distributions are often available either from the Education Departments of the countries considered or online in pages as the World Higher Education Database (WHED). The question is how to calculate the equivalent of grades obtained in the system A to the system B so that a credit rated with a specific point of the scale used in system A corresponds to a credit rated with an “equivalent grade” of system B. Our proposal is based on the statistical concept of quantile, a generalization of the well-known concept of quartile (see for example p. 74 in Clarke and Cooke, 2004). Given a certain grade in system A, this corresponds to a specific quantile of the distribution of frequencies (A). Its equivalent grade in system B will be the grade assigned to the same quantile in the distribution of frequencies (B). Our method is in fact a search for the quantile equivalent position in another system.

The rigorous theoretical formulation is contained in Appendix 1. This appendix contains the necessary technical notation and all the details of the methodology. In spite of providing all this information, we strongly believe that reading it is not necessary in order to neither understand nor apply the method. The reading of the example in the following section should be enough to apply the method. This means that, although the formulation of the methodology is complex, the application of the method is simple and affordable for any individual even without specific training in quantitative methods.

3. Application

We will use real data corresponding to Spain, one of the countries where this scale is used. More specifically, a sample of 819 records of students who have applied for different grants to the Spanish Ministry of Education has been used. In this sample, 236634 credits have been graded, each one with a number from 5.0 to 10.0.

The results are shown in Table 1.

Table 1. Distribution of frequencies in the Spanish grade system

Spanish Grades	Credits (%)	Cumulate Credits (%)	Spanish Grades	Credits (%)	Cumulate Credits (%)
5.0	8.00	8.00	7.6	0.89	65.73
5.1	0.67	8.67	7.7	0.97	66.70
5.2	0.86	9.53	7.8	1.02	67.72
5.3	0.81	10.35	7.9	0.60	68.32
5.4	0.74	11.09	8.0	6.21	74.53
5.5	12.41	23.50	8.1	0.63	75.16
5.6	0.91	24.41	8.2	0.83	75.99
5.7	0.86	25.26	8.3	0.77	76.76
5.8	0.95	26.22	8.4	0.53	77.29
5.9	0.64	26.86	8.5	2.12	79.41
6.0	6.70	33.56	8.6	0.40	79.81
6.1	0.90	34.46	8.7	0.42	80.23
6.2	1.04	35.50	8.8	0.26	80.49
6.3	1.11	36.61	8.9	0.09	80.58
6.4	0.91	37.52	9.0	10.25	90.83
6.5	3.03	40.55	9.1	0.47	91.30
6.6	0.80	41.34	9.2	0.64	91.94
6.7	0.87	42.22	9.3	0.48	92.42
6.8	0.70	42.92	9.4	0.34	92.76
6.9	0.31	43.23	9.5	1.99	94.75
7.0	8.23	51.46	9.6	0.28	95.03
7.1	0.97	52.43	9.7	0.35	95.38
7.2	1.18	53.61	9.8	0.32	95.70
7.3	1.06	54.67	9.9	0.22	95.92
7.4	0.92	55.59	10.0	4.08	100.00
7.5	9.26	64.85			

The Cuban rating system is an example of a short numerical scale with only three values, 3, 4 and 5, with 3 being the minimum score required to pass and 5 the highest possible score. Table 2 shows the percentage of credits for each grade for Cuban students who have applied for the validation of their records to the Spanish Ministry of Education.

Table 2. Distribution of frequencies in the Cuban grade system

Cuban Grades	Credits (%)	Cumulate credits (%)
3.00	12.96	12.96
4.00	56.19	69.15
5.00	30.85	100.00

Table 3 shows the equivalences obtained, so that, for example, credits with a grade of 3 on the Cuban rating system should be assessed with a rating of 5.13 in the Spanish system.

Table 3. Equivalences between the Cuban and the Spanish systems

Cuban Grades	Equivalent Spanish Grades
3.00	5.13
4.00	6.59
5.00	8.91

Application of our methodology gives 5.13 as the equivalent grade of a Cuban 3.00 in the Spanish system. This number has been calculated as the weighted average of the marks obtained in the lowest 12.96% percentage of the Spanish distribution. In order to understand how the method works, specific calculations to obtain this result are shown in Equation 1 below:

$$5.13 = \frac{5.0 \times 8.00 + 5.1 \times 0.67 + 5.2 \times 0.86 + 5.3 \times 0.81 + 5.4 \times 0.74 + 5.5 \times 1.88}{12.96} \qquad (1)$$

Similarly, 6.59 is the Spanish equivalent of the Cuban 4.00 and can be obtained as the weighted average of the marks contained between the 12.96% and the 12.96% + 56.19% = 69.15% percentage of the Spanish distribution, as illustrated in Equation 2 below:

$$6.59 = \frac{5.5 \times 10.53 + 5.6 \times 0.91 + 5.7 \times 0.86 + \cdots + 7.8 \times 0.97 + 7.9 \times 0.60 + 8.0 \times 0.82}{56.19} \qquad (2)$$

Finally, 8.91 is the Spanish equivalent of the Cuban 5.00 and can be obtained as the weighted average of the marks obtained in the highest 30.848% percentage of the Spanish distribution, as illustrated in Equation 3 below:

$$8.91 = \frac{8.0 \times 5.38 + 8.1 \times 0.63 + 8.2 \times 0.83 + \cdots + 9.8 \times 0.32 + 9.9 \times 0.22 + 10 \times 4.08}{30.85} \qquad (3)$$

3.1. Grouping Scientific Fields

So far, the student records have not been differentiated by the areas of knowledge that are attached to university studies. However, it seems logical, a priori, to think that there may be significant differences between the records of students in terms of the scientific area and, therefore, this variable should be taken into account in the method. For the classification of the different degrees by subject area, we have used the ISCED (International Standard Classification of Education), developed by UNESCO. There are eight different groups (Table 4). Frequency distributions can now be obtained for each of the groups. These are omitted in this particular paper but can be obtained by request from the authors. We first test whether the hypothesis of normality can be assumed. We use the Lilliefors test, an adaptation of the Kolmogorov-Smirnov test. We decide against this hypothesis of normality in all groups, as illustrated in table 4.

Table 4. Testing Normality: results

ISCED codes		Lilliefors Test		
		Statistic	gl	p-value
1	Education	0.098	1000	0.000
2	Humanities and Arts	0.133	1000	0.000
3	Social Sciences, Business and Law	0.108	1000	0.000
4	Sciences	0.113	1000	0.000
5	Engineering, Manufacturing and Construction	0.110	1000	0.000
6	Agriculture	0.088	1000	0.000
7	Health and Welfare	0.091	1000	0.000
8	Education	0.120	1000	0.000

To test whether there are significant differences between the grades obtained in different groups, we have used the Kruskal-Wallis test, a nonparametric technique which allows us to discuss whether or not data come from the same population, without assuming normality. As shown in Table 5, we accept that there are differences between the groups.

Table 5. Kruskal-Wallis Test

Chi-square value	163.148
Degrees of Freedom	7
p-value	0.000

To see what the different groups are, we have used the Mann-Whitney test, which tests whether there are significant differences between two ISCED groups from the sample information available. Table 6 shows, in summary, the results obtained.

Table 6. Summary of Mann-Whitney Test

	Groups	p-value	Result		Groups	p-value	Result
1	1 - 2	0.002	Equal	15	3 - 5	0.000	Different
2	1 - 3	0.014	Equal	16	3 - 6	0.234	Equal
3	1 - 4	0.000	Different	17	3 - 7	0.329	Equal
4	1 - 5	0.000	Different	18	3 - 8	0.000	Different
5	1 - 6	0.155	Equal	19	4 - 5	0.008	Equal
6	1 - 7	0.162	Equal	20	4 - 6	0.008	Equal
7	1 - 8	0.000	Different	21	4 - 7	0.012	Equal
8	2 - 3	0.000	Different	22	4 - 8	0.000	Different
9	2 - 4	0.000	Different	23	5 - 6	0.000	Different
10	2 - 5	0.000	Different	24	5 - 7	0.000	Different
11	2 - 6	0.000	Different	25	5 - 8	0.323	Equal
12	2 - 7	0.000	Different	26	6 - 7	0.965	Equal
13	2 - 8	0.000	Different	27	6 - 8	0.000	Different
14	3 - 4	0.192	Equal	28	7 - 8	0.000	Different

Therefore, we could move from the eight initial groups to five: 1-2, 3-4, 5, 6-7 and 8. As the reduction is not significant, we keep the eight original groups in order to follow international standards. As a consequence, our strong recommendation is to recalculate the equivalences taking into account the area of knowledge, keeping the ISCED eight groups. Table 7 shows equivalences between the Cuban and the Spanish systems taking into account the area of knowledge. As expected, different equivalences are obtained.

Table 7. Equivalences between the Cuban and the Spanish systems taking into account different ISCED areas of knowledge

	Equivalent Spanish Grades							
Cuban Grades	ISCED 1	ISCED 2	ISCED 3	ISCED 4	ISCED 5	ISCED 6	ISCED 7	ISCED 8
3.00	5.13	5.26	5.08	5.16	5.01	5.17	5.05	5.01
4.00	6.91	7.14	6.64	6.47	6.35	6.68	6.75	6.41
5.00	9.07	9.34	9.00	8.70	8.70	8.85	8.94	8.51

Figure 1 shows graphically the distribution of the Cuban grades classified by ISCED code.

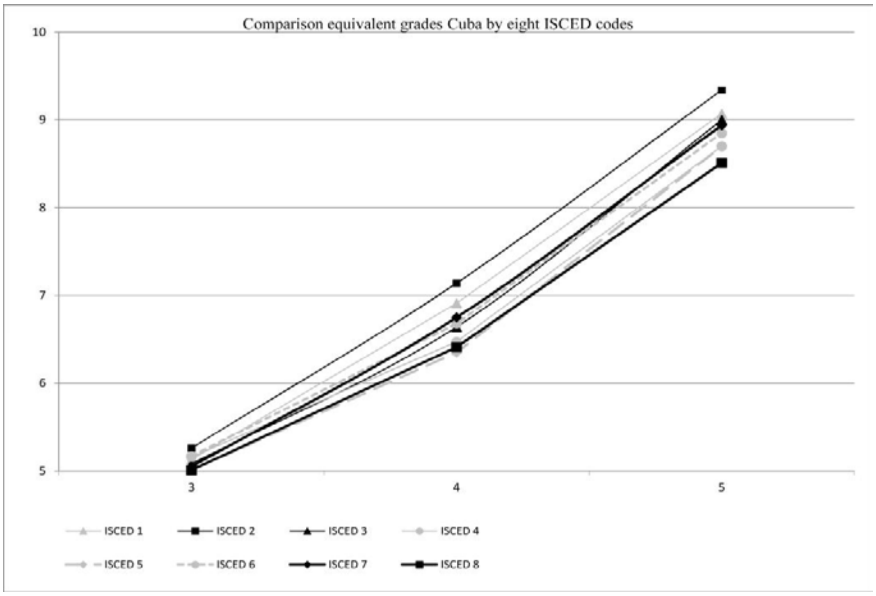

Fig.1. Comparison of equivalent Cuban grades by ISCED codes

4. Conclusions

An objective method, based on the analysis of the distribution of grades, has been introduced for the calculation of equivalences of grades in different systems. The area of knowledge to which the student´s degree is linked should be taken into consideration.

Appendix

Suppose a rating system A (GA), which uses any ordinal scale, and a rating system B (GB), which uses a numerical scale. Table 8 shows information available on these two systems.

Table 8. Distribution of frequencies in the grade systems A and B

Grade System A	nA_i	NA_i	Grade System B	nB_j	NB_j
GA_1	nA_1	NA_1	GB_1	nB_1	NB_1
GA_2	nA_2	NA_2	GB_2	nB_2	NB_2
.....	...	...	.....	...	...
GA_{TA}	nA_{TA}	NA_{TA}	GB_{TB}	nB_{TB}	NB_{TB}

where (similar definitions could be established for system B):

- TA is the number of possible points in the grade system A;
- nA_i is the percentage of credits which have achieved a grade GA_i in the system A;
- NA_i is the cumulative percentage of credits which have achieved a grade lower than or equal to GA_i in the system A.

  Then: $NA_i = \sum_{k=1}^{i} nA_k$; $\sum_{k=1}^{TA} nA_k = 100$; $NT_{TA} = 100$

Take any credit rated GA_i in system A and follow instructions below:

1) Let L be the *first* j such that $NB_j > NA_{i-1}$, where we define $NA_0 = 0$
2) If $NB_L \geq NA_i$ then STOP. The equivalent grade in system B is GB_L
3) If $NB_L < NA_i$ then, for H=L, L+1,… take $n'B_H = nB_H$ while $NB_H < NA_i$, and

   3.1) If $NB_H = NA_i$ then take $n'B_H = nB_H$ and STOP

   3.2) If $NB_H > NA_i$ then take $n'B_H = NA_i - NB_{H-1}$ and STOP

The equivalent grade in system B is $\frac{1}{nA_i} \sum_{H \geq L} n'B_H GB_H$

References

Angoff, W.H. (1984). *Scales, norms and equivalent scores. Princeton, NJ: Educational Testing Service*. (Reprint of chapter in R.L. Thorndike (Ed.), Educational Measurement (2nd ed). Washington, DC: American Council on Education, 1971).

Clarke, G.M. & Cooke, D. (2004). *A Basic Course in Statistics*. Arnold.

Haug, G. (1997). Capturing the Message Conveyed by Grades. Interpreting Foreign Grades. Reprinted from *World Education News & Reviews*, Vol. 10 (2).

Kolen, M.J. (1984). Effectiveness of analytic smoothing in equipercentile equating. *Journal of Educational Statistics,* 9, 209-223.

Linn, R. L. (Ed.). (1993). *Educational measurement*. (3rd. ed.). American Council on Education. Series on Higher Education. Oryx Press.

Petersen, N., Kolen , M. & Hoover, H. (1989). Scaling, norming, and equating. In R. Linn (Ed.), *Educational measurement* (pp. 221-262). New York: American Council on Education. Series on Higher Education. Oryx Press.