

Capturing the Message Conveyed by Grades

Interpreting Foreign Grades

by Guy Haug

Reprinted from World Education News & Reviews, Vol. 10, No. 2, Spring 1997.

Gradings systems differ widely in philosophy and practice from one country to another, and the fair interpretation of foreign grades into national ones is a major issue, both for students returning after a study period abroad and for university staff required to assess the credentials of foreign applicants.

Credential evaluation, credit transfer and grade translation are among the most widely debated and highly sensitive issues in international education, and numerous approaches, solutions, models and formulas have been proposed over the years both in the United States and in Europe.

This article does not intend to propose any particular technique to resolve the issue. It pays more attention to the fundamental needs of interested stakeholders than to the technical tools currently available from professional credit evaluators. Its sole ambition is to recall a few basic rules and principles that tend to be forgotten as the job of translating foreign grades turns into an exercise in accounting or mathematics. The underlying idea in this article is that the first function of grades is to convey a message, and the real challenge in interpreting foreign grades is to render that same message in a different language.

My exposure to the issue of understanding/using foreign grades has been widespread and diversified, but mostly limited to Western Europe and North American systems. In this context, I would distinguish between three main approaches, each guided by a different underlying philosophy.

• **The Inter-university Cooperation Programs (ICPs) developed in the European Union under the ERASMUS program**

Under these exchange schemes set up freely between individual university departments, students spend a study period at a host university abroad and their academic performance there would be fully recognized as part of the degree prepared at the home institution, even though courses abroad may differ substantially from those in the home curriculum.

The basic principle is that "mutual trust and confidence"; grades obtained abroad would be shown on the transcript of the home university. ICPs exchanging large numbers of students among partner universities in several EU countries have gone through an extensive learning process and developed empirical "grading scales" in the forms of charts of the "equivalent" grades at their partner universities. Their specific value is that they are often tailor-made and compare many (if not all) grading systems in use in the EU. Their main limitations are that they are applicable only to short periods of study abroad rather than to entire curricula and that they are negotiated between partner institutions (which entails that they differ substantially from each other: a German 2.3 or an Italian 27 are allocated widely differing foreign equivalents in, for example, the Spanish system, depending on the discipline, institution, and person in charge).

• **The European Credit Transfer System (ECTS)**

ECTS was developed as a pilot scheme under the first phase of the ERASMUS program of the EU and will now be gradually generalized under the new SOCRATES scheme. ECTS has paid considerable attention to the issue of grading, and has introduced a very elaborate "ECTS Grading System" required for use by participating institutions in their ECTS student exchanges.

ECTS goes beyond ICPs, in that it is a whole organized system within which consistency has been sought. The underlying philosophy is that of the equivalence of end products: while the curricula in history, physics, business or engineering may differ in every respect among national systems, the graduates (the "end product") produced by these systems are not all that different. In order to facilitate the transfer of grades

between institutions, "ECTS grades" were introduced with five levels of pass and two levels of fails. They serve as a buffer (or common currency) between different national grades: the host university provides its own national grade and shows the ECTS grade next to the local grade on the student's transcript; the home university in turn uses the ECTS grade and translates it into its national grade, which is used on the student's final transcript.

ECTS offers two distinct advantages: the system is open and can be adapted to all possible national systems (e.g., bridges with Central/Eastern European systems or U.S. grades can be added relatively easily) and it is an interpretative scale rather than a mathematical formula.

• The U.S. Credit Transfer System for Study Abroad

While credit transfer is widespread in the United States, it differs from its younger European counterparts in several important ways: traditional Junior Year Abroad programs are under the direct responsibility of the sending university, and grades are in the U.S. system in order to facilitate the transfer of credits. There are, of course, divergences from this model, especially in cases where students take regular courses taught by the host university and a wide variety of *ad hoc* conversion scales between national and U.S. grades are applied. In many cases, the difficulty of dealing with foreign grades is circumvented as credits are simply given on a pass/fail basis although this penalizes students in good standing by not showing their true achievement. On the other hand, this model has the virtue of a certain type of universality (it is independent of the educational environment in the host country) and the United States has developed considerable professional expertise in assessing credentials and translating grades from all over the world.

Mathematical Formulas Fail to Capture the Message

Both in Europe and in the United States, there have been numerous recent attempts to put together automatic, mathematical formulas that "calculate" foreign grades in the national grading system of the user. In my opinion, these formulas do not produce figures that are a reliable and fair reflection of the message conveyed by the original grade. Their main shortfall is that they cannot adequately deal with certain key characteristics of grading systems:

- Grading systems are not linear and are often characterized by a strongly skewed distribution of grades actually given to students. While American or Italian teachers would use the upper part of their grading scales (albeit in different ways), others (e.g., French and British) in practice hardly ever use the top 20% of their scale. For this reason, proposals based on linear formulas can produce devastating results: I recently saw the case of a German student in France who achieved a 15 (quite a good grade) which was converted into a German 2.5 (a rather mediocre one); on the contrary, a British student who gets a 27/30 in Italy would have every reason to be pleased if that grade were linearly calculated to correspond to a British 90/100!
- Many grading systems are **not continuous**, but divided into several "classes" or "categories" which correspond to broad levels of performance. This means that a small difference in numbers may conceal a substantial difference in meaning when a "class" limit is crossed: in the United Kingdom, a grade of 70 classified as "First Class" is very different from a 69 ("Second Class"), while the same small difference of 1 point is irrelevant between the grades of 54 and 55 (both "Lower Second Class").
- Grading differs **not only between countries**, but there are, as well, marked differences in grading traditions and policies depending on the type and level of the grading institution, the field of study, or even the type of grade (final examination, mid-term, paper, or average computed from various grade items).

Taking France as an example, it is well known that grades at "*classes préparatoires*," which recruit among the best students on their way to "*Grandes Ecoles*," tend to be particularly low, with, for example, 11/20 seen as quite a strong grade, while the pass mark in France is usually an average of 10/20 calculated on all subjects. There may also be minimum pass grades per subject set at a lower level, for example, 8/20.

The distribution of grades tends to be different between certain quantitative fields (with grades distributed

over the whole range) and the non-quantitative fields (where grades are more concentrated in the middle, and the upper part of the scale is seldom used). Thus, even within a given country, a grade may have a "normal," intuitive, abstract meaning which needs to be adjusted (up or downwards) depending on a whole series of factors relating both to who gave it and who interprets it.

From the above observations, my main conclusion is that foreign grades are not just **numbers** that can be calculated by applying a mathematical formula, but a **message** that needs first to be **understood** in the original system and in a second stage **interpreted** by users in their own system.

Simple mathematical formulas with their claim to universality are nothing but a fallacious over-simplification of a reality they fail to capture.

This, however, does not mean that the process of foreign grade interpretation cannot be organized in an efficient, expedient way based on a thorough effort to understand the message that [foreign grades] carry. It is possible to draw up tables ("grade equivalence chart," "grade concordance scale") that render a grade's "normal" or "average" meaning in another grading system, first on a bilateral basis and then in a more multi-lateral context. But this exercise has more to do with the complexity of human language than with mathematics. It takes more listening, modesty and flexibility rather than a doctrinal attitude and a creed in universal formulas/answers. More specifically, the drawing up of tables that can genuinely serve as a basis for interpreting foreign grades is only possible if a certain number of key considerations are observed. The remainder of the article presents six principles that could guide future developments in the area of foreign grade handling.

1. Grade interpretation is no more objective than grading

This is a key consideration: it is a fact of academic life that grades vary, often quite significantly, between institutions, subject areas, and even individual examiners in a given department at a given university. Expectations vary from course to course and from teacher to teacher ... and even over time with the same teacher! Hence, grade conversion scales should not be expected to be more objective than the original grading, and international educators should not be overly sensitive about less-than-perfect conversion scales. Nor should we be overly disturbed that the diverging equivalence tables exist in various contexts of international mobility.

Grade interpretation is no more an exact, objective, universal science than grading itself.

2. Fairness is more important than accuracy

The general attitude towards grade interpretation should be guided by the desire to be fair to students rather than by a vain search for accuracy. In an area marked by subjectivity and diversity, the choice is usually between approximately right and accurately wrong.

But how can fairness be measured and indeed achieved? It seems to me that the only indicator is that the conversion table must provide grades that are in line with the home grades. My experience is that discussions about grade equivalencies are often complicated by emotional reactions where each side insists upon the highest possible foreign grades corresponding to their own grades. This attitude appears to be related to a somewhat defensive, misguided conception of academic pride and leads to a devaluation of foreign grades. Where a dominant partner in an exchange network is able to impose a biased equivalence scheme upon partner institutions, the result is that students from the dominant partner studying elsewhere see their academic performance undervalued when they return home. This can be detected when the performance of students returning from abroad appears to be out of line with either their own previous grades or with those of their classmates who stayed at home.

Structural misinterpretation of foreign grades is unlikely to be detected or corrected easily in the case of one-way mobility. In the case of reciprocal flows, the inevitable effect of a biased conversion scale is that it provides a structural bonus for students moving in one direction while it disadvantages those moving in the opposite direction. These signals are more easily detected in reciprocal exchanges, especially if they involve high levels of student traffic.

3. Grade categories/classes convey core information

In many systems, the full scale of grades is divided not only between pass and fail, but into various "classes" or "categories" corresponding to broad "quality labels" assigned to a certain bracket of numerical grades. Thus, in the United Kingdom, there are "First Class," "Second Class" (divided between upper and lower sub-classes) and "Third Class" performers, while French, German or Spanish students may be labeled in a similar way as, for example, *Passable* (Average), *Gut* (Good) or *Sobresaliente* (Outstanding).

The meaning of these labels in their own context is tainted by culture and tradition. Thus, a British "Third Class" (a pass mark, but usually given only to a relatively small number of very borderline students) is very different from a French *Passable* (a widely-used label that normally applies to the vast majority of pass grades). However tempting it may be, equating *passable* with "Third Class" because they both correspond to the lowest label of "pass grades" would fail to take into account their real meaning.

As a consequence, conversion scales should pay considerable attention to categories/classes of grades. A first priority should be to make certain that this core piece of information is correctly rendered when converting foreign grades; fine-tuning within each particular class/category is only a subordinate exercise: what matters in Britain is whether the grade is a "First" or not, not whether it is a 71 or a 72. This observation is particularly relevant when converting grades from systems using a broad numerical scale into, for example, the U.S. system which usually has only three pass grades (or categories) corresponding to the letters A, B, and C. In the United States, a "D" may also be considered a passing grade, but not for transfer purposes.

The need to pay attention to grade classes reinforces the conclusion that linear methods, which ignore class boundaries, are nothing but fallacious and dangerous over-simplifications. They distort the original message in the same way as a word-for-word check in a bilingual dictionary: for each word there is a corresponding word in the other language, but the sequence of words thus obtained almost certainly means something different (or nothing at all) in the target language.

4. Average grades mean more than individual grades

This is very much related to the previous point: more comprehensive indicators of academic performance abroad convey a more valid message than each of their constituent grades, and should hence receive more attention in the process of interpreting foreign transcripts.

The problem is that in non-linear systems (i.e., in nearly all cases) the mechanical translation of an average grade (using an empirical equivalence chart) will not correspond to the average of the mechanically-translated individual grades from which the average grade was calculated. As a consequence, average grades should be computed in the original system before they are converted into another system. This simple mathematical reality seems confusing to many professionals in international circles. Every now and again, the vain search for a model without this bewildering characteristic brings about deceiving but reassuring proposals based on the simple but wrong assumption of linearity.

5. Reliable conversion scales are transitive

In most cases, institutions need only bilateral conversion scales for incoming/ outgoing students between their own country and one or several foreign countries (e.g., a scale giving U.S. equivalencies for grades from France, Spain, Brazil, etc.). These institutions do not need to convert grades between third countries (e.g., a U.S. university does usually not need to convert Spanish into French grades). Thus, there is no incentive for them to check whether their various bilateral conversion scales are compatible and likely incompatibilities can go unnoticed for a long time.

Yet, there are a few laboratories where grade equivalence needs to be ensured in a multilateral setting and equivalence charts must work simultaneously between all pairs of countries involved. This is the case for a handful of fully integrated, multinational double degree curricula developed under ERASMUS in the European Union, where students go in all directions (e.g. between four partner universities), and their grades

must be converted in a compatible way among all systems involved. The same applies in the case of ECTS, although the situation is slightly different because the common use of "ECTS grades" means in effect that all countries apply only bilateral conversion grades between their own and ECTS grades; yet, a great deal of compatibility between these bilateral scales must exist in order to allow the system to function properly.

The ultimate test of the reliability of equivalence charts is when they are transitive. Transitivity means that the following two exercises produce the same converted grade: (1) a grade from country A is converted into a grade for country B and the grade obtained for country B is converted into a grade for country C; and (2) the same grade from country A is converted directly into a grade for country C.

If, after repeating the exercise various times and in various directions, grades obtained through both calculations are identical or nearly so, then the equivalence charts used for the exercise are unlikely to contain any major structural biases. Developers of all types of grade conversion proposals (be they equivalence tables or mathematical formulas) are invited to submit their proposals to the transitivity test. Usually the results of the test are an invitation to modesty, and sometimes a clear message that the proposed chart needs to be completely reconsidered. Transitivity is of course, all the more difficult to achieve as the number of countries involved grows.

6. Grade interpretation should be done by users

The final interpretation of grades from abroad should be left to the institution that uses them as input for decision making (e.g., to award credits or accept a foreign applicant). In the absence of a universal model for grade interpretation -- even for grades from a particular foreign country -- this is the only way in which the autonomy of each institution can be guaranteed.

What this means in practical terms is that each institution should award grades in its own system and leave the interpretation of those grades in another system to the receiving foreign institution. This basic dual principle is not respected when the grading institution awards grades directly in the system of the using institution (not uncommon in transcripts issued outside of the United States for U.S.-bound exchange students), which in effect imposes pre-translated grades on the using institution, or when the using institution finds its hands bound by an automatic, mechanical conversion model that fails to leave room for interpretation. While conversion should preferably be based on stable tables of equivalencies, these tables only reflect the "normal" or "average" meaning of foreign grades. When there is non-numerical information available (e.g. about "grade inflation" at a given institution), the using institution should have the possibility of adjusting (but not distorting) converted grades to ensure fairness to the student. This may, of course, be misused and open the door to "impressionistic" conversions, but it fundamentally distinguishes grade interpretation from simplistic grade calculation.

In order to safeguard the principle that grades should be interpreted by users and at the same time enhance chances for the correct interpretation of grades, the sending institution should provide information about itself and its grading system. Useful information includes not only maximum and minimum grades, but also grade distribution and class boundaries.

The ECTS grading system is based on a shared code ("ECTS grades") where the encoding is the responsibility of the grading institution and the decoding is left to the using institution. Thus, even in a system based on "mutual trust and confidence" like ECTS, there is some room for interpretation rather than just an automated, numerical exercise. It is also interesting that the network of national academic recognition centers in Europe (known as NARICS and ENICs) is developing a "diploma supplement" appended to transcripts in order to facilitate the interpretation of grades by foreign users. This welcome initiative is jointly supported by the European Union, the Council of Europe and CEPES/UNESCO and should contribute to the education of both graders and grade users and thus reduce the chances that simplistic formulas are used except as a last recourse when nothing else is available.

ECTS Grade	Percent of successful students normally achieving the grade	Definition
A	10%	EXCELLENT - outstanding performance with only minor errors
B	25%	VERY GOOD - above the average standard but with some errors
C	30%	GOOD - generally sound work with a number of notable errors
D	25%	SATISFACTORY - fair but with significant shortcomings
E	10%	SUFFICIENT - performance meets the minimum criteria
FX	-	FAIL - some more work required before the credit can be awarded
F	-	FAIL - considerable work is required

THE CHART BELOW was proposed for a particular consortium of institutions where common examination papers are double marked by home and host campus and hundreds of students are exchanged in all directions annually. It is NOT claimed that this system would be suited for grade conversion outside of this consortium. It neutralizes the British "Third Class" as it was set equal to minimum pass grade in other systems that do not have an equivalent notion as a British "Third."

EXAMPLE OF GRADES/CLASS EQUIVALENCES AMONG ONE CONSORTIUM

GREAT BRITAIN	FRANCE	GERMANY	SPAIN	UNITED STATES
100-90	20.0-18.0	1.0	10.0	
89	17.8	1.0	9.9	
88	TRES 17.6	1.1	9.9	
87	BIEN 17.4	1.1	9.9	
86	17.2	1.2	9.8	A+
85	17.0	1.2	9.8	
84	16.8	1.3	9.7	
83	16.6	1.3	9.7	
82	16.4	1.3	9.6	
81	16.2	1.4	9.6	
FIRST CLASS 80	16.0	SEHR 1.4	SOBRE- 9.5	
79	15.8	GUT 1.5	SALIENTE 9.4	
78	15.6	1.5	9.4	
77	15.4	1.5	9.4	A
76	15.2	1.5	9.3	
75	15.0	1.6	9.3	
74	BIEN 14.8	1.6	9.2	
73	14.6	1.7	9.2	
72	14.4	1.7	9.1	
71	14.2	1.8	9.1	A-
70	14.0	1.9	9.0	
69	13.8	2.0	8.8	
68	13.6	2.1	8.6	
67	13.4	2.2	8.4	B+
UPPER 66	ASSEZ BIEN 13.2	2.3	NOTABLE 8.2	
65	13.0	2.4	8.0	
SECOND CLASS 64	12.8	2.5	7.8	
63	12.6	2.6	7.6	
62	12.4	2.7	7.4	B
CLASS 61	12.2	2.8	7.2	
60	12.0	2.9	7.0	
59	11.8	3.0	6.8	
58	11.6	3.1	6.6	
LOWER 57	11.4	BEFRIEDIGEND 3.2	BUENO 6.4	B-
56	11.2	3.3	6.2	
SECOND CLASS 55	PASSABLE 11.0	3.5	6.0	
CLASS 54	10.8	(AUSREICHEND) 3.6	APROBADO 5.8	
53	10.6	3.7	5.6	
52	10.4	3.8	5.4	
51	10.2	3.9	5.2	C
THIRD CLASS 40-50	10.0	4.0	5.0	
FAIL 39	9.9	4.1	4.9	
38	ECHEC 9.0	NICHT AUSREICHEND 4.3	SUSPENSO 4.7	F
37	8.0	4.5	4.5	
35	7.0	4.9	4.0	
30	6.0	5.0	3.5	
25	5.0	5.1	3.0	
20	4.0	5.3	2.5	
15	3.0	5.5	2.0	
10	2.0	5.7	1.5	
5	1.0	5.9	1.0	
0	0.0	6.0	0.0	