

Analysing students' academic performance in Higher Education in Spain

*Analizando los resultados académicos en la Educación Superior en España*¹

Carlos Rivero (crivero@estad.ucm.es)

Department of Statistics and Operational Research II, Universidad Complutense de Madrid, Madrid, Spain.

Cristina del Campo (campocc@ccee.ucm.es)

Department of Statistics and Operational Research II, Universidad Complutense de Madrid, Madrid, Spain.

Elena Urquia-Grande (eurquiag@ccee.ucm.es)

Department of Financial Economics and Accounting II, Universidad Complutense de Madrid, Madrid, Spain.

Maria-del-Mar Camacho-Miñano (marcamacho@cunef.edu)

Accounting Department, CUNEF (UCM) Madrid, Spain.

David Pascual-Ezama (david.pascual@ccee.ucm.es)

Department of Financial Economics and Accounting II, Universidad Complutense de Madrid, Madrid, Spain.

<http://dx.doi.org/10.12795/EDUCADE.2017.i08.02>

ABSTRACT: The development of the European Higher Education Area has been a process of modernization in many universities. Teaching methodologies have undergone a process of continuous change to meet the demands for high quality leading to a need for enhancement in the learning assessment methodologies as well.

The objective of this study is to analyse student's academic performance measured through coursework vs. final exam and to ascertain the factors that could explain the difference. Regression and variance analysis are carried out over the grades and responses to a questionnaire on a sample of 298 students of different subjects in a Spanish university. The results show that there are differences between continuous assessment and the final examination marks.

KEYWORDS: assessment; coursework; final exam; students' outcomes and preferences.

Artículo. Recibido: 08-03-17 – Versión revisada: 22-06-17, Aceptado: 11-07-17

Licencia Creative Commons BY NC ND · 2017 · Asociación Española de Contabilidad y Administración de Empresas - AECA

¹ This research was partially supported by the Spanish Ministry of Economy and Competitiveness under the RD project Inte-R-LICA (The Internationalisation of Higher Education in Bilingual Degrees) for the period 2014-2016 (Ref. FFI2013-41235-R).

1. INTRODUCCIÓN

The European Higher Education Area (EHEA) has stimulated universities towards a process of continuous improvement. There have been changes in universities strategies such as the teaching methodology which has been modified into a student-centred approach, the use of multimedia resources in education, the development of new scopes or the new degree study plans. Therefore, the process of learning assessment must also change in order to complement these new teaching and learning trend. Furthermore, the EHEA has established the European Credit Transfer System (ECTS) system designed for all European Universities based on homogeneous and transparent evaluation through countries (European Communities, 2009), used for student exchanges in universities around the world. Research in assessment methodology in Higher Education has been defined as serving multiple purposes such as providing program and institutional accountability, improving teaching quality, information about student learning and following student progress (Fletcher, Meyer, Anderson, Johnston and Rees, 2012; Frick, Chadha, Watson and Zlatkowska, 2010).

Many universities have made changes in the assessment methodology to a more holistic system embodying both the students' daily effort with different coursework projects and the final exam (Segers and Doch, 2006). Due to the extensive use of new teaching tools such as simulations, problem-based learning or multimedia materials, all the elements of the learning process: syllabus, teaching methodology, objectives, resources, learning outcomes and type of assessment should be adjusted (Camacho et al., 2016; del Campo and Camacho, 2015). There are few studies with empirical evidence demonstrating how different types of assessment have an impact on students' academic outcomes (Tynjal, 1998). Also, students' grade measurement has an important impact on students' behaviour and, consequently, on learning outcomes (Payne and Brown, 2011; Smith, 2011).

Traditionally, only the final exam was used in the final grade to measure students' learning outcomes but there is a general recognition of the negative and limiting impact of using only those final exams on the students' learning process assessment (Muldoon, 2012). Nowadays, final grading is composed by a sum of different ways of continuous assessment and several exams (interim and final), which measure not only the cognitive competences acquired in the subject but also the development of meta-cognitive skills such as leadership, negotiation, communication or decision making skills (Cano, Chamizo and Martin, 2016). It has to be highlighted that formal final exams refer to closed-book time-constrained written essays, test or exercises similar to the traditional form of assessment while coursework refers to alternative types of assessment including group work, simulations, video and project presentations or long essays about a topic (Camacho-Miñano et al., 2016).

The final exam mark should be the result of the learning process that the coursework had been stimulating. However, a majority of authors confirm coursework results are higher than final exam marks (Bridges et al., 1999 and 2002; Dalziel, 1998; Downs, 2006; Gibbs and Simpson, 2004-05; James and Fleming, 2005; Murdan, 2005; Pascual-Ezama, et al., 2011; Simonite, 2003; Tian, 2007; Yorke, Bridges and Woolf, 2000; Yorke, Cooper and Fox, 1996; Dafouz, Camacho and Urquía, 2014). In general, there are some motives and factors which justify the differences between coursework and exams but the results are not conclusive (Payne and Brown, 2011). There are authors who demonstrate the deviation between coursework and the final exam is explained by the discipline. In

mathematical or quantitative subjects, coursework grades are significantly higher than final exam ones while in qualitative subjects the difference between coursework grading and the final exam is lower (Bridges et al., 2002; Murdan, 2005; Simonite, 2003). Specifically, Yorke et al. (2000) demonstrated that coursework marks were higher than final exams in Computer Science and Mathematics subjects while the difference shortens in Business Studies subjects. Other authors explain the differences depending on the students' personality and the taken course demonstrating that students with a high level of performance tend to do better in the final exams of Anatomy versus Physiology where these same students have lower performance (James and Fleming, 2005). When the assessment criteria accomplish to involve students in different learning activities and their environment, this determines their learning outcomes (Struyen et al., 2008) or their probability to pass the subject (Durán Santomil et al., 2013). However, if course workload is excessive, it will provoke a "surface" learning approach (Yorke et al., 2000; Dafouz, Camacho and Urquía, 2014).

There are studies showing that gender (Woodfield, Earl-Novell and Solomon, 2005), age (Kniveton, 1996), use of information technologies (Chamizo et al., 2015; López Pérez, et al., 2013), cultural differences (Baeten, Dochy and Struyven 2006) or learning strategies (Rivero-Menéndez et al., 2017) are other factors affecting the difference between coursework and exam marks but again with no conclusive results. Additionally, if the subject has been taught in another language (in the majority of the cases English as a medium of instruction, EMI), some researchers have demonstrated that EMI students perform better in coursework than in exams because they have time to express their ideas in essays or projects in groups (Dafouz, Camacho and Urquía, 2014; DeVita, 2002; Jackson, Meyer and Parkinson, 2006).

The objective of the present paper is twofold. Firstly, the authors want to analyse the differences between students' coursework and final examination marks. Secondly, we investigate the factors, such as students' learning strategies, type of degree (Economics, Business Administration or Computer Science), the language in which it is taken (English as a medium of instruction or Spanish), that have an impact on the final academic performance.

Based on the previously mentioned literature, the research presented here aims to provide an answer to the following two research questions:

RQ1: Is the students' coursework mark different from the final examination mark?

RQ2: What factors are associated with the difference between students' final grade and coursework marks?

2. METHOD AND CONTEXT

The context of the study is the University Complutense of Madrid (UCM), the largest university in Spain with over 75,000 students and around 6,000 teachers. During one academic year the university offers 65 different degrees from medicine to computer science, 105 official master's degrees and doctoral programmes. To date the UCM is one of the most recognized universities in the Spanish context according to the rankings.

Participants

The participants were 428 out of 505 freshmen enrolled on the Statistics and/or Financial Accounting courses in the Business Administration, Economics or Computer Science

undergraduate degrees in UCM in one academic year who did the coursework throughout the semester.

The students were asked for their consent to include their data in the present research on the first lecture of the semester, clearly stating that participation was voluntary and data would be treated anonymously in agreement with the Spanish Organic Act on Data Protection (1999). The students were also asked to fill in a pen-and-paper questionnaire during the last week of the semester. None of them declined to cooperate.

Measures

The analysed research variable is the final grade obtained by the already mentioned 428 freshmen enrolled on the Statistics and/or Financial Accounting courses in the Business Administration, Economics or Computer Science undergraduate degrees in the academic year. In Spain, grades range from 0 to 10, where 0 means the worst possible result and 10 the best one, but they are divided into two intervals: grades in [0, 5) mean failure and grades in [5, 10] mean pass. With the Bologna methodology the final exam is not the only component of the final grade. In fact, the final exam (FE), which is a closed-book time-constrained examination, has a weight of only 0.5, 0.6 or 0.7 depending on the bachelor degree (Business Administration, Economics or Computer Science, respectively). The other part of the final grade, called coursework (CW), is composed of active participation, assignments given in usually through the virtual platform (exercises, cases, simulations, real-world problems...) and interim class tests (Chamizo-Gonzalez et al., 2015; Heywood 2000). Also the students have two opportunities in the year to sit for the final exam and pass the subject, while the coursework component is obtained only during the lecturing period.

Furthermore, in Spain students are allowed to re-enrol in a subject where they have failed before up to the fifth call. For 43% of the respondents this was their first enrolment, but for 47% it was their second call, while for the other 10% it was their third, fourth or even fifth call.

Out of the 428 initially enrolled students, only 322 finally sat for the ordinary final exam. Those are the students' results analysed in the following study. Also, students were asked to fill in a questionnaire of 17 items divided into three sections: demographic data (age, gender, nationality and work status), background data (university access examination grade so-called PAU, degree position in university application), academic achievement (number of calls and maths grade), preferred evaluation type, and language medium of instruction (EMI).

Out of the 322 that sat for the final exam only 298 were in class the day the survey was done, although some of them did not answer all of the questions. Those missing data were not considered and the analyses were run on what remains.

Statistical analyses

Two quantitative and ten qualitative factors, coming from the questionnaire, will be taken into account (see Table 1 for a complete description). The authors made this previous distinction because to analyse whether the difference between coursework and final exam is affected by any of the quantitative variables, a regression analysis is performed, while to analyse whether it is affected by any qualitative variables a variance analysis is used.

In order to be able to answer the research questions, regression and variance analyses were carried out on the grades and the answers to the questionnaire. The coursework (CW) and the final exam marks (FE), as well as the difference between them (Diff) were used as dependent variables whereas the other 12 were used as explanatory variables (see Table 1). The variable Diff = CW – FE is defined to study the difference between coursework and final exam. In particular, this variable is defined to study if it is significantly positive. It is a measure of the 'deviation' between the coursework and the final exam mark of each student, so that if it is positive, the coursework mark is higher than the final examination. Furthermore, using this difference allows us to avoid the bias problems among grades coming from different subjects.

Table 1. Variables used in the study

| | Variable | Type |
|-----------------------|-------------------------------------------|--------------|
| Independent variables | Coursework mark (CW) | Quantitative |
| | Final Examination mark (FE) | Quantitative |
| | Diff = CW – FE | Quantitative |
| Explanatory Variables | Subject Final Grade (FG) | Quantitative |
| | Gender | Qualitative |
| | Nationality | Qualitative |
| | Work status | Qualitative |
| | Maths grade | Quantitative |
| | University access examination grade (PAU) | Quantitative |
| | Number of calls | Quantitative |
| | Preferred evaluation type | Qualitative |
| | Teaching language | Qualitative |
| | Motive for degree choice | Qualitative |
| | Degree position in university application | Quantitative |
| | Preferred type of lecturer | Qualitative |
| | Study method | Qualitative |
| | Learning style | Qualitative |
| | Team work preferences | Qualitative |
| | Degree | Qualitative |
| Subject | Qualitative | |

The final grade of a student (FG) is a weighted average between the coursework (CW) and the final examination marks (FE). Specifically $FG = (1-a) \cdot CW + a \cdot FE$, where "a" is the weight of the Final examination mark that can be 0.5, 0.6 or 0.7, depending on the degree the student is doing as explained above. This final grade can be interpreted as

a measurement of the student's ability and effort as the students have to obtain a high grade in both the coursework and the Final examination, if they want a high final grade.

The PAU (from the Spanish for University Access Examination) is a compulsory general examination by knowledge areas to determine the University Access Grade. In some studies, the university access exam is considered a proxy of cognitive skills (Pascual-Ezama et al. 2011).

3. RESULTS

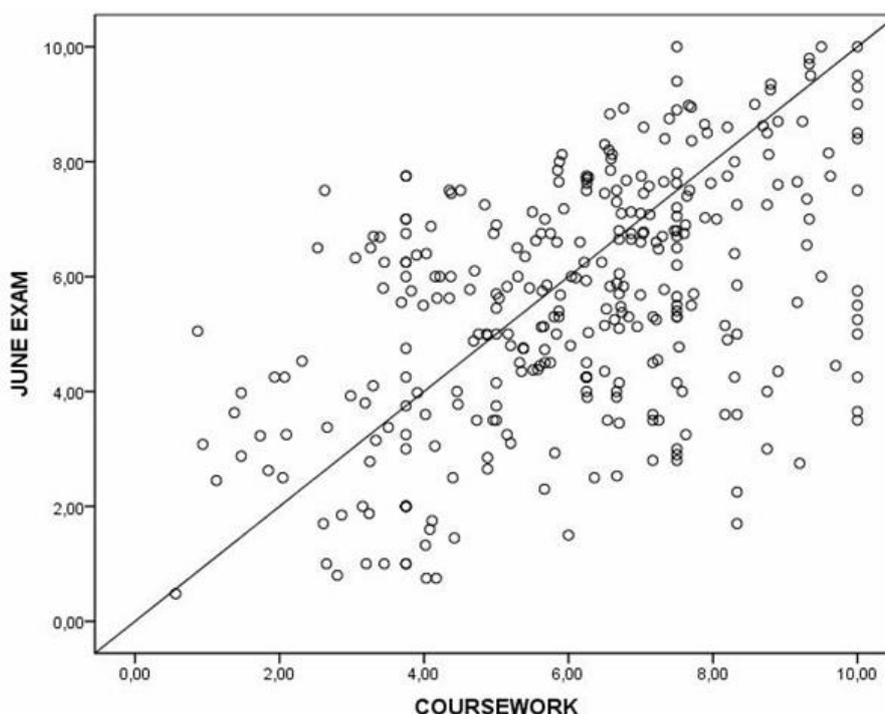
The results of the present study are divided in two parts due to the two proposed research questions. On one hand, the outcome of students' assessment where the coursework and the examination marks will be analysed. Later the factors affecting students' final assessment will be shown.

Relation between coursework and final examination marks

In order to answer the first research question, we will analyse the existing relation between the Final examination and the coursework marks, calculated as a weighted mean of the five to seven (depending on the subject) assignments (consisting of oral presentations, computer lab exercises, simulations, videos and collaborative case studies), active participation and interim test marks.

Figure 1 plots the coursework marks against the final examination marks for the 322 students in the sample. It can be seen that the majority of the points are below the diagonal (CW = FE) indicating that for most of the students the coursework marks are higher than the final examination ones. Specifically, there is 59.6% of the students who achieve a higher coursework mark (CW) than the final exam (FE) one.

Figure 1. Coursework mark against "June examination" mark scatterplot



The Pearson correlation coefficient between coursework marks and final exam grades is 0.7543 with a p-value < 2.2e-16, meaning there is a significant strong positive linear relationship between the two of them. Therefore, as the value of one variable increases, so does the value of the other variable.

Some descriptive statistics of Diff = CW – FE are in Table 2 where “Diff” has a mean of 0.5643 with a 95% confidence interval so, on average, the deviation in marks between coursework and final exam is significantly positive. Concretely, the coursework mark (CW) is higher than the final examination mark (FE).

Table 2. Descriptive statistics of Diff = CW- JE

| | | Statistic | Tip. Error |
|---------|--------------------------------------|-----------|------------|
| Diff | Mean | 0.5643 | 0.11912 |
| | 95% Confidence Interval for the mean | 0.3299 | |
| | Lower bound | 0.7986 | |
| | Upper bound | 0.5206 | |
| | 5% trimmed mean | 0.4850 | |
| | Median | 4.569 | |
| | Variance | 2.13749 | |
| | Std. Dev. | -4.88 | |
| | Minimum | 6.63 | |
| Maximum | | | |

Also, as “Diff” variable values follow a normal distribution (Shapiro-Wilk normality test p-value = 0.01509) a t-Student test is applied to verify whether “Diff” mean is zero or positive. The p-value (p-value = 3.886e-06) of the bilateral test and the confidence interval for the mean (0.3299, 0.7986) included in table 2 show “Diff” mean is significantly positive, so that coursework marks are in average higher than final exam grades.

Factors associated with the difference between students' final grade and coursework marks

In order to answer research question 2, this section is devoted to the analysis of the factors affecting the Diff = CW – FE variable.

It was found that out of the sixteen factors analysed only six had any influence on the “Diff” variable, namely the final grade, the number of exam calls, the preferred evaluation type, PAU, the teaching language and the work status. However, neither the Mathematics grade, the gender, the nationality, the motive for choosing the degree, the degree position in the university application, the preferred type of lecturer, the study method, the learning style, the team work preferences nor the degree or the course were found to have any significant influence on the deviation mark (Diff variable). The six variables that were found to have influence on the “Diff” variable (coursework and final grade differences) will now be analysed in detail.

Final grade

There is a significantly negative correlation between the mark deviation (Diff) and the final grade (FG), therefore a regression analysis is carried out to model the mark

deviation from the final grade. The ANOVA F statistic show there is a significant relation between the estimated regression coefficients (Table 3).

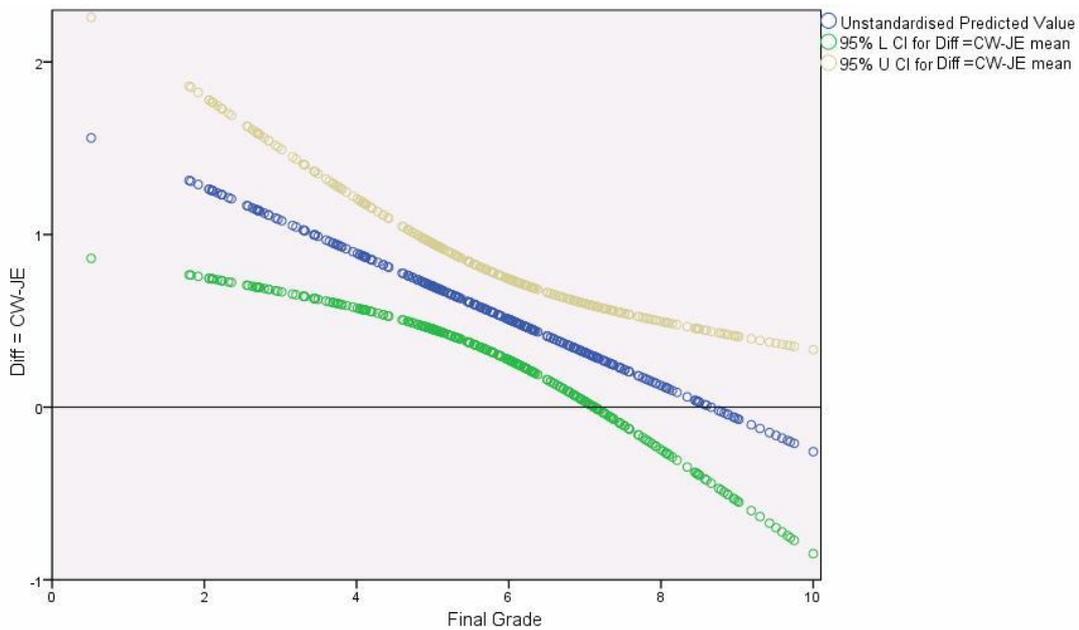
Table 3. Regression Coefficients

| Model | Non-standardised coefficients | | Typified coefficient | T | Sig. | 95% Confidence Interval for B | |
|----------|-------------------------------|------------|----------------------|--------|--------------|-------------------------------|-------------|
| | B | Tip. Error | Beta | | | Lower bound | Upper bound |
| Constant | 1.658 | 0.386 | | 4.296 | 0.000 | 0.899 | 2.417 |
| FG | -0.192 | 0.064 | -0.164 | -2.975 | 0.003 | -0.318 | -0.065 |

a. Dependent variable: Diff

The regression line (in the middle of the figure) is represented in Figure 2 together with the 95% confidence band. For those students whose upper and lower band values are above zero, on average, the mark deviation variable is positive, meaning the coursework mark is higher than the final exam one (CW > FE). The aforementioned situation occurs only for those students with final grade below seven (FG < 7.0) because for those students whose final grade is higher than 7.0, the 95% confidence bands have a negative and a positive part, so no conclusion can be drawn for them. Hence, for students with a final grade above 7 the difference between coursework or final exams is unclear.

Figure 2. Confidence band for variable "Diff" over "final grade" regression line



University Access Grade

There is a significant negative correlation of -0.124 between "Diff" and PAU, so a regression analysis is carried out to model Diff from PAU. The ANOVA F statistic is 3.177 with a signification of 0.076, while the estimated regression coefficients are in Table 4.

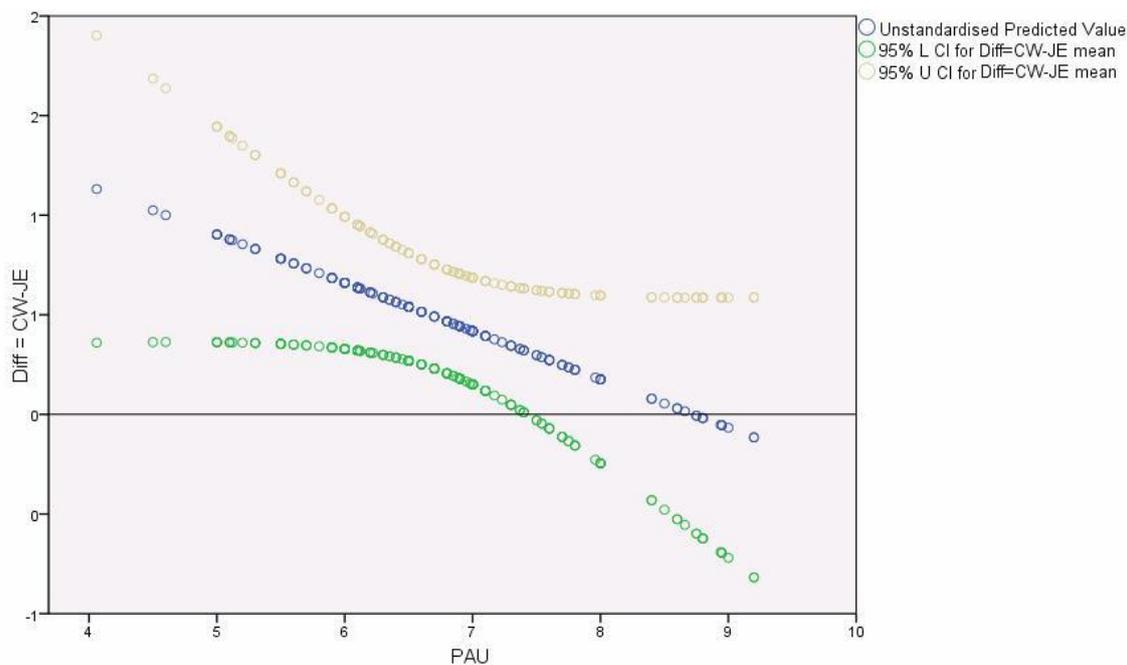
Table 4. Regression coefficients

| Model | Non-standardised coefficients | | Typified coefficients | t | Sig. | 95% Confidence interval for B | |
|----------|-------------------------------|------------|-----------------------|--------|--------------|-------------------------------|-------------|
| | B | Tip. Error | Beta | | | Lower bound | Upper bound |
| Constant | 2.116 | 0.930 | | 2.275 | 0.024 | 0.282 | 3.949 |
| PAU | -0.242 | 0.136 | -0.124 | -1.783 | 0.076 | -0.511 | 0.026 |

a. Dependent variable: Diff = CW-JE

Similarly to the Final Grade previously analysed, the “Diff” variable is significantly positive for those students with a PAU lower than 7.5, and it is inconclusive for those with a PAU higher than 7.5 (see Figure 3). Hence when students obtained in the University Access Examination a mark lower or equal to 7.5 their coursework grades are higher than the final exam, while when the students have obtained an University Access Grade higher than 7.5 the difference between coursework or final exams is unclear.

Figure 3. Confidence band for variable “Diff” over “PAU” regression line



Number of exam calls

As it was stated before, there are students who had previously sat for the exam (from one to four times in our sample) and some others that are taking the subject for the first time. The authors' past experience led them to create a new variable based on the belief that the difference is mainly in their being new to the course or not. Therefore (see table 5), the new variable (retaker) has only two values: zero and one, meaning respectively whether it is the first time the student taking the course (value = 0) or not (value = 1).

Table 5. Number of calls descriptives

| | N | Mean | Std. Desv. | Tip. Error | 95% confidence interval for the mean | | Min. | Max. |
|---------------------|-----|---------|------------|------------|--------------------------------------|-------------|-------|------|
| | | | | | Lower bound | Upper bound | | |
| 0 (First enrolment) | 102 | -0.0249 | 1.63006 | 0.16140 | -0.3451 | 0.2953 | -3.97 | 3.86 |
| 1 (re-enrolment) | 135 | 0.8993 | 2.14983 | 0.18503 | 0.5334 | 1.2653 | -4.88 | 6.63 |
| Total | 237 | 0.5015 | 1.99290 | 0.12945 | 0.2465 | 0.7566 | -4.88 | 6.63 |

As it can be deduced from the ANOVA analysis, if the student had previously taken the course (value = 1), the Diff variable is significantly positive (ANOVA p-value = 0), while the result is inconclusive when the student takes the course for the first time (value = 0). Therefore, if the student had previously taken the course, the coursework mark is significantly higher than the final exam mark.

Preferences over the evaluation method

In question number 15 of the survey the students answered about their preferred evaluation method. It was a closed question, answered by only 235 students, with only two options: just final exam (traditional methodology) or continuous evaluation (final exam weighted with coursework understood as collaborative group work, simulations, or project presentations and real-life exercises).

As it could be expected (see table 6) the Diff variable is significantly positive (ANOVA p-value = 0.048) for those students preferring continuous evaluation, while it is inconclusive for those students preferring only the final exam. Hence, if the student prefers the continuous assessment method, the coursework mark is significantly higher than the final exam mark.

Table 6. Preferences over "evaluation method" descriptives

| | N | Mean | Std. Desv. | Tip. Error | 95% confidence interval for the mean | | Min. | Max. |
|------------|-----|---------|------------|------------|--------------------------------------|-------------|-------|------|
| | | | | | Lower bound | Upper bound | | |
| Continuous | 198 | 0.6106 | 2.00824 | 0.14272 | 0.3291 | 0.8920 | -4.88 | 6.63 |
| Only exam | 37 | -0.0972 | 1.87523 | 0.30829 | -0.7224 | 0.5281 | -3.97 | 4.75 |
| Total | 235 | 0.4992 | 2.00076 | 0.13052 | 0.2420 | 0.7563 | -4.88 | 6.63 |

Language: EMI versus non-EMI

It has to be noted, as it was already stated before, that some of the courses subject to the present study were taught in English as a medium of instruction (EMI), specifically four out of eight were taught in English. It came as a surprise to discover that the "Diff" variable is significantly positive (ANOVA p-value = 0.011) for those courses taught in Spanish (non-EMI) while it is inconclusive for those taught in English (EMI) (see Table 7).

These results show that if the student takes the course in Spanish (non-EMI), the coursework mark is significantly higher than the final exam mark.

Table 7. Language descriptives

| | N | Mean | Std. Desv. | Tip. Error | 95% confidence interval for the mean | | Min. | Max. |
|-------|-----|--------|------------|------------|--------------------------------------|-------------|-------|------|
| | | | | | Lower bound | Upper bound | | |
| | | | | | Non-EMI | 231 | | |
| EMI | 91 | 0.0846 | 1.80810 | 0.18954 | -0.2919 | 0.4612 | -4.88 | 3.86 |
| Total | 322 | 0.5643 | 2.13749 | 0.11912 | 0.3299 | 0.7986 | -4.88 | 6.63 |

Work status

Although the students subject to the present study are freshmen, an item related with work status was included in the questionnaire due to the specificity of the considered degrees (Business Administration, Economics and Computer Science). Indeed, out of the 255 answers to this item, 40% of students declared to be working (Table 8). What is more, Diff is significantly positive (ANOVA p-value = 0) if the student has a job, while it is inconclusive for those students not working. Thus, if the student has a job, the coursework mark is significantly higher than the final exam mark.

Table 8. Work status descriptives

| | N | Mean | Std. Desv. | Tip. Error | 95% confidence interval for the mean | | Min. | Max. |
|---------|-----|--------|------------|------------|--------------------------------------|-------------|-------|------|
| | | | | | Lower bound | Upper bound | | |
| | | | | | Not working | 135 | | |
| Working | 90 | 0.6840 | 2.06898 | 0.21809 | 0.2506 | 1.1173 | -4.00 | 6.63 |
| Total | 255 | 0.4008 | 1.96150 | 0.13077 | 0.1431 | 0.6585 | -4.88 | 6.63 |

Interactions

After considering the "main effect" of each of the independent variables, we are going to study their interactions as the effect of one independent variable on our dependent variable may not be the same at all categories of other independent variables. In statistics, an interaction between independent variables A and B implies that the effect of A depends on the value of B and that the effect of B depends on the value of A.

A backward stepwise selection method was applied to the original model including all the possible interactions between the factors with influence on the dependent variable Diff, so only the significant effects are kept on the new model. The stepwise process ends when none of the effects outside the model has a significant statistic and every effect in the model is significant.

There are significative interactions (see table 9) between the following groups of factors: PAU and retaker; retaker and evaluation[only exam]; retaker and language[sp]; evaluation[only exam] and language[sp]; retaker and work[yes]; FG,

retaker and evaluation[only exam]; PAU, retaker and evaluation[only exam]; PAU, retaker and language[sp]; PAU, evaluation[only exam] and language[sp]; FG, evaluation[only exam] and work[yes]; FG, PAU, retaker and evaluation[only exam]; FG, PAU, evaluation[only exam] and work[yes].

The model's deviance of 1094.2 on 304 degrees of freedom is not significant at the conventional five per cent level, so we have no evidence against this model.

Table 9

| | Estimate | Std. Error | t value | Pr(> t) |
|----------------------------------------|-----------|------------|---------|------------|
| (Intercept) | 8.7355 | 6.5311 | 1.338 | 0.18219 |
| FG | -1.7670 | 1.1264 | -1.569 | 0.11791 |
| PAU | -0.9614 | 0.9903 | -0.971 | 0.33254 |
| retaker | -31.7955 | 10.2299 | -3.108 | 0.00209 ** |
| evaluation[ONLY EXAM] | 32.0806 | 37.8235 | 0.848 | 0.39711 |
| language[SP] | 3.6935 | 5.7263 | 0.645 | 0.51947 |
| work[YES] | -4.6150 | 7.3860 | -0.625 | 0.53262 |
| FG:PAU | 0.2105 | 0.1646 | 1.279 | 0.20203 |
| FG:retaker | 1.9726 | 1.2180 | 1.619 | 0.10653 |
| PAU:retaker | 4.7038 | 1.5319 | 3.071 | 0.00236 ** |
| FG:evaluation[ONLY EXAM] | -5.8852 | 7.3660 | -0.799 | 0.42503 |
| PAU:evaluation[ONLY EXAM] | -4.5243 | 5.5430 | -0.816 | 0.41511 |
| retaker:evaluation[ONLY EXAM] | 218.0317 | 108.2343 | 2.014 | 0.04497 * |
| FG:language[SP] | 0.2275 | 0.3936 | 0.578 | 0.56374 |
| PAU:language[SP] | -0.7330 | 0.8793 | -0.834 | 0.40524 |
| retaker:language[SP] | 22.6616 | 9.6969 | 2.337 | 0.02018 * |
| evaluation[ONLY EXAM]:language[SP] | -107.3717 | 43.9903 | -2.441 | 0.01531 * |
| FG:work[YES] | 0.8484 | 1.2424 | 0.683 | 0.49526 |
| PAU:work[YES] | 0.6384 | 1.0683 | 0.598 | 0.55062 |
| retaker:work[YES] | 1.6356 | 0.6729 | 2.431 | 0.01573 * |
| evaluation[ONLY EXAM]:work[YES] | -117.7770 | 82.8205 | -1.422 | 0.15618 |
| FG:PAU:retaker | -0.2207 | 0.1646 | -1.341 | 0.18108 |
| FG:PAU:evaluation[ONLY EXAM] | 0.8307 | 1.0719 | 0.775 | 0.43905 |
| FG:retaker:evaluation[ONLY EXAM] | -40.2381 | 20.2121 | -1.991 | 0.04753 * |
| PAU:retaker:evaluation[ONLY EXAM] | -33.0583 | 16.2715 | -2.032 | 0.04318 * |
| FG:retaker:language[SP] | -0.7003 | 0.5100 | -1.373 | 0.17092 |
| PAU:retaker:language[SP] | -3.1393 | 1.4951 | -2.100 | 0.03669 * |
| PAU:evaluation[ONLY EXAM]:language[SP] | 16.3376 | 6.7485 | 2.421 | 0.01615 * |
| FG:PAU:work[YES] | -0.1262 | 0.1748 | -0.722 | 0.47080 |
| FG:evaluation[ONLY EXAM]:work[YES] | 43.6173 | 20.3569 | 2.143 | 0.03305 * |
| PAU:evaluation[ONLY EXAM]:work[YES] | 16.9306 | 12.5038 | 1.354 | 0.17688 |
| FG:PAU:retaker:evaluation[ONLY EXAM] | 5.9723 | 3.0027 | 1.989 | 0.04772 * |
| FG:PAU:evaluation[ONLY EXAM]:work[YES] | -6.3685 | 3.0328 | -2.100 | 0.03668 * |

SIGNIF. CODES: 0 '**' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

Interpreting the previous interactions is not easy when there are more than two independent variables and some of them are continuous and others are qualitative. For the sake of clarity, we will describe some interpretation examples. For example, the effect student's preference of evaluation is not the same if the student is retaking the course. It can be observed that being a retaker combined with language of instruction and the preferred type of evaluation is affecting the difference between coursework and final exam achieving higher coursework marks. Also, the effect of the student's preferred type of evaluation is not the same when the language of instruction is English or Spanish together with the different values of PAU.

4. DISCUSSION AND CONCLUSIONS

The objective of this study was to show empirical evidence that there is a difference between student's coursework and final examination marks. Additionally, this research analysed the factors that could explain the difference between those two marks in the subjects under study (Financial Accounting and Statistics from the Business Administration, Economics and Computer Science degrees).

The difference between coursework (weighted average of the required assignments, oral and active participation and interim tests) and final exam marks was analysed. The deviation between both marks is significantly positive being coursework mark higher in average than final exam marks. This result is probably because lecturers are making an effort for the students to achieve a continuous progress and learning along the whole academic year in line with Yorke et al., 2000. Interestingly there is a 40% of the participants whose final exam mark is higher than the coursework mark. These students are analysed further (through their marks in final exam in the subjects and PAU) and seem to be the students who perform better (Final grade > 7 and PAU > 7.5). This can be explained because the best students perform better individually in a unique final exam, however this assumption should be further researched.

In our results the students who preferred continuous evaluation have the higher difference values between final exam and their coursework figures. The student preference over continuous assessment through coursework is in line with other studies (Starr 1970; Woodfield et al. 2005; Dafouz, Camacho and Urquía, 2014) that reported that most students preferred all their marks to come from continuous evaluation, as it is a better evaluator of meta-cognitive abilities and effort. This is in accordance with the students' meta-cognitive strategies that involve students thinking about themselves immersed in the whole learning process, planning for learning, monitoring subject comprehension and self-analysis and self-evaluation in line with Rivero-Menéndez et al., 2017. Additionally, students who preferred to participate and collaborate in class with the teacher (Swain 1985) and with their peers will achieve better academic results in coursework in the medium term (O'Mally and Chamot, 1990). Thus this collaboration with peers and teachers increases the amount of content negotiated and students learn from the exchange in classroom context (Long 1991; Pica 2000; del Campo and Camacho-Miñano, 2015).

The number of calls had a high influence in the final grade and this can be explained because coursework is considered a more reliable approach to assessment and offers more potential to student's deep learning approach when compared with final exams. However, coursework is more vulnerable to plagiarism and impersonation (Yorke et al. 2002) and this can be in line with our results where students who have previously taken the subject already know coursework assignments and do them better.

There is an important issue about students being constrained by their knowledge or skills in a subject learnt in English as a medium of instruction (EMI) (Smith 2011; del Campo et al., 2015). At this point it is important to divide the outcomes in coursework and final exam because the EMI students performed much better in the coursework than in the final exams in line with the DeVita (2002) and Jackson, Meyer and Parkinson (2006) who claim EMI students perform better in the coursework as they have time to adequately express their thinking in a second language while this is not the case in time-constrained exams. Another reason could be that the teacher takes into account the student's context when grading the coursework (Bridges et al., 2002).

The first conclusion of our study is that students prefer continuous assessment because, on average, they get higher marks in coursework than in exams and perceive their learning is deeper. Students do not like exams because of the anxiety and stress of assessment, with only closed-book time-constrained exercises or questions. The second conclusion is that the type of students, the capacity of students' effort and the degree assessment criteria approved by faculties are factors affecting assessment. Interaction between retaking the course, the language of instruction and the preference of evaluation also determine the difference between final exam and coursework.

Our results have clear implications for the stakeholders of Higher Education. Students should know that effort and motivation are essential in their academic outcomes. Thus, the EHEA has been drawing attention to the need to link assessment with teaching and learning objectives and assessment with student learning (Elton and Johnston 2002). In this trend, lecturers and faculty managers should know that assessment criteria established in the syllabus are not trivial. In order to acquire more objective grading it would be a solution for every teacher to share a subject (one teacher for the theoretical framework and one for the practical part) so that they would need to negotiate final grades (Knight 2002; Yorke 2011). Faculties should agree the same assessment criteria for all subjects in order to fairly measure students' performance.

Due to its implications for students, teachers and universities, this research paper contributes to the debate about using new assessment methodologies that requires further research in the future. Also, as future research lines a questionnaire should be made for teachers and students to auto-evaluate their assessment methodology. As limitations of the paper, further longitudinal research is required to contrast our results and to examine agreement in performance for the assessment procedure in Higher Education.

5. REFERENCES

- BAETEN, M., DOCHY, F. & STRUYVEN, K. (2008). Students' approaches to learning and assessment preferences in a portfolio-based learning environment. *Instructional Science*, 36, 359–374.
- BRIDGES, P., BOURDILLON, B., COLLYMORE, D., COOPER, A., FOX, W., HAINES, C. ET AL. (1999). Discipline related marking behaviour using percentages: a potential cause of inequity in assessment. *Assessment & Evaluation in Higher Education*, 24, 285-300.
- BRIDGES, P., A. COOPER, P. EVANSON, C. HAINES, D. JENKINS, D. SCURRY ET AL. (2002). Course marks high, examination marks low: discuss. *Assessment & Evaluation in Higher Education*, 27, 35-48.
- CANO MONTERO, E.I., CHAMIZO GONZÁLEZ, J. & MARTÍN VECINO, T (2016). Coherencia interna del Grado en Administración y Dirección de Empresas: Resultados de Aprendizaje como herramienta de gobernanza. *Revista de Docencia Universitaria*, 14(1), 321-345
- CAMACHO-MIÑANO, M.M; URQUÍA-GRANDE, E.; RIVERO MENÉNDEZ, M.J. & PASCUAL-EZAMA, D. (2016). Recursos multimedia para el aprendizaje de Contabilidad financiera en los Grados Bilingües, *Educación XX1*, 19 (1), 63-89.
- CAMACHO-MIÑANO, M.M. & DEL CAMPO, C. (2016). Useful interactive teaching tool for learning: clickers in Higher Education, *Journal of Interactive Learning Environments*, 24 (4), 706-723.
- CHAMIZO-GONZALEZ, J.; CANO-MONTERO, E.I.; URQUIA-GRANDE, E. & MUÑOZ-COLOMINA, C.I. (2015). Educational data mining for improving learning

- outcomes in teaching accounting within Higher Education, *International Journal of Information and Learning Technology*, 32 (5), 272-285.
- DAFOUZ, E., CAMACHO, M. & URQUIA, E. (2014). 'Surely they can't do as well': a comparison of business students' academic performance in English-medium and Spanish-as-first-language-medium programmes. *Language and Education*, 28(3), 223-236.
- DALZIEL, J. (1998). Using marks to assess student performance: Some problems and alternatives. *Assessment and Evaluation in Higher Education*, 23, 351-366.
- DEL CAMPO, C.; CANCER, A.; PASCUAL-EZAMA, D. & URQUÍA-GRANDE, E. (2015). EMI vs. Non-EMI: Preliminary Analysis of the Academic Output within the INTE-R-LICA Project. *Procedia - Social and Behavioral Sciences*, 212, 74 – 79.
- DEVITA, G. (2002). Cultural equivalence in the assessment of home and internal business management students: a UK exploratory study. *Studies in Higher Education*, 27, 221-231.
- DOWNS, C. (2006). What should make a final mark for a course? An investigation into the academic performance of first year Bioscience students. *Assessment and Evaluation in Higher Education*, 31, 345-364.
- DURÁN SANTOMIL, P.; MASIDE SANFÍZ, J. M.; CANTORNA AGRA, S. & RODEIRO PAZOS, D. (2013). ¿Es el nuevo sistema de evaluación del EEES realmente diferente del sistema tradicional?: Un análisis empírico del rendimiento académico en una asignatura de contabilidad. *Educade: Revista de educación en contabilidad, finanzas y administración de empresas*, 4, 77-96.
- EUROPEAN COMMUNITIES. (2009). *ECTS Users' Guide*, Luxembourg: Office for Official Publications of the European Communities. http://ec.europa.eu/education/tools/docs/ects-guide_en.pdf. Accessed 31 July 2014.
- ELTON, L. & JOHNSTON, B. (2002). *Assessment in Universities: a critical review of research*. LTSN Generic Centre. <http://eprints.soton.ac.uk/59244/1/59244.pdf> Accessed January 15, 2013.
- FLETCHER, R.B., MEYER, L.H., ANDERSON, H., JOHNSTON, P. & REES, M. (2012). Faculty and Students Conceptions of Assessment in Higher Education. *Higher Education*, 64, 119- 133.
- FRICK, T.W., CHADHA, R., WATSON, C. & ZLATKOVSKA, E. (2010). Improving course evaluations to improve instruction and complex learning in higher education. *Educational Technology Research & Development*, 58, 115-136.
- GAMMIE, E. & MATSON, M. (2007). Group Assessment at Final Degree Level: An Evaluation. *Accounting Education: an international journal*, 16 (2), 185–206.
- GIBBS, G. & SIMPSON, C. (2004-05). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1, 3-30.
- GRASHA, A. (1996). *Teaching with Style. A practical guide to enhancing learning by understanding teaching and learning style*. Pittsburgh, PA: Alliance Publishers.
- HEYWOOD, J. (2000). *Assessment in Higher Education*. London: Jessica Kingsley.
- JAMES, D. & FLEMING, S. (2005). Agreement in student performance in assessment. *Learning and Teaching in Higher Education*, 1, 32-50.
- JACKSON, L., W. MEYER & PARKINSON, J. (2006). A study of the writing tasks and reading assigned to undergraduate science students at a South African University. *English for Specific Purposes*, 25, 260-281.
- KNIGHT, P. T. (2002). Summative assessment in higher education: practices in disarray. *Studies in Higher Education*, 27, 275–286.
- KNIVETON, B. H. (1996). Student perceptions of assessment methods. *Assessment and Evaluation in Higher Education*, 21(3), 229-238.
- LONG, M. (1991). Focus on form: A design feature in language teaching methodology. In, K. de Bot, D. Coste, R. Ginsberg & C. Kramsch (Eds.) *Foreign Language Research in Cross-Cultural Perspective* (pp. 39-52). Amsterdam: John Benjamins.

- LÓPEZ-PÉREZ, M.V., PÉREZ-LÓPEZ, M.C., RODRÍGUEZ-ARIZA, L. & ARGENTE-LINARES, E. (2013). The influence of the use of technology on student outcomes in a blended learning context. *Educational Technology Research & Development*, 61, 625-638.
- MARTON, F. AND SÄLJÖ, R. (1976a). On qualitative differences in learning I, Outcome and process. *British Journal of Educational Psychology*, 46, 4-11.
- MARTON, F. AND SÄLJÖ, R. (1976b). On qualitative differences in learning II, Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology*, 46, 115-127.
- MULDOON, R. (2012). Is it time to ditch the traditional university exam? *Higher Education Research & Development*, 31, 263-265.
- MURDAN, S. (2005). Exploring relationships between coursework and examination marks: a study from one school of pharmacy. *Pharmacy Education*, 5, 97-104.
- O'MALLEY, J. M. & CHAMOT, A. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- PASCUAL-EZAMA, D., CAMACHO-MIÑANO, M.M., URQUIA GRANDE, E. & MÜLLER, A. (2011). ¿Son los nuevos criterios de evaluación en el marco del EEES adecuados para valorar el rendimiento académico de los alumnos? Experiencia en Contabilidad Financiera. [Are the new assessment criteria appropriate to evaluate students' academic achievement? An experience from Financial Accounting]. *Revista Educade*, 2, 67-83.
- PAYNE, E. & BROWN, G. (2011). Communication and practice with examination criteria. Does this influence performance in examinations? *Assessment and Evaluation in Higher Education*, 36, 619-626.
- PEACOCK, M. (2000). Learning Styles and Teaching Style Preferences in EFL. *Perspectives*, 12/2000 spring
- PICA, T. (2000). Tradition and transition in English language teaching methodology. *System*, 28, 1-18.
- RIVERO-MENÉNDEZ, M.J.; URQUÍA-GRANDE, E.; LÓPEZ SÁNCHEZ, P. AND CAMACHO-MIÑANO, M.M. (forthcoming 2017). Motivation and learning strategies in Accounting: Are there differences in EMI versus non-EMI students? *Spanish Accounting Review*.
- RITTER, L. (2000). The quest for an effective form of assessment: the evolution and evaluation of a controlled assessment procedure (CAP). *Assessment and Evaluation in Higher Education*, 25, 307-320.
- SASANGUIE, D., ELEN, J., CLAREBOUT, G., VAN DEN NOORTGATE, W., VANDENABEELE, J. & DE FRAINE, B. (2011). Disentangling instructional roles: the case of teaching and summative assessment. *Studies in Higher Education*, 36 (8), 897-910.
- SEGERS, M. AND DOCHY, F. (2006). Enhancing student learning through assessment: Alignment between levels of assessment and different effects on learning. *Studies in Educational Evaluation*, 32(3), 171-179.
- SIMONITE, V. (2003). The impact of coursework on degree classifications and the performance of individual students. *Assessment and Evaluation in Higher Education*, 28, 459-470.
- SMITH, C. (2011). Examinations and the ESL student – more evidence of particular disadvantages. *Assessment and Evaluation in Higher Education*, 36, 13-25.
- SPANISH ORGANIC ACT ON DATA PROTECTION. (1999). Boletín Oficial del Estado (BOE), 298, 43088-43099 (also available at <http://www.boe.es/boe/dias/1999/12/14/pdfs/A43088-43099.pdf>, in Spanish)
- STARR, J.W. (1970). Student opinion on methods of assessment. *Educational Review*, 22, 243-253.
- STRUYVEN, K, DOCHI, F., JANSSENS, S., SCHELFHOUT, W. & GIELEN, S. (2006). The overall effects of end-of-course assessment on student performance: a comparison

- between multiple choice testing, peer assessment, case-based assessment and portfolio assessment. *Studies in Educational Evaluation*, 32, 202-222.
- SWAIN, M. (1985). Communicative Competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass and C. Madden (eds.) *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- TIAN, X. (2007). Do assessment methods matter? A sensitivity test. *Assessment and Evaluation in Higher Education*, 32, 387-401.
- TYNJAL, P. (1998). Traditional studying for examination versus constructivist learning tasks: do learning outcomes differ? *Studies in Higher Education*, 23, 173-190.
- WOODFIELD, R., EARL-NOVELL, S. & SOLOMON, L. (2005). Gender and mode of assessment at university: should we assume female students are better suited to coursework and males to unseen examinations? *Assessment and Evaluation in Higher Education*, 30, 35-50.
- YORKE, M. (2011). Summative assessment: dealing with the 'measurement fallacy'. *Studies in Higher Education*, 36, 251-273.
- YORKE, M., BARNETT, G., BRIDGES, P., EVANSON, P., HAINES, C., JENKINS, D. ET AL. (2002). Does Grading Method Influence Honours Degree Classification? *Assessment and Evaluation in Higher Education*, 27, 269-279.
- YORKE, M., BRIDGES, P. & WOOLF, H. (2000). Mark distributions and marking practices in UK higher education. Some challenging issues. *Learning in Higher Education*, 1, 7-27.
- YORKE, M., COOPER, A. & FOX, W. (1996). Module mark distributions in eight subject areas and some issues they raise. In N. Jackson (Ed.), *Modular higher education in the UK in focus* (pp. 105-107). London: Higher Education Quality Council.