Assessment of transferable competences in computing

Jorge E. Pérez-Martínez, Gregoria Blanco-Viejo
Technical University of Madrid, Madrid, Spain
{jeperez, gblanco}@eui.upm.es

Edmundo Tovar-Caro, Agueda Arquero Hidaldo
Technical University of Madrid, Madrid, Spain
{etovar, aarquero}@fi.upm.es

Carmen Vizcarro-Guarch, Taísa Borges Grün
Universidad Autónoma de Madrid, Madrid, Spain
carmen.vizcarro@uam.es, taisabg@hotmail.com

***Abstract**: The European Credit Transfer and Accumulation System (ECTS) is the credit system for higher education used in the European Higher Education Area (EHEA), which involves all the countries engaged in the Bologna Process. This paper describes a study which is part of the project of the Bologna Experts Team-Spain and was carried out with the following aims: 1) designing some procedures for the assessment of transferable competences; and 2) testing some basic psychometric features that an assessment device with some consequences for the subjects being evaluated needs to prove. We will focus on the degrees of Computing. The sample of students (20) includes first year students from the Technical University of Madrid. In this paper, we will report some results of data analyses carried out to this moment on reliability and validity of the task designed to measure problem solving.*

Introduction

The assessment of competences or learning outcomes is a key concept in the European Credit Transfer and Accumulation System (ECTS) since credits are awarded when the assessment shows the competences which were aimed to be acquired (European Communities, 2009). ECTS is the credit allocation system for higher education used in the European Higher Education Area (EHEA), which involves all the countries engaged in the Bologna Process, 47 at this point in time. Most Bologna countries have adopted ECTS by law for their higher education systems (European Communities, 2009). In Spain, a decree passed in 2007 (Ministerio de Educación y Ciencia, 2007) establishes the transferable competences which any student with a university degree must have developed; these include understanding basic and gradually more advanced texts, problem solving, looking for, selecting and using information to solve problems or making decisions and the capacity to learn independently.

This paper describes a study which was part of the project of the Bologna Experts Team-Spain (http://www.expertosbet.es/) and was carried out with the main aims of gaining experience in the assessment of learning outcomes, designing some procedures for the assessment of transferable competences and testing some basic psychometric features of the proposed task. It is part of a larger study, but only partial data will be discussed here.

In order to achieve the overall goals of our study, participants from different fields of knowledge (Biology, Psychology, Computing and Economy) were invited to take part in it. These participants came from 5 different universities, so that the use of criteria and standards could be compared. In this paper, we shall focus on the degrees of Computer Engineering, Software Engineering, Computing and Mathematics & Computer Science. A sample of over 200 students from these degrees took part in the general study, but only the data of 20 students taking the task purported to measure problem solving will be reported here. We will only report some results of data analyses carried out to this moment including different forms of reliability and validity of the assessment device. This work is still under

progress and only initial data will be discussed. Finally, we will discuss our experience in the use of these procedures.

Context and background

In the context of higher education, a competence may be understood as *the combination of skills, knowledge, attitudes, values and abilities that underpin effective and/or superior performance in a professional area* (European Communities, 2009). In this way, when we try to assess student performance, we are interested in assessing not only knowledge, as has been the case in traditional education, but also *what the student is able to do (and how)* using this knowledge. By *how*, we understand adhering to disciplinary methodological standards and values. Thus, competence or learning outcomes assessment includes the assessment of knowledge, but is not limited to it. It is normally assessed through complex, representative disciplinary tasks that imply knowledge and are often complemented with students' reflections whereby students justify the decisions they have taken on a theoretical and/or disciplinary base, and take into account their consequences or the values that inform them.

The starting point for this study were the basic and general transferable competences mentioned by the Decree 1393/2007 (Ministry of Education and Science, 2007) which every higher education graduate should have developed. They were selected since they are common to all grades although every discipline is expected to further introduce its own particular colouring and nuances. For this reason, they were considered at the same time to be a good basis for independent work and for making interesting comparisons. This fact gave place to two additional questions. On the one hand, we could learn about the particularities of the assessment of learning outcomes regarding different disciplines; on the other, if the structure used for the tasks was similar, we could explore to which extent assessment criteria and standards were used in similar ways.

Research questions

As mentioned above, our aim was to design assessment procedures to assess the basic competences which all graduates must have mastered by the end of their undergraduate academic life according to the Spanish law. This should be complemented by the development of assessment criteria that would allow enough objectivity when correcting and eventually grading students' work. We also tried to validate the tasks as appropriated for the assessment of these basic competences in different ways.

Theoretical framework

In line with the standard procedures described in the literature (e.g., Wass, Van der Vleuten, Shatzer & Jones, 2001), the design of these assessment procedures has included a number of steps: 1) a detailed analysis of the facets included in each competence; 2) designing assessment tasks covering these competences and aspects and including different levels of difficulty to adapt to the development of the student along the two first academic years. We have developed general tasks, which allow us to compare students across the different disciplines, as well as specific tasks for each discipline participating in the study; 3) developing and discussing assessment criteria which allow reliability when grading students' work; 4) determining the basic psychometric properties that any measurement device should show, such as inter-rater agreement, internal consistency and validity (content as well as discriminant validity), following the Standards for educational and psychological testing (AERA, APA, NCME, 1999). In this paper, we will report, by way of example, on the work we carried out regarding one on these general competencies, namely problem solving. This is part of a larger study and although the original study included other aims, no other analyses will be reported here.

Method

Participants

In the larger study, one degree was selected from each of the five branches of knowledge (Arts and Humanities, Social Sciences, Engineering, Health Sciences and Sciences, although no participants from the branch of Art and Humanities volunteered to take part in the study. We also tried to have at least two higher education institutions for each degree in order to make significant comparisons. The Technical University of Madrid participates with two centres and four different degrees in Computer

Sciences: Software Engineering, Computer Engineering, Computing and Mathematics & Computer Science. In order to adhere to limitations imposed for this paper, only the data secured in the two former degrees will be reported here. Table I describes the sample participating in this study.

TABLE I. STUDENTS TAKING PART IN THE PILOT PROJECT

Degree	Men	Women	Total
SE	12	4	16
CE	2	2	4

SE: Software Engineering; CE: Computer Engineering

Participants were between 17 and 24 years of age (mean = 18.5). During the selection period, a total of 14 students had chosen their degree as their first option and 6 as their third. At high school, 19 students studied Science and Technology and 1 studied Humanities. Regarding the country of origin, 18 of the students were Spanish (13 of them came from Madrid) and 2 were Ecuadorian

Procedure

Students participated on a voluntary basis and their work was recognised by 1 ECTS, taking into account they not only took this task but also all the general as well as the specific tasks for each competence.

Measurements and instruments

This section includes, by way of example, the description of one of the competences we have worked with, namely problem solving, our analysis of the facets it implies, the task designed to measure it and of the assessment criteria developed with to aims; on the one hand, making grading more objective and, on the other, breaking down the task so specific feedback can be given to students when the task is used for competence development.

Competence description

An approximate translation of the competence states that students are expected to able to apply knowledge to their career or vocation in a professional way and show the competences acquired which are normally showed when developing and presenting arguments or when solving problems related to their field of study.

Analysis of the facets

Our analysis of this competence is as follows:

1. Presentation and argumentation of a point of view or opinion
   a) Understanding two or more implied sides or opinions on a given issue
   b) Identifying conflict points
   c) Identifying the information needed to support an argument
   d) Organizing the information required to elaborate an argument
   e) Presenting the information by using the adequate format
2. Problem solving
   a) Identifying problems
   b) Developing solution strategies
   c) Determining the information required
   d) Application of knowledge to solve the problem
   e) Evaluating the solution

Since the students undertaking this task are first year students of any computing degree, a subject which is taught in the first semester of all these degrees has been chosen. This subject was Programming.

Task designed

The task (CB2) created to measure this competence is show below.

"During a football championship, the number of goals scored by a certain footballer in the different matches was counted in order to carry out statistical monitoring of his performance. The number of

goals per match, which was either a positive or null whole number, was presented in a computer system. This consisted of a programme with a repetitive structure which recognized the number of goals and showed the amount of them scored by this footballer. Some examples of the way to input the information and the results obtained are given below:

Input:	Output:
3 5 0 4 0 -1	Number of goals: 12
Input:	Output:
0 -1	Number of goals: 0
Input:	Output:
-1	Number of goals: 0

The value -1 means the end of the data input, it is not a data itself and it must always appear in the input".

The following tasks regarding the text above mentioned were performed so as to meet as many aspects as possible:

1. Determine which loop (counter, sentinel) would you use. Explain and justify your choice.
2. Which condition ends the loop?
3. How should the variable associated to this condition be started and ended?
4. Which is the repeating process?
5. How should this process be started and updated?
6. The input of information has been modified and it identifies now the number of goals scored in each match, taking into account that a total of 5 matches have been played. For example:

Input:	Output:
3 5 0 4 0	Number of goals: 12

Question a. Now, what kind of loop would you choose (counter, sentinel…) and why?

Question b. Why have you not choose the alternative loop?

Assessment criteria

Regarding the assessment criteria, we decided that questions 1, 6a and 6b should be graded using the values (0, 2, 4): 0 for inadequate answers, 4 for excellent answers and 2 for intermediate, whereas the values (0, 1, 2), with the same meanings stated above, should be used in the rest. Questions 1, 6a and 6b have a stronger weitgh since they involve a deeper knowledge of the subject which is being assessed. Knowledge in control structures is required for this task but in a natural language taking into account the students' entrance level knowledge. The test was carried out during the third week of October, when students were supposed to have acquired the required knowledge.

Data analysis and results

Data regarding content validity will be presented in a separate report. Briefly, the task was sent to 2 independent experts asking them several questions such as: 1. which one of the basic competences prescribed by the ministry would you think this task measures? 2. Together with a listing of our facets analysis, which of the following facets do you think this competence includes? Please tell us which ones you think are not included in this task or which facets you think are lacking in our analysis. 3. The same as above for assessment criteria.

The results showed the agreement for question 1 was almost perfect and when it was not, we could easily accept the expert's suggestions so they enriched our task analysis. For questions 2 & 3 agreement was above 75 % and, again, many of their comments could be used to improve our analysis. So we may conclude the expert's judgments validate our tasks.

Two research interns were chosen in order to grade computing specific tasks in the University School of Computer Science. They were previously trained to understand and use the grading criteria. With this end, they discussed the criteria with the professors responsible for the tasks. However, given the size of the sample (20 students) we decided that both interns should make corrections together following the assessment criteria previously established; if their evaluation did not coincide, they were asked to discuss their arguments in order to find the conflict points and reach an agreement.

Tables II to V describe some of the analyses carried out to gather information on some basic psychometric properties of the tasks. Table II shows a descriptive analysis of the data as well as an estimation of its difficulty level. As can be seen, the mean values obtained are quite high while the standard deviation is higher for questions 1 and 6b. The difficulty of the items (p-value) is medium for items 3, 4 & 5 and low for the rest. It must be born in mind that these results have been obtained with a small sample of students who volunteered.

TABLE II. DESCRIPTIVE VALUES FOR THE TASK

Questions	Minimum value	Maximum value	Mean	S.D.	*p-value*
1	0	4	3,10	1,51	0,77
2	0	2	1,80	0,61	0,90
3	0	2	1,35	0,81	0,67
4	0	2	1,40	0,75	0,70
5	0	2	1,30	0,73	0,65
6a	2	4	3,70	0,73	0,92
6b	0	4	2,90	1,51	0,72
Total	10	20	15,55	3,05	

As for the internal consistency of the task, Table III shows the corrected internal consistency index as well as Cronbach's alpha. As can be gathered from these results, some of the tasks suggested do not behave in an adequate way to measure the competence concerned. However, this gives us a valuable information on the items that should be removed (items 5 and 6b) to increase reliability.

TABLE III. CICI Y ALFA DE CRONBACH

Questions	CICI	Cronbach' α
1	0,187	0,327
2	0,533	0,142
3	0,220	0,304
4	0,494	0,116
5	-0,079	0,473
6a	0,102	0,370
6b	-0,133	0,506

Finally, Table IV shows the correlations of the overall score in this task with 3 different criteria. Of course these criteria should be considered very approximate, since they do not intend to measure competences. Instead, they represent 2 traditional gradings of the students. However, we did not have an independent measure of competences and expected at least some correlation with these measures. We also added an overall teacher's judgment on student's competence which was carried out by their current teachers by the end of the term.

TABLE IV. INTRACLASS CORRELATIONS OF THE ÍTEMS WITH VARIOUS GRADES (UNIVERSITY ENTRANCE, MEAN GRADE IN CURRENT TERM AND TEACHERS' OVERALL JUDGMENT

Questions	University entrance	Mean grade for current term	Teachers judgments
Total_CB1	-0,288	-0,007	0,425
Total_CB2	0,197	0,035	0,052
Total_CB3	-0,186	0,158	0,471

Total_CB4	0,182	0,471	0,128
Total_CB5	0,227	0,606**	0,673**

As can be seen, results show a significant correlation only for total task score and mean grade and teacher judgment at current term. This could be expected since the total score is a stronger measure than any individual item it contains. On the other hand, the size of the correlation is medium, as expected.

Discussion

We must say all teachers involved were satisfied with the experience. We have a task consisting of a meaningful activity for these students which apparently taps one of the basic competences they should develop. However, some comments seem in order. In the first place, some teachers opposed to working with this task in the context of their classrooms because they understood this was an intolerable a loss of time. Taking into account competence based education needs to introduce competence assessment for development, as well as certification, this seems to show some teachers' mentality has not undergone the needed changes to adapt to the educational reform we are experiencing yet. Of course we understand these tasks should be included in the regular work of students in the different subject matters they have. Otherwise, this is an add on that completely lacks meaning and students should not be expected to be motived to work on these tasks.

Following the thread of this experience, there is a problem which has been raised at Spanish universities: the evolution from the assessment of contents to the assessment of competences, which is sometimes misinterpreted, as it is thought that competences must be assessed on the one hand and contents on the other, separately. Carrying out these tasks has revealed that this is not the case and assessments should not be separated. Learning outcomes should be assessed simultaneously with the contents. This is the answer to usual questions among professors, such as the following: how should I assess competences? When assessing competences, how do I know if students have learn what they "should" have learned? The answer is, of course, that the tasks used to measure competences should imply knowledge.

For instance, only two of the tasks (CB2, CB4) designed for Computer Science could be applied to some subjects included in the curriculum (Programming). Even so, carrying out these tasks required an agreement among the professors who teach the subject concerned and the agreement of their department, as a considerable amount of time is required in order to carry out these. It is worth highlighting this issue, which poses a logistical problem whose solution is generally difficult to find, since the assessment of transferable competences has not been seriously considered in these subjects.

References

American Educational Research Association, American Psychoogical Association & National Council on Measurement in Education (1999). Standards for educational and psychological testing. Washington, DC: American Psychological Association.

European Communities (2009). *ECTS Users' Guide*. Accessed at http://ec.europa.eu/education/lifelong-learning-policy/doc/ects/guide_en.pdf on 16 Jun 2011.

Ministerio de Educación y Ciencia (2007). "REAL DECRETO 1393/2007, de 29 de octubre, por el que se establece la ordenación de las enseñanzas universitarias oficiales", BOE núm. 260, de 30 octubre 2007, pp. 44037-44048

Wass, V., Van der Vleuten, C., Shatzer, J. & Jones, R. (2001). Assessment of clinical competence. *The Lancet*, 359, 945-949

Acknowledgements

This paper has been partially funded by the projects "Life Long Learning Program, DG Education & Culture, Team of Bologna Experts, Spain" as well as by the project IE106110103 from the Technical University of Madrid. We would like to thank Rafael Miñano, Alberto Tejero, Celia Fernández and Ana Palomar for their collaboration in the design of some tasks.

Copyright statement