Cooperative learning of requirements engineering through an international educational scenario enabled by the MOY programme

Autor/res/ras: Juan Manuel Carrillo de Gea[1], Joaquín Nicolás[1], José Luis Fernández Alemán[1], Ambrosio Toval[1], Sofia Ouhbi[2], Ali Idri[2]

Institución u Organismo al que pertenecen:

[1]Software Engineering Research Group, Regional Campus of International Excellence "Campus Mare Nostrum", University of Murcia, Murcia

[2]Software Project Management Research Team, ENSIAS, Mohammed V Souissi University, Rabat, Morocco

Indique uno o varios de los seis temas de Interés: (Marque con una {x})

{x} Enseñanza bilingüe e internacionalización

{ } Movilidad, equipos colaborativos y sistemas de coordinación

{x} Experiencias de innovación apoyadas en el uso de TIC. Nuevos escenarios tecnológicos para la enseñanza y el aprendizaje.

{ } Nuevos modelos de enseñanza y metodologías innovadoras. Experiencias de aprendizaje flexible. Acción tutorial.

{ } Organización escolar. Atención a la diversidad.

{ } Políticas educativas y reformas en enseñanza superior. Sistemas de evaluación. Calidad y docencia.

Idioma en el que se va a realizar la defensa: (Marque con una {x})

{x} Español **{ }** Inglés

Abstract.

The International Excellence Campus for Higher Education and Research of the Region of Murcia, and the Mediterranean Office for Youth (MOY) programme are new initiatives that offer opportunities for designing educational activities in which can take part international students enrolled in academic degrees at different universities. Besides, a significant rise in distributed and collaborative software development has been observed in recent years (Global Software Development, GSD), which involves space, time and socio-cultural distances and requires new techniques, tools and practices to meet new challenges and opportunities. In addition, poor requirements are one of the most common causes of project failure in any domain. Projects which devote more resources to Requirements Engineering (RE) result in lower costs and lower deviations of their planning. Therefore, the relevance of education and training the future systems and software professionals in RE activities and techniques, in particular in GSD environments, must be stressed. We have conducted an educational innovation activity based on teaching RE in co-located and GSD contexts. This activity has been carried out in the form of an experiment with students. This paper presents the scenario in which this educational activity is framed as well as some preliminary results of this experiment.

Keywords: education, experiment, requirements engineering

1. Introduction

Campus Mare Nostrum 37/38 is the International Excellence Campus for Higher Education and Research of the University of Murcia (UMU) and the University of Cartagena (UPCT). The campus is a joint effort of international organizations, research centers, technology parks, companies and the administration, which seeks to transform the Region of Murcia into a pole of international, high-quality education, science, business and culture in the Mediterranean area (CMN, 2014). This offers the opportunity of designing educational activities in which can take part international students enrolled in academic degrees at different universities.

The Mediterranean Office for Youth (MOY) is a multilateral pilot program that facilitates mobility of students enrolled in Master and PhD, in the Mediterranean countries. The MOY programme's aim is to promote exchanges and mutual knowledge and to contribute to economic and social development of its sixteen member countries. In this perspective, it aims to:

- Develop academic exchanges in the priority sectors of development of the partner countries consistent with their labour markets;
- Facilitate the mobility of the best Master/PhD students from the Mediterranean area through the certification of educational excellence resulting in co-diplomas and a system of mobility grants;
- Promote a first professional experience abroad, with the creation of a platform for internships and jobs.

The MOY programme (MOY, 2014) has enabled an academic relationship between the University of Murcia (UMU) and the Mohammed V Souissi University (UMV). Specifically, contacts and collaboration between the UMU's Faculty of Computer Science and the UMV's ENSIAS has been established.

A significant rise in distributed and collaborative software development has been observed in recent years (López, Nicolás, & Toval, 2009; Lima Peixoto, Nicolas Audy, & Prikladnicki, 2010). As a result, the software industry is now truly global. The diversity of cultures and the dispersion in time and space involved in outsourced software development projects require new techniques, tools and practices from various disciplines to meet the challenges and opportunities offered by Global Software Development (GSD) (Damian, Lanubile, & Oppenheimer, 2003).

Poor requirements are one of the most common causes of project failure in any domain (The Standish Group, 2009). A typical estimation for a regular project is to devote around 10% of its effort to requirements. However, the most successful software projects in banking and telecommunications among 15 projects analysed, were those who allocated more than 28% of their resources to requirements (Hofmann & Lehner, 2001). Furthermore, a study of National Aeronautics and Space Administration (NASA) projects corroborates that those which devoted more than 10% of their resources to RE resulted in lower costs and lower deviations of their planning (Hooks & Farry, 2001). Thus, it is necessary to stress the relevance of education and training the future systems and software professionals in RE activities and techniques.

Requirements are generally specified in industry using unstructured natural language (NL) (D. Ott, 2012), since it is easier to understand by all the stakeholders than other non-textual notations, even when they lack any kind of technical training. However, NL is inherently ambiguous and can lead to different interpretations depending on the context (Mavin, 2012). This drawback will be therefore exacerbated when globally distributed stakeholders are involved in the process, as regards language, social and cultural differences, in addition to problems resulting from tacit knowledge related to requirements (Noll, Beecham, & Richardson, 2010). Nevertheless, this is not an impediment for systems and software development to become more and more globalised nowadays, as shown by the growth of GSD in recent years.

In the above context, we have posed an educational innovation activity based on teaching RE in a co-located and GSD context. Moreover, this activity has been carried out in the form of an experiment with students. This paper presents the scenario in which this educational activity is framed as well as some preliminary results of this experiment, which provide a basis for analysing the work of the co-localized and distributed groups.

2. Methodology

The students were provided with the following documentation:

- Task statement.
- Specific assignment of each team.
- Internationalisation (i18n) standards (ISO/IEC 24751, ISO 9241-151, W3C and CWA 14928X).
- Generic i18n requirements catalogue in the form of a software requirements specification (SRS).
- Traceability matrix including the relationships between the requirements in the catalogue.
- Brief explanation of the foundations of the techniques to be applied.
- Questionnaire.

The tasks to be completed by the participants were concerned with two strategies for training in requirements specification skills, namely specification from source documents and specification from catalogues of reusable requirements. This signified that the participants had to elicit a set of requirements from a specific section of an i18n standard (Task 1.a, non-reuse task) and reuse a set of requirements related to a particular topic using an i18n requirements catalogue (Task 1.b, reuse task). The participants also filled in a questionnaire in which they provided us with the solutions of the tasks (input to measure effectiveness), reported the time they had spent on it (which is needed to calculate productivity) and also their perceptions of the techniques and results (used to measure difficulty, speed, quality and understanding). These indicators helped us to determine the skills and knowledge acquired by the students in a practical, learning by doing scenario.

The assignments set in this study cover different learning objectives. In terms of the cognitive domain of Bloom's Taxonomy (Bloom, Furst, Hill, & Krathwohl, 1956), the correspondence between the cognitive learning levels and some of the educational activities is shown in Table 1.

Category and educational activities
Knowledge. Memorise concepts such as software requirement, specification, quality requirement, traceability, requirements catalogue, and so on
Comprehension. Understand the activities of requirements specification from source documents and from catalogues of reusable requirements. For example, comprehend the requirements reuse process
Application. Use the "requirement" concept in a new situation. For example, collaborate with a distributed team-mate to specify requirements
Analysis. Identify key information of a paragraph in a standard to specify quality software requirements, or apply a simple reuse process to achieve the same goal. For example, discuss and agree with a team-mate on the meaning of an statement in a source document
Synthesis. Build new requirements using two requirements specification techniques. For example, specify new requirements starting from a standard
Evaluation. Choose a suitable NL statement for specifying a requirement, or search the reusable requirements catalogue and resolve the traceability relationships and variation points found in the reusable requirements. For example, avoid ambiguity, incompleteness and inconsistency

Table 1. Educational activities in the cognitive domain of Bloom's taxonomy

3. Experiment

This section gives detailed information on the experiment designed and conducted during the fall term of 2012.

3.1. Participants

A total of 31 computer science and engineering students either in their last years of university or in their first years of postgraduate studies with similar training and experience in RE and mixed gender, age and educational background gave their verbal consent to participate in the study. They had a good command of technical English in order to be able to properly follow the instructions and fit in with the international focus of the experiment. A formal sample size calculation was not performed due to a lack of information from previous studies. However, efforts were made to recruit the largest possible sample.

3.2. Design

The experimental tasks were carried out by teams of two people. There were three participation modalities: (1) global (GLO), one student from the University of Murcia (UMU) was paired up with one student from the Mohammed V Souissi University

(UMV); (2) co-located in Murcia (CLM), two students from the UMU were paired up; and (3) co-located in Rabat (CLR), two students from the UMV were paired up. The working language was English, although informal communication in the native language was allowed for participants in the co-located modalities. Nevertheless, results had to be presented in English, in both the global and co-located modalities.

The study employed a randomised controlled design. After recruitment, 31 students were randomly divided into groups. There were 15 participants from the UMV and 16 participants from the UMU. We defined 14 teams of two students and one team of three students, since the number of participants was odd. Seven teams (15 students) were assigned to the co-located modality (CLM or CLR), while eight teams (16 students) were assigned to the global modality (GLO). Moreover, four teams (eight students) were appointed to CLM groups and three teams (seven students) were appointed to CLR groups.

The period for the execution of the experiment was two weeks. It started on November 26th, 2012, and this coincided with the date on which the task statement and the rest of the documentation were handed out, and the procedure of the experiment was explained to the participants. The estimated effective time that they had to devote to completing the task was in the order of hours. However, during the whole two weeks period, the participants needed to meet for the first time and jointly agree on face-to-face or virtual work meetings and schedule them accordingly, depending on personal time constraints.

The participants were encouraged to interact, collaborate and work together to complete the tasks, thus fostering a climate of mutual support and guidance. They were also encouraged to discuss and agree with their team-mates on each aspect of the tasks (i.e. agree on meetings, comprehension of the bases of the techniques, problem solving activities, etc.). The participants included in a co-located team (CLM or CLR) had to set up at least one face-to-face meeting with their team-mates. They were also asked to strictly adhere to our instructions and recommendations, since this was key to the correctness of the experimental procedure.

3.3. Outcome Instruments and Measures

A total of six variables were studied, and a Likert-type scale was used in five of them: *effectiveness, productivity, difficulty, speed, quality and understanding*. We can classify our variables into one of the following two types (Abrahão, Insfran, Carsí, & Genero, 2011): performance-based variables, which measure how well subjects are able to use a requirements specification method (*effectiveness, productivity*); and perception-based variables, which measure how effective subjects believe a requirements specification method is (*difficulty, speed, quality, understanding*).

The researchers attained effectiveness by objectively assessing the outcome of the task execution (1-Very good, 2-Good, 3-Fair, 4-Poor). This evaluation was carried out by the first author of this paper, since the existence of different raters can result in disagreements about measurement results of the same object (Fleiss, Levin, & Paik, 2003). Deviations and inconsistencies (e.g. variations in the procedures used to carry out the experiment, interpreting the results and presenting them) which may be affected by experimenter's bias, were therefore mitigated. Productivity is defined as output divided by the effort required to produce that output (Maxwell & Forselius, 2000). However, the notion of output is not straightforward for software (Premraj, Shepperd, Kitchenham, & Forselius, 2005). In our analysis the measure collected is *requirements per hour* (Seyff, Graf, Maiden, & Grünbacher, 2009), which reflects the amount of requirements specified per hour per team, i.e. Productivity = requirements/hour. The remaining variables represent the participants' subjective perceptions, which were gathered using a 5-point Likert scale. Difficulty is a measure of how hard the techniques are (1-Easy, 5-Difficult). Speed assesses the rapidness of the techniques (1-Quick, 5-Slow). Quality reflects the perceived quality of the requirements obtained by using the techniques (1-High, 5-Low). Finally, understanding is a measure of the comprehension of the resulting requirements after using the techniques (1-Good, 5-Bad). Productivity, difficulty and speed are therefore focused on the process, and effectiveness, quality and understanding provide insights into the product. Each variable is analysed separately for each modality of participation, namely co-located and global.

3.4. Results

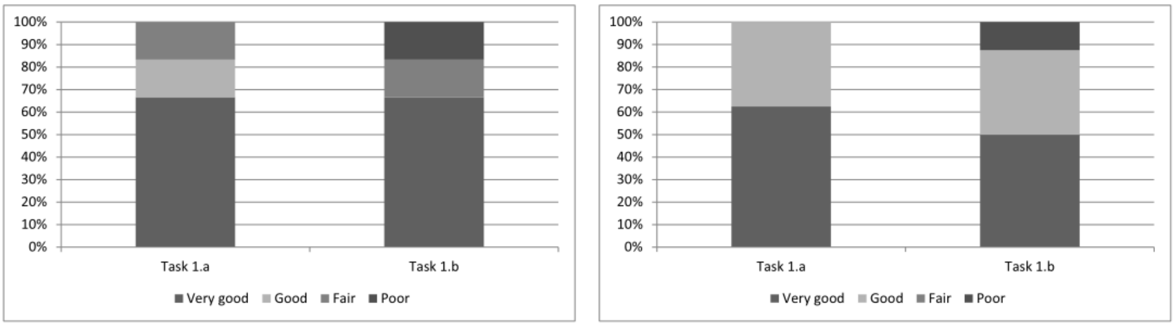
The 31 students who initially participated in the study had diverse ages, gender, educational backgrounds, and grade averages in the previous academic year. Most participants were male (61.3%, n = 19). Participants were aged between 21 and 25 years, with most belonging to the 22-year age group. The final participation rate of the experiment was 93.55% (29 out of 31 participants), given that one CLR team made up of two students was discovered to be an outlier in the sample and was eventually discarded.

A summary of the statistics is presented in Table 2 in order to describe the variables of the study. Measures of central tendency (arithmetic mean) and dispersion (standard deviation) are included. Figures 1-6 depict our variables from a descriptive point of view. We use two types of graphs. On the one hand, the stacked bar graph is an extension of the bar graph which provides us with a method with which to display data from a pair of qualitative variables or a single quantitative and a qualitative variable. The boxplot, meanwhile, is a graphic display method that is concerned with the data's symmetry and skewness, and provides us with numeric measures of central tendency, location and spread of the variables (L. Ott & Longnecker, 2010).

Task	**Non-reuse**						**Reuse**					
Modality	Co-located			Global			Co-located			Global		
Statistic	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Effectiveness	13	1.46	0.78	16	1.38	0.50	13	2.00	1.35	16	1.75	1.00
Productivity	13	26.54	23.57	16	10.78	9.18	13	13.14	7.91	16	8.82	4.95
Difficulty	13	1.92	0.76	16	2.13	1.09	13	2.62	1.04	16	2.56	1.21
Speed	13	2.08	0.64	16	2.38	1.31	13	2.62	1.19	16	2.44	1.21
Quality	13	2.38	1.12	16	2.25	1.13	13	2.54	0.52	16	2.63	0.89
Understanding	13	1.92	0.95	16	2.00	1.15	13	2.15	0.55	16	2.13	0.72

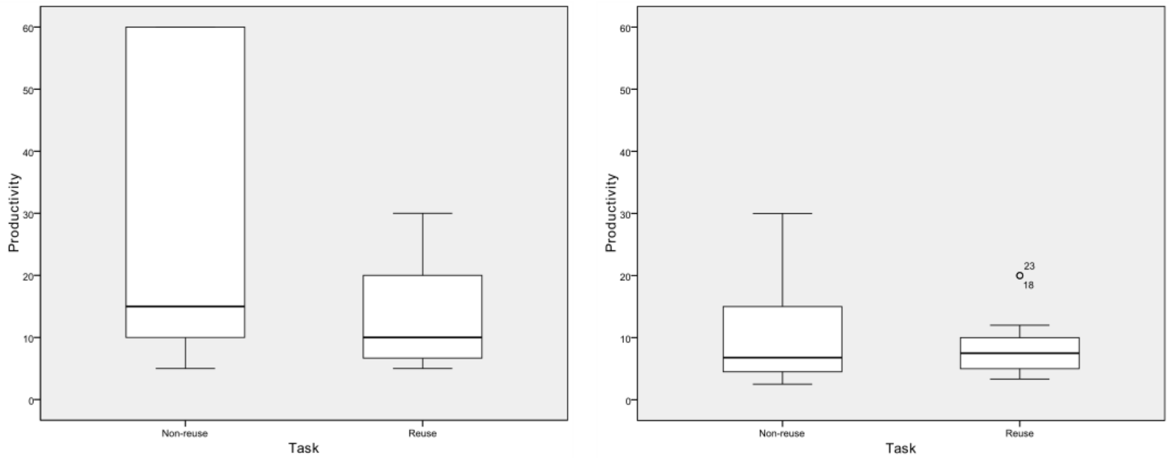
Table 2. Descriptive statistics. N: sample size, SD: std. deviation

Stacked bar graphs can be used to show the percentage different sub-groups contribute to each separate category. For example, in the case of Figure 3 (a), the bars representing the individual categories (i.e. Task 1.a and Task 1.b) are all the same size —which corresponds to the value of 100%— and the relative contribution of the sub-groups (i.e. 1-Easy, 2-Somewhat easy, 3-Neutral, 4-Somewhat difficult, 5-Difficult) is different for each category. Thus, in the case of the Task 1.a category, 2-Somewhat easy has the highest importance —more than 45%—, followed by 1-Easy —more than 30%— and 3-Neutral —more than 20%—. 4-Somewhat difficult and 5-Difficult have no contribution, or 0%. The same applies to the interpretation of the other stacked bar graphs. Furthermore, the boxplot was only used to represent the variable productivity, given its quantitative nature.



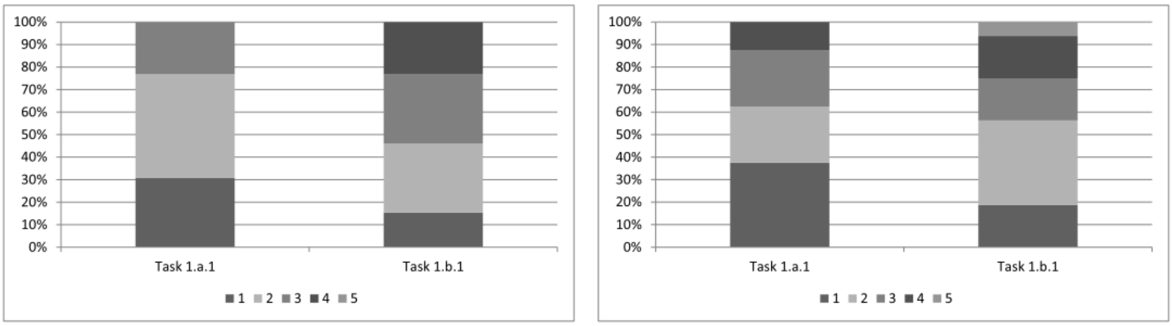
(a) (b)

Figure 1. Effectiveness (1-Very good, 4-Poor): (a) co-located students and (b) global students



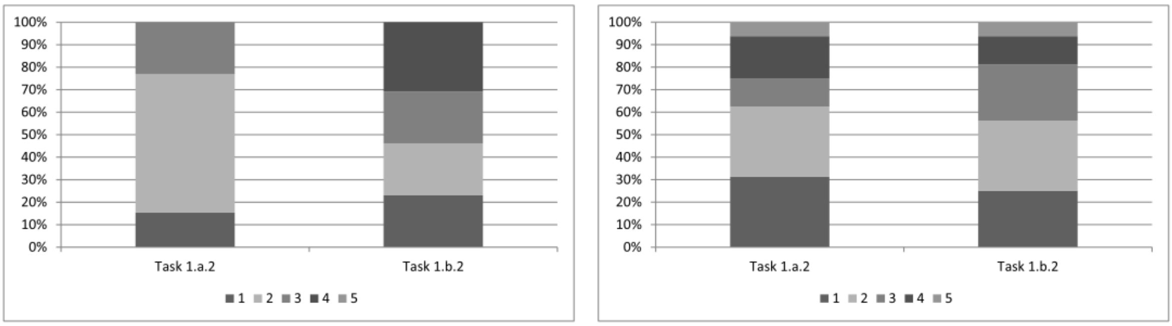
(a) (b)

Figure 2. Productivity (requirements per hour per team): (a) co-located students and (b) global students



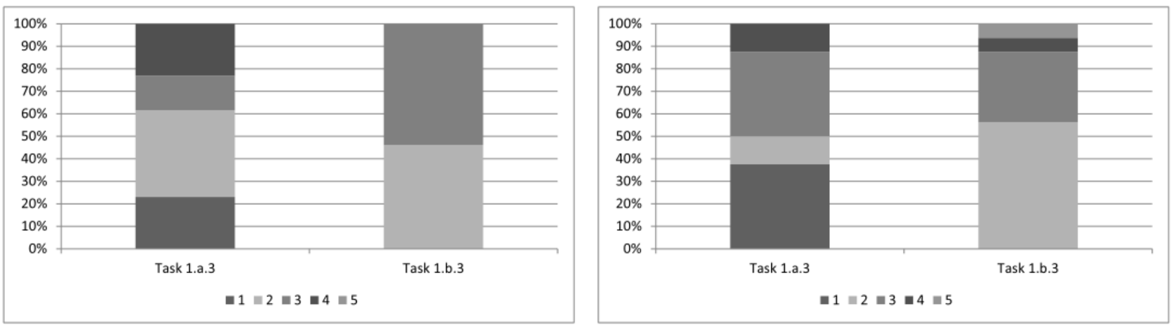
(a) (b)

Figure 3. Difficulty (1-Easy, 5-Difficult): (a) co-located students and (b) global students



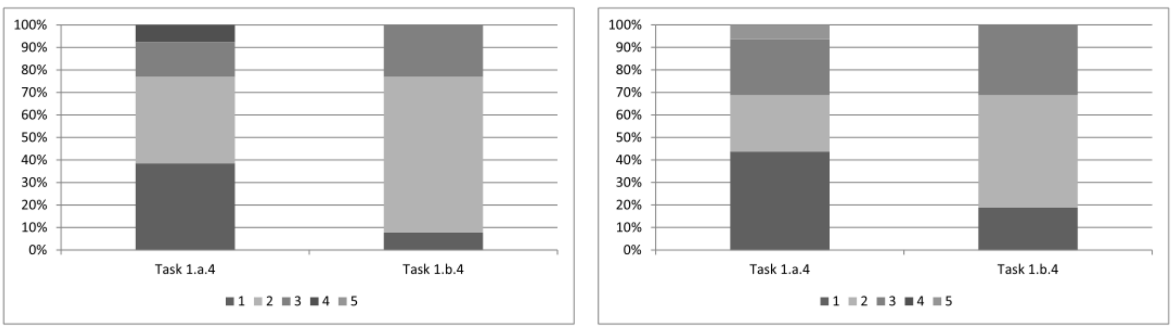
(a) (b)

Figure 4. Speed (1-Quick, 5-Slow): (a) co-located students and (b) global students



(a) (b)

Figure 5. Quality (1-High, 5-Low): (a) co-located students and (b) global students



(a) (b)

Figure 6. Understanding (1-Good, 5-Bad): (a) co-located students and (b) global students

4. Conclusions and future work

In order to summarise the descriptive information presented above, we can state that the global teams performed slightly better than the co-located teams in both the non-reuse and reuse tasks, according to the effectiveness measure. This can be explained by the fact that the productivity was clearly lower in the case of the global teams. More time devoted to requirements normally implies better results (Hofmann & Lehner, 2001; Hooks & Farry, 2001). Concerning the specification of requirements from source documents, the participants' subjective perception of the difficulty, speed and understanding was a little better in the case of the co-located teams. Perceived quality was, however, better in the case of the global teams. On the other hand, the participants' subjective perception of the task of specifying requirements from the catalogue of reusable requirements was generally better in the case of the global teams, with the exception of quality, which was better perceived by the co-located teams. Nevertheless, the differences between them are very slight, particularly in the case of understanding. In summary, depending on geographical dispersion, we found opposite students' perceptions of the educational approaches, although the differences are mostly small or very small. The statistical dispersion or variability of the scores is higher in the case of difficulty and speed than in the case of quality and understanding for both co-located and global teams.

With regard to the comparison between the outcomes of the educational strategies studied in this paper, our results show that the students assimilated the traditional approach —specification of requirements from source documents— better than the reuse-based approach. This occurred throughout all the variables of the study.

In future work, more detailed statistical data analysis will be performed, including statistical inference techniques, to gain insights into the relevance of the differences between the non-reuse and reuse tasks, and between the co-located and global teams. This will allow us to generalise the results of our experiment and draw relevant conclusions.

References.

Abrahão, S., Insfran, E., Carsí, J. A., & Genero, M. (2011, Aug.). Evaluating requirements modeling methods based on user perceptions: a family of experiments. *Inf. Sci.*, 181(16), 3356–3378.

Bloom, B., Furst, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives: Handbook i, the cognitive domain*. Addison-Wesley.

CMN (2014). *What is CMN?* January 2014. URL: http:// http://www.campusmarenostrum.es/que_es_cmn_en.html

Damian, D., Lanubile, F., & Oppenheimer, H. L. (2003). Addressing the challenges of software industry globalization: the workshop on global software development. In Icse '03: Proc. of the 25th int. conf. on softw. eng. (pp.793–794). Washington, DC, USA: IEEE Comput. Soc. Press.

Fleiss, J. L., Levin, B. A., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed.). J. Wiley.

Hofmann, H. F., & Lehner, F. (2001, Jul). Requirements engineering as a success factor in software projects. *IEEE Softw*., 18(4), 58–66.

Hooks, I. F., & Farry, K. A. (2001). *Customer-centered products: creating successful products through smart requirements management*. New York, USA: AMACOM.

Lima Peixoto, C. E., Nicolas Audy, J. L., & Prikladnicki, R. (2010). Effort estimation in global software development projects: preliminary results from a survey. In *Proc. of the 5th IEEE int. conf. on glob. softw. eng.* (pp. 123–127). Washington, DC, USA: IEEE Computer Society.

López, A., Nicolás, J., & Toval, A. (2009). Risks and safeguards for the requirements engineering process in global software development. In *Proc. of the 4th IEEE int. conf. on glob. softw. eng.* (pp. 394–399). Washington, DC, USA: IEEE Comput. Soc. Press.

Mavin, A. (2012, March). Listen, then use EARS. *IEEE Softw*., 29(2), 17–18.

Maxwell, K. D., & Forselius, P. (2000, Jan). Benchmarking software development productivity. *IEEE Softw.*, 17(1), 80–88.

MOY (2014). *About the MOY*. January 2014. URL: http://www.officemediterraneendelajeunesse.org/en/about-the-moy/about-the-moy

Noll, J., Beecham, S., & Richardson, I. (2010). Global software development and collaboration: barriers and solutions. *ACM Inroads*, 1(3), 66–78.

Ott, D. (2012). Defects in natural language requirement specifications at Mercedes-Benz: an investigation using a combination of legacy data and expert opinion. In *Proc. of the 20th IEEE int. requir. eng. conf.* (pp. 291–296). Los Alamitos, CA, USA: IEEE Comput. Soc. Press.

Ott, L., & Longnecker, M. (2010). An introduction to statistical methods and data analysis. 6th ed., pp. 56–139. Brooks/Cole, Cengage Learning.

Premraj, R., Shepperd, M., Kitchenham, B., & Forselius, P. (2005). An empirical analysis of software productivity over time. In *Proc. of the 11th IEEE int. symp. on softw. metr.* (pp. 37–46). Washington, DC, USA: IEEE Comput. Soc. Press.

Seyff, N., Graf, F., Maiden, N., & Grünbacher, P. (2009). Scenarios in the wild: experiences with a contextual requirements discovery method. In M. Glinz & P. Heymans (Eds.), *Requirements engineering: Foundation for software quality* (Vol. 5512, pp. 147–161). Springer Berlin Heidelberg.

The Standish Group (2009). *Chaos summary 2009. the 10 laws of chaos*. The Standish Group International, Inc.