This is a postprint version of the following published document:

Munoz Merino, P. J., Alario Hoyos, C., Munoz Organero, M., Kloos, C. D. & Molina, M. F. (2018). The Effect of Different Features for Educational Computer-Based Competition Environments. *IEEE Transactions on Learning Technologies*, 11(4), pp. 468–477.

DOI: <https://doi.org/10.1109/TLT.2017.2763595>

© 2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

The Effect of Different Features for Educational Computer-based Competition Environments

Pedro J. Muñoz-Merino, Senior Member, IEEE, Carlos Alario-Hoyos, Mario Muñoz-Organero, Carlos Delgado Kloos, Senior Member, IEEE, Manuel Fernandez Molina

Abstract—Educational computer-based competition environments need to be designed with a set of features that enhance the learning process. Although recently some frameworks for the design of educational computer-based systems (e.g., educational games) have been proposed, they do not focus on the details of the competition. Therefore, the design of educational computer-based competition environments is still an open issue. We propose the 4NESS framework for the design of such competition environments. This framework classifies the competition features into four dimensions: easiness, fairness, awareness, and adaptiveness. We designed the ISCARE educational computer-based competition environment implementing these four dimensions to evaluate and validate some included features for competition, and conducted experiments during three consecutive years. Students considered the computer-based competition environment with the proposed features, including a modification of the Swiss system, to be very easy to use, fair for pairing students, moderately fair for calculating scores, with good awareness of their state during the competition and that of their peers, and with contents of appropriate level, especially for the group working with adaptive contents. Adaptation of contents made students increase their learning gains in 0.5 sigmas during the competition. Pairings were seen as with students of similar levels. And final scores were moderately fair: while the order tends to be similar as compared to the results in the post-test, students' grades with high performance tend to be lower in the competition environment.

Index Terms— computer environments, competitive environments, competition, education

1 INTRODUCTION

COMPETITION has traditionally been used in many different areas, such as sports and gaming. The use of competition in education is, however, a controversial topic. There are authors who support that competition only brings bad consequences and should be avoided in the classroom [1]. One of the main criticisms for using competition in the classroom is that it promotes in learners the achievement of one's own goals, even though this may cause negative emotions on classmates [2], such as anxiety for getting good results, or frustration on people who lose [3][4][5]. Moreover, theories of social identity manifest that those participants who are part of a group may give preferential treatment to others when they are perceived to be in the same ingroup, forcing direct competition with those from the outgroup [6][7]. For this reason, computer-based competition environments have been generally underappreciated in education in favor of collaborative environments, where the importance is on the goal achieved by a team of students [8].

However, computer-based competition environments have also reported significant positive effects in many studies, increasing learners' performance [9][10] and

motivation [11][12][13]. Moreover, different features to try to mitigate negative emotions, such as anxiety and frustration, while using computer-based competition environments have been implemented and evaluated with successful results. Examples of these features are the introduction of virtual pets instead of the own students' images [3], the combination of competition with cooperation or collaboration [14][15][16], or the assignment of opponents with a similar knowledge level using the Equal Opportunity Tactic (EOT) approach [17].

The introduction of Intelligent Tutoring Systems (ITSs) for managing and enabling competition in education can help to improve the learning experience and alleviate the negative issues that may derive from the use of computer-based competition environments in the classroom. Actually, there are already tutors in the literature that implement competition in several ways. Some of them are incorporated into existing tools, such as Cool-it ([18], which is integrated in Matlab and compares students' solutions with respect to an ideal one, giving students a level of achievement. There are other tutors that are based on game boards, such as Joyce [19], where students' advancement on the board depends on answering questions correctly. Another example tutor is QUEST [9], which applies collaboration in addition to the competition, giving scores depending on the time to answer compared to the opponents.

Arguably, the use of computer-based competition environments generates a substantial discrepancy in the

• Pedro J. Muñoz-Merino, Carlos Alario-Hoyos, Mario Muñoz-Organero, Carlos Delgado Kloos and Manuel Fernández Molina are with the Department of Telematics Engineering, Universidad Carlos III de Madrid, Av. Universidad 30, 28911 Leganés (Madrid), Spain Email: pedmume@it.uc3m.es, calario@it.uc3m.es, munozm@it.uc3m.es, cdk@it.uc3m.es, manuferna@gmail.com

educational community. And although the effect of computer-based competition environments on students' emotions has been extensively analyzed in the literature, there is still no agreement on the appropriate way to design computer-based competition environments for education. It is therefore necessary to gain insights into the design principles to be followed by environments that support competition in education, as well as into the effect of different features included in these environments. In particular, it is necessary to gain insights into the effect of design principles and different features on a) students' satisfaction; b) students' final rankings; and c) students' learning gains.

In general, the design of educational computer-based competition environments is an open issue. In the last years, several frameworks have been proposed for the design of some educational computer-based systems such as educational games. For example, authors in [20] proposed three different components: motivation, learning and gaming. Similarly, authors in [21] enumerate different learning mechanics. Based on these frameworks, different computer-based educational games have been designed and evaluated, such as shopC [22], or the one proposed in [23]. However, these are frameworks that focus specifically on educational games, and do not enter into details about how the competition should be designed. There is a need for frameworks that help in the design and analysis of computer-based competition environments, such as which type of competition might be better, the fairness, or if the adaptation of contents can work in combination with the competition. Only some aspects of the competition in education have been studied such as the fact that problem solving attitude and creative thinking are the top skills that can be reinforced with competition environments [24].

Therefore, there is a need for studying all the aspects related to the design of computer-based competition environments. This paper gathers design principles and features for educational computer-based competition environments in four main dimensions under the 4NESS framework. These four dimensions are: easiness, fairness, awareness, and adaptiveness. Next, specific implementations of these four dimensions are provided in an example computer-based competition environment, named ISCARE. This environment is used in real educational settings, and serves as the base to analyze and compare the effect of the implementation of these dimensions on computer-based competition environments. This allows us to infer conclusions about the effect of design principles and different features on computer-based competition environments. These features would not be feasible to implement without a computer, since e.g., the proposed modification of the Swiss system requires a computational power. Specifically, the main research questions are the following:

- 1) Is a specific design of a computer-based competition environment easy to use?
- 2) Which type of competition system do students prefer and do they consider fair their scoring and adaptive

matching?

3) Were students aware of their rankings and of the evolution of their opponents in live during their interaction with the competition environment?

4) What is the difference on learning gains when using the competition with adaptation of contents with respect to not adapting the contents?

2 4NESS

4NESS is a design framework that aims to cover four main dimensions that should be present in educational computer-based competition environments in order for students to get a pleasant experience and increase their motivation and learning. These four dimensions come from different fields that have impact on the design of educational computer-based competition environments, and are: easiness, fairness, awareness and adaptiveness (see Figure 1). Easiness comes from the Human-Computer Interaction (HCI) field, as technologies that are easy to use are more likely to be adopted [25], and is included in the framework since the aim is to design computer-based environments with a high acceptance. Fairness comes from the psychology field, as procedural justice is related to learner motivation [26], and is included in the framework since the aim is to design educational competition environments that promote motivation for all learners using them. Awareness comes from the education field, as self-awareness fosters student reflection, which is a mechanism to enhance learning [27], and is included in the framework since the aim is to design effective educational environments. Adaptiveness comes from the ITS field, as it helps to deal with a wide variety of students with different goals by personalizing instruction [28], and is included in the framework since the aim is to design computer-based competition environments that can be tailored to learner's needs.

Easiness refers to the property of simplicity (ease) of use of a computer-based environment in terms of end-user satisfaction with such environment to achieve a certain goal. The ease of use is one of the main aspects studied by HCI, which seeks to improve the usability of computer interfaces and systems in general [25]. Educational computer-based environments, in general, must be easy to use and easy to learn so that learners can incorporate them seamlessly into the suite of tools they use as part of their learning process. The ease of use of a certain technology has an important impact on the cognitive load of the learner, which is the amount of information the learner must process to complete a task using the given technology [29]. The easier the use of the technology, the lower the cognitive load and, in general, the faster the learner can complete the task [30]. Cognitive Load Theory (CLT) aims to decrease extraneous cognitive load [31], so that learners can focus on the primary task, instead of devoting great efforts in trying to understand how the technology they are using works. The design of computer-based competition environments should take best practices learned from the HCI field in the design of computer-based environments, and from

Fig. 1. 4NESS design framework for educational computer-based competition environments.

the CLT field in the design of educational computer-based environments. In the case of educational computer-based competition environments the dimension of easiness can be implemented, for instance, in a clear and attractive graphical user interface that includes only the key functionality students need to compete against each other.

Fairness refers to the property by which students consider that the bases of the competition are established looking for social justice, considering developed skills and acquired knowledge of the individuals with respect to the group of competitors. The fairness in the bases of the competition includes: the fairness in the type of tournaments, the fairness in the pairings, and the fairness in the scoring systems. Regarding the type of tournament, elimination tournaments (knockout) allow ensuring a winner within a few rounds even for a large number of participants, and promote the belief that less skilled participants can end up higher in the ranking [32]. Round-robin tournaments (all-play-all) are typically considered as fairer than elimination tournaments, in the sense that more-able students are more likely to win, although round-robin tournaments become impractical as the number of competitors becomes larger [33]. Swiss-system tournaments are an alternative non-elimination approach aimed at reducing the number of rounds in round-robin tournaments. Regarding pairings, highly unbalanced pairings tend to be seen as unfair, and can cause discouragement and frustration in less-able students [17]. In order to reduce these negative feelings and make fairer pairings, [17] proposed Equal Opportunity Tactic (EOT) for competition in education, achieving a reduction in the effect of individual ability difference on the perceived performance between more-able and less-able students. Finally, regarding the scoring system, fairness can be improved considering the actual performance of each opponent and the type of victory (narrow, overwhelming, etc.), instead of only giving points to the winner.

Awareness refers to the property of the educational

computer-based competition environment by which the student is constantly receiving meaningful information of his own progress and of his progress compared to the peers, as learners often wish to know their work is of the same level as that of their peers. A successful awareness strategy relies on collecting low-level information from the computer-based environment, processing this information, and showing appropriate visualizations to the learner [34]. Awareness is intended to foster three additional stages in the learning process: reflection (the learner asking questions based on the visualized data), sense-making (the learner answering these questions) and impact (the learning changing his behavior, if he considers it is necessary) [34]. The visualization of data is intended to increase learners' self-reflection and is especially useful in self-regulated learning contexts, where the learner is in control of his own learning process [35]. The visualization of data for awareness purposes in the case of educational computer-based competition environments can include, for example, the absolute position of the learner in the leaderboard, and the relative position among peers with similar level of knowledge. Other examples of relevant data for the learner are, for example, the temporary restrictions that apply when competing, e.g., the remaining time for himself (or the opponent) to answer the questions.

Adaptiveness refers to the property by which the educational computer-based competition environment changes its behavior to tailor to each student needs, getting information from the student himself and from the competition he is running; this can include, for example, adapting the questions delivered to competitors to their level of knowledge of the subject. Research on adaptive hypermedia defines different types and techniques of adaptiveness [36][37] which can be applied to computer-based competition environments with the aim to provide intelligence to the learning process and contribute to the improvement of learning effectiveness. This improvement of learning effectiveness through adaptiveness has been extensively demonstrated in the case of ITSs with increases of sigmas in learning within the range from 0.3 to 1.7 (with respect to traditional instruction with no tutoring) [38]. Other interesting comparisons can be applied between different versions of the same ITS or different ITSs, such as tutor strategies to prevent gaming the system [39]. Furthermore, several studies reported on the effectiveness of learning with specific ITSs, such as PACT tutor [40] or AutoTutor [41], measuring the difference between a post-test and a pre-test, but without a comparison with other instructional methodologies or ITSs. There are several evaluations involving competition ITSs, but much more effort should be done for analyzing their effectiveness in terms of learning gains and comparison with different situations. The study in [42] reported a survey on students' own learning perception of computer-based competition environments, obtaining positive results. The work in [9] analyzed students' satisfaction and achievement between experimental group with competition and control group, obtaining better final grades for the competition group,

but without using a pre-test and post-test methodology. Finally, adaptiveness has not only been applied to ITSs but also to Learning Management Systems (LMSs), such as in Moodle, where Artificial Intelligence techniques are applied for adaptation without providing competition features [43], or as in [42].

3 A COMPUTER-BASED COMPETITION ENVIRONMENT EXAMPLE: ISCARE

ISCARE (Information System for Competition based on pRblem solving in Education) [44] is a competition environment developed for educational purposes and includes the four dimensions of the 4NESS framework. Table 1 presents the main features of the ISCAR competition environment and how these relate to the dimensions defined in the 4NESS framework.

Figure 2 shows a screenshot of a student interacting with ISCAR. The left frame presents one by one the questions that the student has to answer. The questions can be multiple choice (several answers and one is correct), multiple response (several answers and several are correct) and fill in the blank. Students can receive feedback about correct or incorrect resolution after answering each question. The questions cannot only be to memorize concepts but they can involve thinking and reasoning about different issues. Students can learn while solving the different questions.

The right frame of Figure 2 shows competitors main information. The top frame shows context information about the tournament, round and number of problems to be solved in the round. According to the screenshot, this is the third round of the tournament and, therefore, each student has already competed against two other opponents in previous rounds. The competition is synchronous for each round, i.e. all students compete at the same moment for each round, so all students should be online at the same established moment.

In each round, each student can obtain two points: one depending on the number of right answers and one depending on the outcome of the competition against the opponent on a five-level system (overwhelming victory, narrow victory, draw, narrow defeat, overwhelming defeat). Thus, after round 2, each student can have obtained a maximum of 4 points in the tournament. Competitors in round 3 are those with the closest scoring after round 2. Each student can see his opponents' progress on live in real time in a synchronous way (identifier, name, tournament points, round points and

Fig. 2. A student interacting with ISCAR competition environment.

progress solving exercises). This way, there is a type of interaction between both users, which can see their evolution on real time. At present, other real time interactions of the students (e.g. by chat or collaborative activities) are not implemented but would be a nice future work to explore.

After each round, each student is aware of his/her global progress on the tournament and the different finished round results (with the scoring). Each student can also be aware of the progress of all the participants. But in addition, the student is also aware of his/her learning progress and the opponent in real time during the active round, since the student can see his/her scoring evolution on the current round and the active current exercise. This information changes dynamically after each student attempt.

From Figure 2, time left is 6 minutes and 24 seconds for both students at this moment. All students have the same total time for solving questions in each round. Student with id 100052714 is solving the sixth question and has obtained 1 out of 15 points until that moment, while student with id 1000052718 has not started solving questions yet. When the round finishes, each student will be able to see his absolute and relative position in the overall rankings of that round and of the global tournament. A complete description of ISCAR can be seen in [44].

4 MATERIAL AND METHODS

4.1 Experimental Design

TABLE 1
ISCARE MAIN FEATURES AND RELATIONSHIP WITH THE
FOUR DIMENSIONS OF THE 4NESS FRAMEWORK

4NESS dimension	ISCARE main features
Easiness	GUI: Simple and clear interface with only the basic information students need, and simple interaction with the environment.
Fairness	Swiss system: Non-elimination tournament format. Each tournament has a set of rounds (fewer than in a round-robin system). In each round, students are matched to compete one-on-one and without repeating previous opponents.
	Adaptive matching: In each round, students are matched with opponents of a similar knowledge level (based on scoring).
	Scoring system: In each round, one point depends on students' performance (percentage of right answers in a scale from 0 to 1), and another point depends on the outcome of the competition on a five-level system: overwhelming victory (1-0), narrow victory (0.75-0.25), draw (0.5-0.5), narrow defeat (0.25-0.75), overwhelming defeat (0-1)
Awareness	Rankings: Students can see their absolute and relative position in different rankings (rounds, tournaments)
	Opponent progress: Shown on live including the remaining time per round to complete a set of questions.
Adaptiveness	Adaptive questions: Questions received by competitors in each round are adapted to their level of knowledge of the subject.

The ISCARE competition environment was used to support learners of a course named Computer Architecture Laboratory during three consecutive editions. Computer Architecture Laboratory is a fourth-year course in the Telecommunication Engineering degree at Universidad Carlos III de Madrid, Spain. This course covers topics related to computer architecture, including shell script, makefile, system calls and the FAT file system. It has approximately 28 hours of classes distributed in 14 sessions. The course is very practical, with only four hours for theoretical sessions, and the remaining hours for practical sessions. During practical sessions students are provided with practical exercises that they have to solve and submit for evaluation.

4.2 Participants

Participants in the experiment were students from three consecutive editions of Computer Architecture Laboratory. A total of 13 students from year 1, 38 students from year 2, and 28 students from year 3, interacted with the ISCARE competition environment ($N = 79$). Most of the students were between 21 and 23 years old. We consider that a student participated in the experiment if he/she interacted with ISCARE. Nevertheless, not all the students who interacted with ISCARE filled in the final survey or took the pre-test and post-test used for evaluating the experience, as these were optional activities carried out at the beginning or

TABLE 2
SETUP FOR EACH EDITION OF THE COURSE ON COMPUTER
ARCHITECTURE LABORATORY

Setup	Year 1	Year 2	Year 3
Number of students	13	38	28
Number of sessions / hours	2 / 4	1 / 2	1 / 2
Number of rounds	5	4	4
Time per round (minutes)	10	10	10
Number of exercises per round	12	12	12
Use of a final survey	Yes	Yes	Yes
Use of pre-test and post-test	No	Yes	Yes
Adaptive questions (<i>adaptiveness</i>)	No	No	Yes

end of the sessions.

4.3 Setup of the Experience for each Edition

Table 2 shows a summary of the setup for the three editions of Computer Architecture Laboratory, including similarities and differences. The number of sessions for the experiment was two in the first edition (4 hours) and one for the last two editions (2 hours each). The first edition had one tournament with five rounds (shell script, makefile, system calls, FAT file system, and an overall review of all topics), while the second and third editions had one tournament as well, but only with the first four rounds. There were two main reasons for the removal of the fifth round in the last two editions: one reason was to be able to do everything in just one session and in a controlled setting (this way we can remove external effects such as students studying between the first and the second sessions); another reason was that most students solved correctly most of the exercises in the fifth round, so this round did not seem to add new skills.

The number of exercises per round was 12 in all cases, and the time to solve all the exercises was 10 minutes per round. In the three editions students filled in a survey about the experiment and the implementation of the four dimensions from 4NESS in ISCARE. The last two editions included also a pre-test and a post-test, because we wanted to compare learning gains in two different scenarios: without adaptation of contents (year 2) and with adaptation of contents (year 3). It is important to note that in all the editions students were monitored by teachers during the sessions so that they could not cheat (e.g., copying from each other or browsing the Web) while using the competition environment. In addition, it is important to note that the proposed questions in ISCARE, as well as in the pre-test and post-test usually required reasoning on the concepts and they cannot only be usually solved by memorization of a previous answer to a question. For example, a first multiple choice question can ask about a shell script loop using one example, while a second question can also ask about the shell script loop concept but using a different example. Memorizing the answer to the first question does not guarantee that the student will answer correctly the second question. Instead, the important thing is understanding the underlying concept and reasoning on it. The student might acquire the concept solving the first question because he/she can understand the concept but not because of memorization of the correct answer.

4.4 Methodology

A survey was designed to evaluate the meaningfulness of the experience and the implementation of the four dimensions of the 4NESS framework in IS CARE. The survey included a set of statements that had to be assessed by students using a 7-Likert scale. The 7-Likert scale is widely used in the literature and has several advantages over the 5-Likert scale [45], such as a more accurate measure. We used students' answers to the survey from the three editions (years 1, 2 and 3).

The final classification of students in the competition environment (rankings) contributed to the assessment of fairness, particularly of fairness in the scoring system and fairness in adaptive matching: to see if students were really paired with other students with similar knowledge level. Rankings from the 38 students in year 2 were used for this assessment, as this year had the highest number of participants. Students from different editions shall not be merged in a ranking analysis.

A pre-test and post-test design was used to compare learning gains and contribute to the assessment of adaptiveness (adaptive questions) in the competition environment.

This comparison was done between year 2 and year 3, as these two years share the main experimental setup (i.e. one session and four rounds). A total of 32 and 25 learners took the pre-test and post-test in year 2 and year 3, respectively.

5 RESULTS AND DISCUSSION

5.1 General students' satisfaction with the experience and competition environment

Students filled in a survey containing three direct statements about their general satisfaction with the experience and with the competition environment. These three statements were intended to find out whether the learning context was meaningful to them, and so, to confirm the validity of the case study as a testbed for this research. A 7-Likert scale was used in the three statements, ranging from 1 (totally disagree) to 7 (totally agree). Table 3 shows these statements and the results. Results show that, in general, students liked and were satisfied with the experience ($M=5.39$, $SD=1.42$), would like to repeat the experience ($M=5.17$, $SD=1.66$) and liked the competition environment ($M=5.49$, $SD=1.48$). These results show that competition driven by a computer-based environment that implements the four dimensions of the 4NESS framework can represent a satisfactory experience for students.

5.2 Evaluation of different features of the competition environment

This subsection evaluates the implementation of the four dimensions of 4NESS in computer-based competition environments. This evaluation is intended to gain insights on whether the four dimensions had the desired effect on

TABLE 3
STUDENTS' ANSWERS ABOUT GENERAL SATISFACTION WITH THE EXPERIENCES AND COMPETITION ENVIRONMENT

STATEMENT	TOTAL	M	SD
I liked and I am satisfied with the experience with the competition environment	77	5.39	1.42
I would like to repeat the experience with the competition environment in other courses	77	5.17	1.66
I liked the used competition environment	77	5.49	1.48

TABLE 4
STUDENTS' ANSWERS ABOUT THE EASINESS OF THE COMPETITION ENVIRONMENT.

STATEMENT	TOTAL	M	SD
The interaction with the competition environment was easy	75	5.97	1.13

students. In addition, these results allow comparing students' perceptions on some dimensions with respect to others, which can give a hint of the features that work better for computer-based competition environments, and the ones that should be changed or adapted according to students' perceptions.

5.2.1 Easiness

A statement in the survey was presented to the students to evaluate the implementation of the easiness in the competition environment, as defined in Table 1. Table 4 presents the results of students' answers. Students perceived the environment as easy to use ($M=5.97$, $SD=1.13$), which validates the implementation of the easiness dimension in the competition environment.

5.2.2 Fairness

The implementation of fairness in the competition environment includes three features, as defined in Table 1. Table 5 presents the set of statements used to evaluate these three features. The first two statements refer to the tournament type, and compare the Swiss system with eliminatory and round-robin systems. The third statement refers to the used of adaptive matching. The final statement serves to assess the scoring system.

Students liked more on average the adaptation of the Swiss system implemented in the competition environment, rather than an eliminatory or round-robin system. However, while the preference is clear against an eliminatory system ($M=5.78$, $SD=1.35$), it is not so strong against a league system ($M=4.66$, $SD=1.63$). It is noteworthy the trade-off regarding the selection of the system and the number of participants, as round-robin systems become impractical as the number of students increases, due to the large number of rounds. A combination of the adaptation of Swiss system used here with some type of league (e.g., grouping students in groups of five according to their classification) might be implemented to try to improve students' perceptions.

TABLE 5
STUDENTS' ANSWERS ABOUT THE FAIRNESS OF THE
COMPETITION ENVIRONMENT.

STATEMENT	TOTAL	M	SD
I like this tournament type more than an eliminatory system	72	5.78	1.35
I like this tournament type more than a round-robin system	70	4.66	1.63
The opponents I played against were of a similar level as me	73	5.58	1.29
I consider the scoring system used as fair	73	4.62	1.70

The matching algorithm was well assessed ($M=5.58$, $SD=1.29$), as students perceived that their opponents had a similar knowledge level. It is important to note that in the first round students are paired randomly, while in subsequent rounds the competition environment gets more and more information to provide fairer pairings.

The scoring system was moderately positively assessed by students ($M=4.62$, $SD=1.70$). However, the answers to this statement showed high diversity. Some redefinition of the scoring system might be considered to try to increase students' perceptions of fairness about the scoring system.

Among the 38 students who interacted with the ISCARE competition environment in year 2, the grades in the environment and the grades in the post-test of 32 of them were taken into account for the complementary analysis of fairness, as well as for the analysis of adaptiveness (subsection 5.2.4). These 32 students were the ones who completed the pre-test, post-test, final exam, and survey question about adaptiveness. The remaining six students interacted with the ISCARE competition environment, but were not taken into account in this complementary analysis as they did not complete one or more of the aforementioned tests and questions.

Figure 3 shows the relationship between students' final grades in the competition environment (X-axis) with respect to students' final grades in the post-test (Y-axis). The Pearson correlation between both variables is 0.543.

Based on these results, we can observe a general tendency according to which students' grades in the competition environment are related with students' grades in the post-test. However, there are several factors that prevent this relationship from being higher.

- There are several students with high grades in the post-test (some of them even got the maximum grade of 10 points). However, it is very difficult to get the maximum grade in the competition environment due to the scoring system. For example, if several students master the whole syllabus of the course, they will still lose 0.5 points in each match in case of tie.

- The exercises in the competition environment for this case study were of a level of difficulty such that, on average, more than half of the exercises were solved correctly by the students. As one point of the scoring is obtained from the resolution of exercises and one point of the scoring is obtained from the match result, then, on average, students lost more points due to their results

Fig. 3. Students' grades in the competition environment (X-axis) versus students' grades in the post-test (Y-axis).

against the opponent than due to answering incorrectly. For this case, a possible solution would be to include a correction coefficient to increase students' global scores.

- The scoring system tends to increase the grade of students with a lower level. For example, two students that are paired together but that do not answer correctly any of the questions would still receive 0.5 points each.

The Pearson correlation between students' final grades in the post-test and the average of opponents' final grades in the post-test is 0.405. The fact that this relationship is not higher can be explained by the low number of rounds played (four). The first round assigns pairings randomly, and a minimum number of rounds are required to get to match students of similar levels. There is, however, a trade-off for the total number of rounds: a large number does not allow conducting a controlled experiment, while a low number does not guarantee fair pairings.

All in all, the positive answers of learners to the statements, together with the analysis of students' post-test grades and students' grades in the competition environment allow to validate the implementation of the fairness dimension in the competition environment, but indicating that this dimension is moderately fair, and that a correction factor may be included to increase fairness regarding the scoring system.

5.2.3 Awareness

The implementation of awareness in the competition environment includes two features, rankings and opponents' progress, as defined in Table 1. Table 6 presents the statement used to evaluate each of these two features.

Students liked to see the rankings and classification of all the class after each round ($M=6.18$, $SD=1.00$). Students also assessed positively the possibility of seeing their own evolution together with their opponent evolution while they are solving the exercises ($M=5.70$, $SD=1.60$). The positive answers of learners to these statements validate the implementation of the awareness dimension in the

TABLE 6
STUDENTS' ANSWERS ABOUT THE AWARENESS OF THE
COMPETITION ENVIRONMENT.

STATEMENT	TOTAL	M	SD
I liked to see the partial ranking after each round	74	6.18	1.00
I liked to see my opponent evolution on live for each round	73	5.70	1.60

competition environment.

5.2.4 Adaptiveness

Adaptiveness was implemented in the competition environment by presenting students questions adapted to their level of knowledge on that particular topic. The implementation of this dimension in the competition environment is validated through a statement that student assessed, and through students' learning gains using a pre-test/post-test design with control and experimental groups. Students in the experimental group received questions adapted to their level of knowledge, while students in the control group received random questions.

First of all, students were asked about their perception related to the questions they answered using the competition environment. Table 7 shows the results for the experimental and control groups. The mean for students in the experimental group was 5.56 (SD=0.82), while the mean for the students in the control group was 4.84 (SD=1.51). Applying a one-tailed independent t-test for this variable between both groups ($t=2.29$, $p=0.013$, $df=49.77$), there is a statistically significant difference in favor of the experimental group. So, as expected, there is a more favorable opinion of the appropriateness of the difficulty level of the exercises to solve when there is an adaptation of the exercises, and students were able to perceive it.

Learning gains were calculated using the formulas in [46] for the cases in which a student got a post-test grade higher than his pre-test grade (formula a), and in which a student got a post-test grade lower than his pre-test grade (formula b). Therefore, learning gains will range within the interval $[-1, 1]$. Having a better post-test score was the most common case: 23 out of 25 for the experimental group, and 27 out of 32 for the control group. The formulas for calculating learning gains are the following: formula a (top) and formula b (down)

$$\frac{(post_test - pre_test)}{10 - pre_test} \quad (a)$$

$$\frac{(post_test - pre_test)}{pre_test} \quad (b)$$

The means of learning gains for the experimental group (group with adaptive questions) and for the control group (group with random questions) are 0.68 and 0.41 respectively. Applying an independent t-test between learning gains for experimental group and control group

TABLE 7
STUDENTS' ANSWERS ABOUT THE ADAPTIVENESS OF THE
COMPETITION ENVIRONMENT.

STATEMENT	TOTAL	M	SD
The exercises that I had to solve in the different rounds were of an appropriate difficulty level (<i>control group</i>)	32	4.84	1.51
The exercises that I had to solve in the different rounds were of an appropriate difficulty level (<i>experimental group</i>)	25	5.56	0.82

led to a statistically significant difference between both groups in favor of the experimental group ($t=2.39$, $p=0.010$, $df=54.71$). A learning increase of 0.55 sigmas is obtained in favor of the experimental group.

The mean of the pre-test for the control group was 6.90 (SD=1.65), while the mean for the experimental group was 6.88 (SD=1.79). Applying an independent t-test between the pre-test of control and experimental groups ($t=0.057$, $p=0.954$, $df=49.64$) showed no statistically significant difference between students' initial knowledge level in both groups. The high scores on the pre-test for both groups can be explained because both pre-test and post-test were part of the continuous evaluation and students knew it in advance, so they took these tests seriously and studied previously to the session.

A Pearson correlation between learning gains and the appropriateness of the difficulty level of exercises revealed that there is a statistically significant relationship between both variables ($r=0.469$, $p<0.001$), so as students learning gains were higher, their perceptions about the appropriateness of the difficulty level of exercises increases. This relationship is higher for the experimental group ($r=0.468$) than for the control group ($r=0.282$). This can be explained because as exercises are adapted to their knowledge level, students of the experimental group better perceive the appropriateness of the difficulty level of exercises, regardless of the learning gains they obtained.

When the adaptation capabilities of exercises are used in the competition environment, learning gains increase, and the opinions of students about the appropriateness of the difficulty level of exercises are better. The answers of learners to the statement, together with the pre-test/post-test analysis validate the implementation of the adaptiveness dimension in the competition environment, resulting computer-based competition environments that implement this dimension more effective than those which do not. Adaptation of exercises within the ISCARE competition environment resulted better, in terms of learning gains, than not using the adaptation capability. It is important to note that both (control and experimental) groups included adaptation of matches as a fairer way to pair students with a similar knowledge level (fairness dimension), and therefore the adaptiveness dimension only refers to exercises here.

6 CONCLUSION

The discrepancy on the use of computer-based competition environments in education is considerable,

with some papers claiming on their drawbacks, especially regarding students' negative feelings and predisposition to the formation of groups that favor peers belonging to the group versus those who are outside the group, and other papers highlighting their benefits for the learning process. We think that competition in education can be good but only if there are the adequate features during the learning process. An open issue is how to design properly competition features for computer environments. This paper sheds some light on the use of computer-based competition environments in education analyzing the effect of different features of computer-based competition environments and classifying these features in four dimensions that conform the 4NESS framework: easiness, fairness, awareness and adaptiveness. These four dimensions are particularized and implemented in a competition environment named ISCARE, which has been used to promote students' learning and motivation for three consecutive years of a course on Computer Architecture Laboratory. It is noteworthy that the 4NESS framework is not intended to be exhaustive covering each and every one of the possible dimensions of educational computer-based competition environments, but to try to simplify design decisions for technology providers with four main dimensions to take into account.

Although the results obtained from this case study are limited by the particular context of use (e.g., area of knowledge and students' profiles), the experience of using ISCARE throughout the three years in a course on Computer Architecture Laboratory has been found meaningful by learners, validating the pertinence of this study. The implementation of the four dimensions in ISCARE with specific features of the competition using a modification of the Swiss system of competition is positively evaluated according to students' satisfaction collected through a survey (four dimensions), easiness to use the environment (easiness dimension), students' final rankings in the competition environment and in the course (fairness dimension), students' awareness regarding rankings and opponents evolution in live (awareness dimension), and students' learning gains through control and experimental groups (adaptiveness dimension).

In addition, as far as we know, the effect of adaptiveness on learning gains has not been previously studied for computer-based competition environments. This work shows that the adaptation of exercises produced an increase of 0.5 sigmas comparing the group with adaptive questions (experimental group) and the group with random questions (control group). This result is consistent with learning gains obtained when comparing the adaptive condition in ITSs that do not

include competition [38]. We would like to remark that many previous works studied the impact on learning gains of adaptive questions (see e.g., [47]) and there is a review that includes comparison of learning gains involving educational systems with adaptation of questions [38]. But the novelty regarding this aspect is that none of these systems and studies included the competition element. Even if the studied adaptation condition is on the exercises, the fact of having a competition might (or not) have an effect on the learning gains. We should also note that we are measuring short-term learning since the post-test is done just after the intervention. A "delayed-post-test" or even several of them at different instants of time might have also been included to measure long-term learning.

There is a moderate relationship between students' post-test scores and their grades in the competition environment, as well as between students' post-test scores and opponents' post-test scores. These results also validate the fairness of the competition environment. However, the relationship is not perfect since for example the scoring system tends to narrow the gap between students with higher abilities and students with lower abilities.

Students' answers to the survey are consistent with the results from students' interactions with the computer-based competition environment: an increase in learning gains is obtained when using adaptive contents, which is in line with a better consideration of the appropriateness of exercises. In addition, the considerations of fairness about the scoring system and pairings are consistent with the comparison between post-test grades and students' grades in the competition environment.

As future work, we would like to run experiments using competition in massive environments, e.g. such as the ones enabled by MOOCs (Massive Open Online Courses). These experiments would allow extracting conclusions depending on different user features such as culture or age.

Moreover, an interesting future research line direction would be the extension of the learning analytics support for competition. For example, it would be very interesting to include the detection of new student indicators related to self-regulation and self-reflection and make students aware of these indicators, and study also the effect of the competition in relation with these indicators.

ACKNOWLEDGMENT

The authors acknowledge the eMadrid Excellence Network, funded by the Madrid Regional Government with grant No. S2013/ICE-2715, the Spanish Ministry of Economy and Competitiveness Project RESET (TIN2014-

53199-C3-1-R), and the European Erasmus+ projects MOOC-Maker (561533-EPP-1-2015-1-ES-EPPKA2-CBHE-JP), SHEILA (562080-EPP-1-2015-BE-EPPKA3-PI-FORWARD), COMPETEN-SEA (574212-EPP-1-2016-1-NL-EPPKA2-CBHE-JP), and COMPASS (2015-1-EL01-KA203-014033).

REFERENCES

- [1] A. Kohn, *No contest: The case against competition*, Houghton Mifflin Harcourt, 1992.
- [2] J. C. Hong, M. Y. Hwang, C. H. Lu, C. L. Cheng, Y. C. Lee, C. L. Lin, "Playfulness-based design in educational games: a perspective on an evolutionary contest game," *Interactive Learning Environments*, vol. 17, no. 1, pp. 15–35, 2009.
- [3] Z. H. Chen, C. Y. Chou, G. Biswas, T. W. Chan, "Substitutive competition: Virtual pets as competitive buffers to alleviate possible negative influence on pupils," *British Journal of Educational Technology*, vol. 43, no. 2, pp. 247–258, 2012.
- [4] M. Ediger, "Cooperative Learning versus Competition: Which Is Better?" *Journal of Instructional Psychology*, vol. 23, no. 3, pp. 204–300, 2001.
- [5] P. J. Muñoz-Merino, M. Fernández Molina, M. Muñoz-Organero, C. Delgado Kloos, "Motivation and Emotions in Competition Systems for Education: An Empirical Study," *IEEE Transactions on Education*, vol. 57, no. 3, pp. 182–187, 2014.
- [6] H. Tajfel, (Ed.) *Social identity and intergroup relations*. Cambridge University Press, 2010.
- [7] S. A. Haslam, N. Ellemers, S. D. Reicher, K. J. Reynolds, M. T. Schmitt, "The social identity perspective tomorrow: Opportunities and avenues for advance," In *Rediscovering Social Identity: Key Readings*. Psychology Press, Taylor & Francis, 2010.
- [8] D. A. Stapel, W. Koomen, "Competition, cooperation, and the effects of others on me," *Journal of Personality and Social Psychology*, vol. 88, no. 6, pp. 1029–1038, 2005.
- [9] L. M. Regueras, E. Verdú, M. F. Muñoz, M. A. Pérez, J. P. de Castro, M. J. Verdú, "Effects of competitive e-learning tools on higher education students: A case study," *IEEE Transactions on Education*, vol. 52, no. 2, pp. 279–285, 2009.
- [10] P. J. Muñoz-Merino, M. Fernández Molina, C. Delgado Kloos, "Analyzing Learning Gains in a Competition Intelligent Tutoring System," *Proc. Intelligent Tutoring Systems Conference* (pp. 662–663). Springer Berlin Heidelberg, 2014.
- [11] J. C. Burguillo, "Using game theory and competition-based learning to stimulate student motivation and performance," *Computers & Education*, vol. 55, pp. 566–575, 2010.
- [12] N. E. Cagiltay, E. Ozcelik, N. S. Ozcelik, "The effect of competition on learning in games," *Computers & Education*, vol. 87, pp. 35–41, 2015.
- [13] G. Robles, J. M. Gonzalez-Barahona, A. Moral, "A synchronous on-line competition software to improve and motivate learning," *Proc. Of the IEEE Global Engineering Education Conference, EDUCON 2012* (pp. 1–8), 2012.
- [14] B. Silva, R. N. Madeira, "A study and a proposal of a collaborative and competitive learning methodology," *Proc. Of the IEEE Global Engineering Education Conference, EDUCON 2010* (pp. 1011–1018), 2010.
- [15] S. Attle, B. Baker, "Cooperative Learning in a Competitive Environment: Classroom Applications," *International Journal of Teaching and Learning in Higher Education*, vol. 19, no. 1, pp. 77–83, 2007.
- [16] C. P. Lin, S. S. Young, H. C. Hung, "The game-based constructive learning environment to increase English vocabulary acquisition: Implementing a Wireless Crossword Fan-Tan Game (WICFG) as an example," *Proc. Of the Fifth IEEE International Conference on Wireless, Mobile, and Ubiquitous Technology in Education, WMUTE 2008* (pp. 205–207), 2008.
- [17] H. N. Cheng, W. M. C. Wu, C. C. Y. Liao, T. W. Chan, "Equal opportunity tactic: Redesigning and applying competition games in classrooms," *Computers & Education*, vol. 53, no. 3, pp. 866–876, 2009.
- [18] J. M. Pfothenauer, D. J. Gagnon, M. J. Litzkow, C. C. Blakesley, "Designing and using an on-line game to teach engineering," *Proc. Of the 39th IEEE Frontiers in Education Conference, FIE 2009* (pp. 1–5), 2009.
- [19] L. J. Chang, J. C. Yang, F. Y. Yu, "Development and evaluation of multiple competitive activities in a synchronous quiz game system," *Innovations in Education and Teaching International*, vol. 40, no. 1, pp. 16–26, 2003.
- [20] Y. K. Chua, R. S. Balkunje, "An Exploratory Study of Game-based M-learning for Software Project Management," *Journal of Universal Computer Science*, vol. 18, no. 14, pp. 1933–1949, 2012.
- [21] S. Arnab, T. Lim, M. B. Carvalho, F. Bellotti, S. Freitas, S. Louchart, A. De Gloria, "Mapping learning and game mechanics for serious games analysis," *British Journal of Educational Technology*, vol. 46, no. 2, pp. 391–411, 2015.
- [22] M. Julian Mateos, P. J. Muñoz-Merino, C. Delgado Kloos, D. Hernández-Leo, D. Redondo-Martínez, "Design and evaluation of a computer based game for education," *Proc. Of the IEEE Frontiers in Education Conference, FIE 2016* (pp. 1–8), 2016.
- [23] M. J. Callaghan, N. McShane, T. Teill, P. Raspail, "Practical application of the Learning Mechanics-Game Mechanics (LM-GM) framework for Serious Games analysis in engineering education," *Proc. Of the 13th International Conference on Remote Engineering and Virtual Instrumentation (REV)* (pp. 391–395), 2016.
- [24] S. A. Sukiman, H. Yusop, R. Mokhtar, N. H. Jaafar, "Competition-Based Learning: Determining the Strongest Skill that Can Be Achieved Among Higher Education Learners," In *Regional Conference on Science, Technology and Social Sciences* (pp. 505–516). Springer Singapore, 2016.
- [25] A. Dix, "Human-computer interaction". Springer US, 2009.
- [26] R. M. Chory-Assad, "Classroom justice: Perceptions of fairness as a predictor of student motivation, learning, and aggression," *Communication Quarterly*, vol. 50, no. 1, pp. 58–77, 2002.
- [27] L. McAlpine, C. Weston, "Reflection: Issues related to improving professors' teaching and students' learning," *Instructional science*, vol. 28, no. 5, pp. 363–385, 2000.
- [28] P. Phobun, J. Vicheanpanya, "Adaptive intelligent tutoring systems for e-learning systems," *Procedia-Social and Behavioral Sciences*, vol. 2, no. 2, pp. 4064–4069, 2010.
- [29] P. A. Kirschner, "Cognitive load theory: implications of cognitive load theory on the design of learning," *Learning and Instruction*, vol. 12, pp. 1–10, 2002.
- [30] A. P. Chalmer, "The role of cognitive theory in human-computer interface," *Computers in Human Behavior*, vol. 19, pp. 593–607, 2003.
- [31] F. Paas, A. Renkl, J. Sweller, "Cognitive load theory and instructional design: Recent developments," *Educational psychologist*, vol. 38, no. 1, pp. 1–4, 2003.

- [32] D. R. Appleton, "May the best man win?" *The Statistician*, vol. 44, no. 4, pp. 529-538, 1995.
- [33] N. M. Bengtson, "An illustration of some basic probability concepts: determining probabilities of winning in single elimination tournaments," *Mathematics and Computer Education*, vol. 44, no. 2, pp. 92-104, 2010.
- [34] K. Verbert, E. Duval, J. Klerkx, S. Govaerts, J. L. Santos, "Learning analytics dashboard applications," *American Behavioral Scientist*, vol. 57, no. 10, pp. 1500-1509, 2013.
- [35] S. Govaerts, K. Verbert, E. Duval, A. Pardo, "The student activity meter for awareness and self-reflection," *Proc. Of the Computer-Human Interaction*, CHI 2012 (pp. 869-884), 2012.
- [36] P. Brusilovsky, "Methods and techniques of adaptive hypermedia," *In Adaptive hypertext and hypermedia* (pp. 1-43). Springer Netherlands, 1998.
- [37] N. Henze, W. Nejdl, "Logically characterizing adaptive educational hypermedia systems," *Proc. Of the International Workshop on Adaptive Hypermedia and Adaptive Web-based Systems*, AH 2003 (pp. 20-24), 2003.
- [38] K. VanLehn, "The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems," *Educational Psychologist*, vol. 46, no. 4, pp. 197-221, 2011.
- [39] R. Baker, K. R. Koedinger, A. T. Corbett, A. Z. Wagner, S. Evenson, "Adapting to when students game an intelligent tutoring system," *Proc. Of the Intelligent tutoring systems* (pp. 392-401). Springer Berlin Heidelberg, 2006.
- [40] V. Aleven, K. R. Koedinger, H. C. Sinclair, J. Snyder, "Combating shallow learning in a tutor for geometry problem solving," *Proc. Of the Intelligent Tutoring Systems Conference* (pp. 364-373). Springer Berlin Heidelberg, 1998.
- [41] A. C. Graesser, P. Chipman, B. C. Haynes, A. Olney, "AutoTutor: An intelligent tutoring system with mixed-initiative dialogue," *IEEE Transactions on Education*, vol. 48, no. 4, 612-618, 2005.
- [42] T. Y. Chang, Y. R. Ke, "A personalized e-course composition based on a genetic algorithm with forcing legality in an adaptive learning system," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 533-542, 2013.
- [43] V. Caputi, A. Garrido, "Student-oriented planning of e-learning contents for Moodle," *Journal of Network and Computer Applications*, vol. 53, pp. 115-127, 2015.
- [44] P. J. Muñoz-Merino, M. Fernández Molina, M. Muñoz-Organero, C. Delgado Kloos, "An adaptive and innovative question-driven competition-based intelligent tutoring system for learning," *Expert Systems with Applications*, vol. 39, no. 8, pp. 6932-6948, 2012.
- [45] K. Finstad, "Response interpolation and scale sensitivity: Evidence against 5-point scales," *Journal of Usability Studies*, vol. 5, no. 3, pp. 104-110, 2010.
- [46] V. Aleven, V. M. McLaren, I. Roll, K. R. Koedinger, "Toward Tutoring Help Seeking: Applying Cognitive Modeling to Meta-Cognitive Skills," *Proc. Of the Intelligent Tutoring Systems Conference* (pp. 227-239). Springer Berlin, Heidelberg, 2004.
- [47] M. Barla, M. Bieliková, A.B. Ezzeddinne, T. Kramár, M. Šimko & O. Vozár, "On the impact of adaptive test question selection for learning efficiency," *Computers & Education*, 55(2), 846-857, 2010.

Pedro J. Muñoz-Merino is a lecturer and researcher at the Universidad Carlos III de

Madrid, where he is the Director of the Master in Telematics Engineering. In 2009 he got his PhD in Telematics Engineering from the Universidad Carlos III de Madrid. He obtained his accreditation in May 2012 as Associate Professor by the ANECA agency from the Spanish Ministry of Education. Pedro has received several awards for his work on educational technologies. He is author of more than 90 scientific publications and has participated in more than 30 research projects, coordinating some of them with private companies. His skills and experience include research and development in learning analytics, educational data mining, evaluation of learning experiences, user studies, gamification or Intelligent Tutoring Systems.

Carlos Alario-Hoyos is a Visiting Associate professor in the Department of Telematics Engineering at the Universidad Carlos III de Madrid. He received M.S. degree in Telecommunication Engineering and PhD in Information and Communication Technologies from Universidad of Valladolid, Spain, in 2007 and 2012 respectively. His skills and experience include research and development in MOOCs, social networks, collaborative learning, and evaluation of learning experiences, among others.

Mario Muñoz-Organero received the M.Sc. degree in telecommunications engineering from the Polytechnic University of Catalonia, Barcelona, Spain, in 1996, and the Ph.D. degree in telecommunications engineering from the Universidad Carlos III de Madrid, Leganes, Spain, in 2004. He is a Professor of telematics engineering at the Universidad Carlos III de Madrid. He has also more than four years of experience working for the telecommunications industry in companies such as Telefonica R&D and Lucent Technologies, both in Madrid, Spain. He has participated in several European-funded projects such as E-LANE and OSAMI. He has been the lead researcher of the Spanish-funded HERMES, ARTEMISA, HAUS, COMINN, REMEDISS and IRENE projects. He was also the IP of the GEEWHEZ EU-FP7 project and is currently an MC member of TU1305 EU funded COST action about social networks and travel behavior.

Carlos Delgado Kloos received the PhD degree in Computer Science from the Technische Universität München and in Telecommunications Engineering from the Universidad Politécnica de Madrid. He is full professor of Telematics Engineering at the Universidad Carlos III de Madrid, where he is the director of the GAST research group, director of the UNESCO Chair on "Scalable Digital Education for All", and vicepresident for Strategy and Digital Education. He is also

the Coordinator of the eMadrid research network on Educational Technology in the Region of Madrid. He is the Spanish representative at IFIP TC3 on Education.

Manuel Fernández computer science more than 7 years the development management of

Molina is a engineer with of experience in and solutions and

projects related to the Information Technologies.