Chapter 9

CALCULATED QUESTIONS AND E-CHEATING: A CASE STUDY

Juan Carlos González de Sande
ETSIS de Telecomunicación, Universidad Politécnica de Madrid, Spain

ABSTRACT
Many learning management systems permit to configure a questionnaire based on an existing item bank. This item bank should be large enough in order to assure that the students do not know the questions (and the corresponding right answer without any study) after several colleagues have solved the questionnaire. A way to minimize this problem is by creating a very large item bank (several thousands of items). In many engineering and science disciplines is an easy task to automatically generate random numerical variants of the same question. The answer of such question is numerical and it is obtained after some calculation using one or more parameters that are randomly assigned by the learning management system. This type of questions is called "calculated questions". We have noticed that, even using calculated questions, there are some students that correctly answer the questionnaire in such a fast time that make the instructors think they have obtained some unfair advantage. During the time that some of these questionnaires was open, we have introduced a new calculated question and followed the evolution of the wrong/right answers over time. We have focused our attention on the students that solved the questionnaire in a fast time. Results show that after a few hours and after the first tenth of students have answered the new question, a surprisingly high proportion of students that solve the questionnaire in a fast way, answer the new question correctly.

Keywords: online questionnaires, calculated questions, computer based assessment, e-cheating.

1. INTRODUCTION

During last decades, several forms of e-learning have been adopted progressively to some extent in all educational levels. For example, in United States, the proportion of higher education students taking at least one online course has steadily increased from less than 10% to more than 30% over the first decade of the 21st century (Bowen, Chingos, Lack, & Nygren, 2013). This is also the case for European Educational Institutions where the use of e-assessment has been increased sharply since the beginning of the present century (Whitelock, Road, & Ripley, 2007).

The delivery of online test that are automatically assessed in real time is one of the most popular tools that e-learning provides to students and teachers. These online tests could be part of a complete online course or of a blended learning based course. In both cases, web based tests could be intended for assessment purposes, in order the students could check their progress (in real time, with instant feedback, anytime and anywhere), or for fulfilling both objectives (Whitelock et al., 2007). When used for assessment purposes, the main drawback of online tests is the ease of cheating (Bedford, Gregg, & Clinton, 2011; de Sande et al., 2010; King & Case, 2014; Rowe, 2004; Young, 2012). Additional disadvantages has been also pointed out as, for example, bias towards tech-savvy students over non-technical students and lack of social interaction between teachers and students.

Learning management systems (LMS) usually include several features that can be used to partially thwart online cheating: the use of username and password to acces to the test, the use of large item bank, randomization of the questions, limit the time when the test is available, limit the time to complete the test, etc. (Cluskey, Ehlen, & Raiborn, 2011; King & Case, 2014; Smith, 2013). Many other measures have been proposed to reduce the online cheating, for example, academic dishonesty policy dissemination, setting a cheating trap in the web, the use of a class mole, statistical analysis to detect common errors, webcam surveillance, live or remote proctoring and/or the use of lockdown browsers (Bedford et al., 2011; Cluskey et al., 2011; Moten, Fitterer, Brazier, Leonard, & Brown, 2013). However, many of this strategies have a high cost (Cluskey et al., 2011) or are contrary to one of the great advantages of e-learning: the possibility of learn anytime and anywhere (Whitelock et al., 2007).

Regarding to the randomization of the tests, LMS usually offer several ways to create tests with some degree of randomness (de Sande 2010, 2011; Guimarães Pereira & Scheuermann, 2007; Montes, Deza, & de Sande, 2011; Pachler, Daly, Mor, & Mellar, 2010). Typical ways to obtain different quizzes are picking questions from a large item bank, changing the order in which the questions are presented and changing the order in which the possible answers are presented (at least for the case of multiple choice questions). It has been observed that when using a not too large item bank, some students solved the questionnaires in a suspiciously reduced time and obtained a good result (de Sande et al., 2010). A way to easily increase the item bank (till several thousands of items) is to automatically generate numerical variations of a set of base questions (de Sande, 2010, 2011; Montes et al., 2011). However, even in this case, after several semesters using the same item bank, it has been observed that some students obtained good results in surprisingly short time which is an indication of dishonest behavior (Moten et al., 2013).

The main goal of this work is to analyze if the use of calculated questions could avoid or mitigate the online cheating. Following this introductory section, the Background will be described in Section 2. The design and method used to develop the present study will be presented in Section3. The main results of the work will be analyzed and discussed in Section 4. Some future research directions are mentioned in Section 5. Finally, Section 6 is devoted to the concluding remarks of this work.

2. BACKGROUND

Since 2009/10 academic year, the Higher School of Telecommunications and System Engineering (ETSIST) at the Technical University of Madrid (UPM) offers several degrees adapted to the European Higher Education Area. Both the courses content and the learning and teaching methods have been updated and a blended learning scheme have adopted in most of the courses included in such degrees. This is the case for Signals and Systems course that is a mandatory course of the Electrical and Electronics Engineering studies. Moodle platform (Moodle, 2014) has been selected by the UPM as the LMS for delivering different course material: course information including schedule, expected learning outcomes and competences, assessment methods, course content, slides, solved exercises, proposed homework, etc. Most of students (over 95%) enrolled in Signals and Systems course choose to follow a continuous assessment method instead of being assessed by means of a final exam exclusively. The continuous assessment method proposed in Signals and Systems course includes automatically assessed online tests delivered via Moodle (de Sande, Godino-Llorente, Osma-Ruiz, & S’enz-Lechon, 2012).

A large item bank (around 5000 different items for each questionnaire) has been developed during 2010 and it has been used to deliver online tests since 2010/11 academic year. After two consecutive semesters using the same item bank, it was noticed that a large group of students solved the test in a surprisingly short time and obtained very good results (de Sande et al., 2012). The high marks obtained in the test by a considerable large group of students (around forth of students) that fulfilled the test in a short time made the author suspicious that some students were cheating somehow when they fulfilled the online tests. This was the reason for which the present work has been carried out.

3. DESIGN AND METHOD

The present study has been developed in ETSIST at UPM. Signals and Systems is a mandatory course of the Electrical and Electronics Engineering studies. Most of students enrolled in the course are sophomore students. The course is divided into four different subjects and at the end of each subject a questionnaire is delivered via Moddle. The marks obtained in the online tests counts for obtaining the final mark. The weight of each individual questionnaire ranged from 3 to 5% of the final mark. These online tests were delivered during two to four days and the students could solve it anywhere and anytime during that period. The time to solve a test was 30 min since the test was opened by each student. The test corresponding to the same subject of the course (Fourier analysis of continuous time signals and systems) has been selected for this study. The execution time as well as the date and hour when the test was taken as well as the final marks of this test have been analyzed for several semesters.

Since 2010/11 academic year, a large item bank (around 5000 items for each questionnaire) has been used to deliver the questionnaires. Each questionnaire includes 10 calculated questions and each question is selected from a set of items created as numerical variations (100 different variations) of several base items (from 3 to 8 depending on the question). An example of calculated question for Signals and Systems is given in Figure 1. This type of question can be classified as an analyzing process according to the Bloom's taxonomy (Krathwohl, 2002). The same example could be used changing "the magnitude" by "the phase", so two different base questions with similar difficulty are created. By using different wordings of the same question or substituting "angular frequency (in rad/s)" by "frequency (in Hz)" and so on, it is easy to obtain a set of 3 to 8 base questions with similar difficulty, and then create from 300 to 800 different items. For each student, only one of these items is selected to create one question of the questionnaire. It is expected that the students should derive the expression for the solution (where the parameters should be substituted by their numerical values) and afterwards they should evaluate this expression for the particular values of the parameters. Note that the students should find and introduce a numerical answer to each questions, they are not asked to select an option. The same procedure is followed for each of the ten questions of the test. Additionally, the questions are randomly arranged.

Figure 1. Example of calculated question. 100 randomly generated sets of parameters a_j, b_j,(j=0,1,2) and ω_0 together with their corresponding right answers are stored in the item bank.

Question: Given a linear and time invariant system characterized by the equation $a_0y(t)= b_0x(t) +b_1x'(t) +b_2x''(t) - a_1y'(t) - a_2y''(t)$, find the magnitude of its frequency response for the angular frequency ω_0 (in rad/s).

Solution:

$$|H(j\omega_0)| = \sqrt{\frac{(b_0 - b_2\omega_0^2)^2 + b_1^2\omega_0^2}{(a_0 - a_2\omega_0^2)^2 + a_1^2\omega_0^2}}$$

The online tests were delivered during two to four consecutive days that were announced at least three weeks in advance via Moodle and during the previous week in classroom. Students could solve the tests anywhere and anytime within the allowed period. They could use their notes, books, etc., but were asked to solve the test individually.

Moodle platform registers all the activities carried out by each student along the course. The delivered online tests were analyzed. During the first semesters analyzed in this work (academic years since 2011/12 and 2012/13), only the execution time for each student as well as the mark obtained were studied. A total of 102 and 168 students solved the test during the 2011/2012 and 2012/2013 academic years, respectively. Graphs of these data for the two semesters are presented in Figs 2 and 3. Many students solved the test in a short time and obtained a good mark. It is reasonable to think that most of those students knew the general solution for most of the base questions so they only had to substitute the numerical values for finding the right answer to each question. Probably, some of students enrolled in the course for second time had collected a set of solutions, or may be, groups of students worked together to solve the tests (Young, 2012) and found the solutions for different sets of questions.

In the 2013/14 academic year, a set of items used for a given question was replaced by a new set that was created following the previously described procedure. The date and hour when each student opened his/her test, as well as the time employed to solve it were analyzed and compared to test mark and to the right/wrong answer to the new question. A total of 76 students, 13 females and 63 males, solved the test under the described conditions.

4. RESULTS AND DISCUSSION

Figure 2 shows that, during the first semester studied, corresponding to the 2011/12 academic year, 23 out of 102 students solved a 10 items test in less than 12 min and most of them obtained a mark over 9 (except 4 outliers). The mean mark for this group of students was 8.4 (on a 0 to 10 scale), considerably higher than the mean mark for all the students that solved the test during that semester (7.4). It is expected that most of students did not take unfair advantage in solving the test. However, it is difficult to believe that a student could correctly solve 10 questions like the example given in Fig. 1 in less than 12 min without any unfair help (19 students, nearly the fifth of the total number of students, did it and obtained a test mark over 9).

A similar behavior is observed in Fig. 3 for the marks obtained by the students of the following academic year. The number of students that carried out the test in a short time is larger than in the previous case, up to 40 out of 168 students fulfilled the test in less than 12 min and obtained a mark over 9. In this case, the mean mark for the group of students that solved the test in less than 12 min was 8.8 while the mean mark of all the students that solved the test was 6.8. In this case, the difference between the mean mark of the students that solved the test in a short time and that of the all students is even larger than during the previous semester.

Figure 2. Performance of students of the 2011/12 academic year in a test including 10 calculated questions.

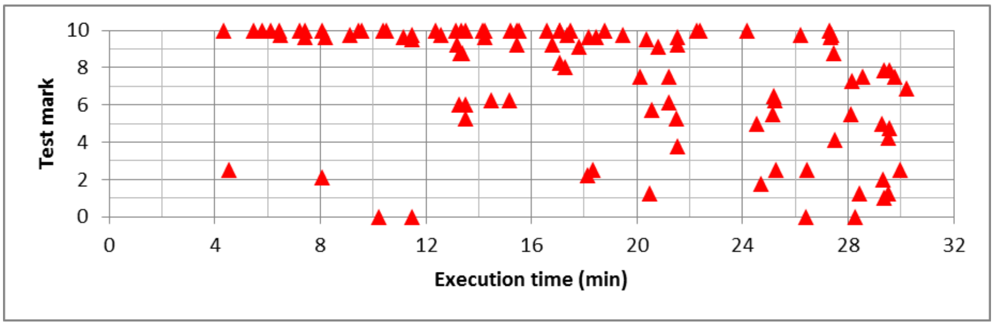
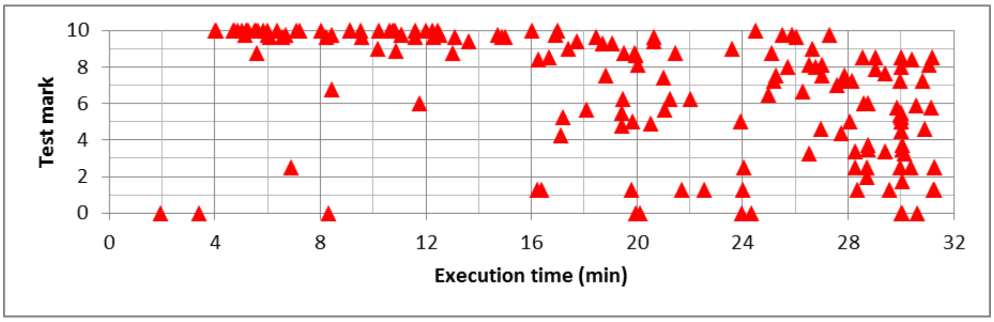


Figure 3. Performance of students of the 2012/13 academic year in a test including 10 calculated questions.



During 2013/14 academic year, the test under study was slightly changed: a set of items for creating a given question was changed by a new set of items. Three different base calculated questions of similar difficulty (wording variations or asking for different aspects of the same issue) were introduced in such a way that 300 new items were automatically created as numerical variations of the base questions. The LMS randomly selected just one new item anytime a student opened the test. The date and hour when each student opened the test, as well as the execution time, total mark of the test and the answer to the new question were analyzed. The students were divided in groups depending on their answer to the new question and labelled either as 'R' for right answer or 'W' for wrong answer. Depending on the execution time, label 'F' was assigned to those students who fulfilled their test in less than 15 min and label 'S' to those who fulfilled their test in more than 15 min. Then four groups of students appeared: students that solved the test in less than 15 min and give the right (wrong) answer to the new question, denoted by R-F (W-F) and red diamonds (blue triangles) in the graphs of Figs. 4 and 5, and students that solved the test in more than 15 min and give the right (wrong) answer, denoted by R-S (W-S) and orange circles (green squares).

Figure 4 shows the total mark obtained by the students as a function of the time employed to fulfil the test during the third semester studied (2013/14 academic year). It could be observed a similar shape of the graph to those of Fig. 2 and 3. However, there is a slight difference for those students that fulfilled the test in a fast way (label 'F'); their mean mark is slightly lower (7.9) than those of the equivalent group in the previous two semesters. In fact the fastest students obtained a mark around 8 instead of over 9 as it was observed in Figs. 2 and 3.

Figure 4. Right and fast (R-F), wrong and fast (W-R), right and slow (R-S), or wrong and slow (W-S) answer to the new question vs execution time. In vertical appears the overall mark obtained by each student in this test.

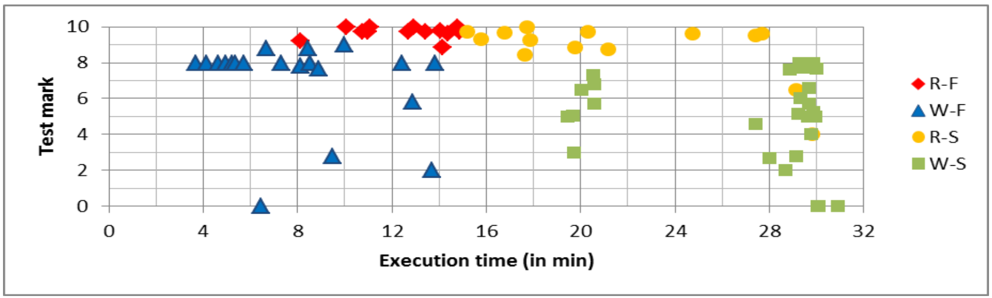
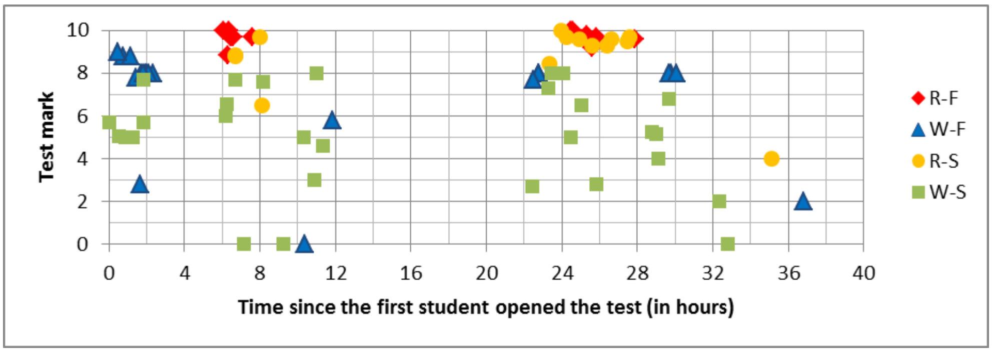


Figure 5 shows the total mark obtained for each student as a function of the time when they opened the test. Then the evolution of the right and wrong answers to the new question can be appreciated. Again, the data are presented for the same four groups as in Fig 4.

Figure 5. Right and fast (R-F), wrong and fast (W-R), right and slow (R-S), or wrong and slow (W-S) answer to the new question vs time since the first student opened his/her test. In vertical appears the overall mark obtained by each student in this test.



It can be observed that the first right answer does not appear till six hours later than the first student opened his test. Moreover, up to 17 students opened the test and give a wrong answer to the new question before the first right answer appears. However, 45% of students that opened and solved the test after the first right answer was given, correctly answered the new question. Assuming that this percentage represents the probability that a student correctly answers that question, it is hard to find a sequence of 17 consecutive wrong answers (the probability of such fact would be less than 4×10^{-5}).

Considering only the students that opened the test after the first right answer to the new question had appeared, it can be observed that: i) 59 % of students that solved the test in less than 15 min, correctly answered the new question, while only 39% of those that solved the test in more than 15 min, correctly answered the new question. This result does not mean that all students that solved the test in a short time behaved in unfair way, but it is a suspicious result. In the same way, it is possible that some students that solved the test in a long time also took some unfair advantage of the conditions for doing the online tests.

On the other hand, the evolution of the right/wrong answers to the new questions also suggests that students that opened the test during the first 5 hours and cheated probably used collected solutions from the previous years. May be some students were provided with the old solutions but were not warned about the changes (blue triangles around 22 and 30 hours in Fig. 5). Those students provided with solutions from the previous semesters that opened the test during the first hours were probably surprised by the new question, so they answered with an old wrong rule. Probably, some of them warned to others students that possibly worked together to find the general solution of the new question (Young, 2012). After they found the solution, they probably disseminated it among the rest of students interested in cheating.

There is no doubt that calculated questions help instructors to ask for higher level cognitive processes, as application, analysis and synthesis or evaluation, in the Bloom's taxonomy (Krathwohl, 2002). The randomization of the parameters included in each question, could partially help in the objective of reduce the online cheating. However, students could be able to elaborate a recipe for each question type (the used formula to answer the question) and collect them together in library test files. Then many students that opened the test afterwards and students of the following semesters could use a single sheet with a set of recipes to solve the test and obtain a high score.

5. FUTURE RESEARCH DIRECTIONS

Online test intended for assessment, and in general, online learning presents unquestionable advantages for students interested in learning, especially for those who have difficulties due to distance, disability, illness, or work commitments (Whitelock et al., 2007). Online learning also permits to develop potential high-quality instruction with a considerable cost reduction for educational institutions, (Bowen et al., 2013). However, online test assessment, and in general online assessment, also presents great advantages for students interested in cheating (de Sande et al., 2010; King & Case, 2014; Moten et al., 2013; Rowe, 2004; Watson and Sottile, 2010; Young, 2012). This is only one of the reasons why it is necessary to develop rigorous studies on learning outcomes for students enrolled in online courses or courses that include some kind of online assessment.

6. CONCLUSION

The comparison of the execution time of online tests and the marks obtained by students in such tests, made the author suspicious about the way the students solved the tests. An experiment was designed to confirm this. A new question was introduced in a given test and both the time taken to solve it and the sequence of right/wrong answers to the new question was analyzed. An extremely improbable sequence of 17 consecutive wrong answers to the new question was found at the beginning of the period given for solving the test. After the first right answer was given (about 6 hours after the first student opened his/her test), the percentage of students that correctly answered the new question was significantly higher (59 %) for the group of students that solved the test in a short time than for those that solved it in a longer time than 15 min (39%). Calculated questions could be used to easily create questions that require higher level cognitive processes. Then, instructors have a powerful tool to engage and help students interested in learning. Cheating may be harder when these calculated questions are used to create questionnaires, but dishonest behavior of students is not prevented by only using such type of questions. Further research is necessary in order to develop secure, valid, and reliable online assessment.

REFERENCES

Bedford, D. W., Gregg, J. R., & Clinton, M. S. (2011). Preventing online cheating with technology: A pilot study of remote proctor and an update of its use. *Journal of Higher Education Theory and Practice*, *11*(2), 41-59.

Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2013). Online learning in higher education. *Education Next*, *13*(2), 59-64.

Cluskey Jr., G. R. Ehlen, C. R., & Raiborn, M. H. (2011). Thwarting online exam cheating without proctor supervision. *Journal of Academic and Business Ethics*, *4*, 1-7.

de Sande, J. C. G. (2010). Automatic generation of signals and systems exercises. In L. Gómez Chova, D. Martí Belenguer, & I. Candel Torres (Eds.), *Conference proceedings of the 3rd International Conference of Education, Research and Innovation* (pp. 1991-1998). Madrid, Spain: IATED.

de Sande, J. C. G. (2011). Computer-based training tool for signals and systems exercises. *International Journal of Engineering Education, 27*, 1150-1157.

de Sande, J. C. G., Fraile, R., Arriero, L., Osma-Ruiz, V., Osés, D., & Godino-Llorente, J. I. (2010). Cheating and learning through web based tests. In L. Gómez Chova, D. Martí Belenguer, & I. Candel Torres (Eds.), *Conference proceedings of the 3rd International Conference of Education, Research and Innovation* (pp. 5004-5009). Madrid, Spain: IATED.

de Sande, J. C. G., Godino-Llorente, J. I., Osma-Ruiz, V., & S'enz-Lechon, N. (2012). Analysis of the validity of e-assessment and self-assessment in formal assessment in electrical and electronics engineering studies through a case study. In IEEE Computer Society. (Ed.), *Proceedings of the 12th IEEE International Conference on Advanced Learning Technologies* (pp. 31-35). Institute of Electrical and Electronics Engineers.

Guimarães Pereira, A., & Scheuermann, F. (2007). *On e-testing: An overview of main issues* (JRC Tecnical Notes). Italy: European Commission. Retrieved March 24th, 2014, from: http://publications.jrc.ec.europa.eu/repository/bitstream/JRC36015/reqno_jrc36015_background%20note%20on%20e-testing%5B2%5D.pdf

King, D. L., & Case, C. J. (2014). E-cheating: Incidence and trends among college students. *Issues in Information Systems*, *15*(1), 20-27.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, *41*(4), 212-218. doi:10.1207/s15430421tip4104_2

Montes, J., Deza, J., & de Sande, J. C. G. (2011). *Signals and systems students' trainer.* In L. Gómez Chova, D. Martí Belenguer, & A. López Martínez (Eds.), *EDULEARN11 Proceedings CD: 3rd International Conference on Education and New Learning Technologies* (pp. 3801-3808). Barcelona, Spain: IATED.

Moodle. (2014). *Moodle: A free, open source code course management system for online learning*. Retrieved September 1, 2014, from http://moodle.org

Moten Jr., J., Fitterer, A., Brazier, E., Leonard, J., & Brown, A. (2013). Examining online college cyber cheating methods and prevention measures. *Electronic Journal of e-Learning, 11*(2), 139-146.

Pachler, N., Daly, C., Mor, Y., & Mellar, H. (2010). Formative e-assessment: Practitioner cases. *Computers & Education, 54*(3), 715-721.

Rowe, N. C. (2004). Cheating in online student assessment: beyond plagiarism. *Online Journal of Distance Learning Administration VII*(II). Retrieved March 31, 2014, from http://www.westga.edu/~distance/ojdla/summer72/rowe72.html

Smith, S. (2013). *Controlling cheating in online courses: A primer*. Pearson Education. Retrieved October 21, 2014, from http://www.pearsonmylabandmastering.com/global/northamerica/assets/pdf/controlling-cheating-in-online-courses.pdf

Watson, G., & Sottile, J. (2010). Cheating in the digital age: Do students cheat more in online courses? *Online Journal of Distance Learning Administration, XIII*(I). Retrieved October, 21 2014, from http://www.westga.edu/~distance/ojdla/spring131/watson131.html

Whitelock, D., Road, M., Ripley, M. (2007) *Effective practice with e-assessment: An overview of technologies, policies and practice in further and higher education*. Joint Information System Committee (JISC). Retrieved September 1, 2014, from http://www.jisc.ac.uk/media/documents/themes/elearning/effpraceassess.pdf

Young, J. R. (2012, June 3). *Online classes see cheating go high-tech*. Article posted to The Chronicle of Higher Education. Retrieved September 1, 2014, from http://chronicle.com/article/Cheating-Goes-High-Tech/132093/

AUTHOR(S) INFORMATION

Full name: Juan Carlos González de Sande
Institutional affiliation: ETSIS de Telecomunicación, Universidad Politécnica de Madrid
Institutional address: Campus Sur-UPM, Ctra de Valencia, km 7, 28031 Madrid, Spain
Biographical sketch: Juan Carlos González de Sande received the M.Sc. degree from Universidad Autónoma de Madrid (UAM), Madrid, Spain, in 1987, and the Ph.D. degree from Universidad Complutense de Madrid (UCM), Madrid, Spain, in 2004, both in applied physics. From 1989 to 1994, he was with the Department of Optics, UCM, as an Assistant Professor. From 1994 to 1996, he was with the Physics Department, Universidad Politécnica de Madrid (UPM), Madrid, Spain, and in 1996 he joined the Circuits and Systems Department, UPM, as an Associate Professor. His research interests include engieneering education, optical physics, optoelectronics, and optical characterization of materials.