

# Mixed e-Assessment: an application of the student-generated question technique

Salvatore Cannella and Elena Ciancimino

Faculty of Engineering  
University of Palermo  
Palermo, 90128, Italy  
{cannella, eciancimino}@unipa.it

Mónica López Campos

School of Engineering  
University of Seville  
Camino de los Descubrimientos s/n. 41092 Seville, Spain  
monlopcam@alum.us.es

**Abstract**— This paper presents the experiment of a mixed summative-formative evaluation in an asynchronous distance higher education context. In the experiment, performed in the distance course of Total Quality Management, Industrial Engineering, University of Guadalajara (Mexico), students are guided to formulate questions on specific topics. Student-generated questions are evaluated through an ad-hoc quantitative tool, specifically designed for the purpose: the four-criterion Observation Matrix. The experiment shows: (1) that it is possible to improve the Higher Education evaluation process and formalise students' skill in a more thoroughly way than with traditional evaluation; (2) how the student-educator interaction in a distance learning environment can be enriched through the mixed assessment

**Keywords:** *mixed formative-summative assessment; distance learning; continuous improvement in higher education; quantitative assessment tool*

## I. INTRODUCTION

Student questioning is a strategy to promote higher-order thinking and to improve learning [1]. Koch and Eckstein [2] showed that students who were taught to generate their own questions achieved more learning outcomes than students who used only teacher's questions. In 1992, King [3] found that training students to generate specific questions and then attempt to answer them is more effective than training in other study techniques. Foos et al. [4] showed that generating potential test questions while preparing for an examination is a very effective technique leading to high performance. Besides, students' questions are a precious source of information on the critical issues in a course [5]. Biddulph et al. [6] and Dori and Herscovitz [7] suggest including question generation in the formal evaluation process. This could be reasonably considered one the aspects of the Darwinian evolution of the learning and teaching paradigm.

Recently, Bergman [8] performed a multistage formative assessment experience realised with Ph.D. students, consisting in formulating "good" questions in areas determined by the examiner [8]. Bergman concluded that this reverse-question evaluation model resulted very interesting for it brought to light learners' "hidden" skills and for the different perception of the evaluation conditions it could lead to by reversing the classical role of questioner and answerer.

Inspired by Bergman's research, Ciancimino and Cannella [9] ran a mixed summative-formative assessment pilot

experiment at University of Palermo (Italy). The experiment consisted in planning and realising a two-stage examination within the Business Process Modelling course of the MSc in Management Engineering. Students had to generate written questions about topics of the course selected by the educator. These questions were subsequently answered by the lecturer and become the core of formalised class debates among students, mediated by the educator. Unlike in Bergman's experiment, in the pilot test at University of Palermo the examiner utilised an assessment tool specifically designed for the purpose: the Observation Matrix. This Observation Matrix allowed the educator to assess the students' questions on the basis of a list of binary criteria. A value of 1 in case of compliance of the criterion or a value of 0 else has to be assigned to each of the criteria established during the experiment design: pertinence to the assigned topic, terminology used to formulate the question, level of complexity, multidisciplinary approach. From the comparison of results of the traditional examination and of the reverse-question evaluation model a number of cases presenting an opposite trend was observed. The experience showed that some of the students that achieved a high mark in the classical exam were low-performers in the reverse-question evaluation. As well, other students had an excellent mark in the reverse-question evaluation but were low-ranked in the classical exam. The pilot experiment provided the educator with a significant feedback, useful for a strength/weakness analysis of the course and for a better organisation of the topics in the course.

This paper presents the experiment of a mixed summative-formative evaluation in an asynchronous distance higher education context. In the experiment, run in the distance course of Total Quality Management, Industrial Engineering, University of Guadalajara (Mexico), students are guided to formulate questions on specific topics. Student-generated questions are evaluated through the ad-hoc quantitative tool, specifically designed for the purpose, namely the four-criterion Observation Matrix [9]. To the authors' knowledge, it is reasonable to consider this experiment the first attempt to evaluate student-generated questions through a distance learning platform.

The experiment shows: (1) that it is possible to improve the Higher Education evaluation process and formalise students' skill in a more thoroughly way than with traditional evaluation;

(2) how the student-educator interaction in a distance learning environment can be enriched through the mixed assessment.

## II. DESIGN OF THE EXPERIMENT

The mixed reverse-question evaluation was developed through a referenced technique for designing and executing examination processes: the Four Processes Architecture [10]. This theoretical framework defines four phases, each of which defines a specific process within the examination, from the learning outcomes selection to the final evaluation of examinees. The phases of Four Process Architecture are summarised in the followings.

(i) Activity Selection: consists in selecting and sequencing tasks to be accomplished by the students. The lecturer of the MSc distance course of Total Quality Management selected three topics from the programme of study: (a) Origins and Evolutions of Quality Management, (b) Statistical Tools for Quality Management, and (c) Total Quality Control. During the six-month course, the students were required to accomplish the routine weekly activities (reports, questionnaires, numerical problems), and monthly verifications. At the same time, each two months the mixed reverse-question model was adopted for the three topics.

(ii) Presentation process: is responsible for presenting the task to the participants. The mixed reverse-question evaluation model was presented to the students during the first month of the course as part of the formal evaluation of the course. Students were required to generate one question for each of the three topics selected in the phase (i) of Almond's framework. On the Moodle platform modality and scheduling of the exam were detailed, including explication of the evaluation criteria and examples. Exemplars of "good" and "bad" questions, chosen so as to illustrate what distinguishes high quality from low [11], were also showed to students, in order to provide a general reference of the level of performance they were required to reach.

(iii) Response processing, has the objective to provide evidence about the participant's current knowledge, skills and abilities. This process was aimed at evaluating the student-generated questions according to the criteria presented in Section 3: pertinence, terminology, level, multidisciplinary. In this phase the educator recollected the questions and performed a preliminary analysis on their characteristics with reference to the identified criteria. At the end of each mixed reverse-question evaluation session, the educator completed a document in the Moodle distance learning platform for each student with a qualitative evaluation of the questions, including a detailed commentary on their performance and advices on how to improve it. As well, in the same document the educator answered the student's question.

(iv) Summary Scoring: consists in converting the qualitative observations into a numerical value representative of the evaluator's beliefs about the participant's knowledge, skills, and abilities. In this phase the educator filled the Observation Matrix for each student (fig. 1).

OBSERVATION MATRIX				
Student name:			Student ID:	
Course:				
Questions	PERTINENCE	TERMINOLOGY	LEVEL	MULTIDISCIPLINARITY
1	—	—	—	—
2	—	—	—	—
3	—	—	—	—

Figure 1: Observation Matrix

The resulting data were gathered in a spreadsheet file and the evaluation of each question was obtained through equation 1. The average value of the assessments of the three questions was the final grade in the mixed reverse-question evaluation.

## III. EVALUATION CRITERIA

In the literature, student-generated questions were categorised and assessed differently, but with a shared emphasis on higher-order thinking skills [12]. Dori and Herscovitz [7] created a quantitative method to obtain a single numerical value, indicative of the level of complexity of students' questions. Marbach-Ad and Sokolove [13] developed empirically an eight-level taxonomy to categorise students' questions for increasing order of thinking. Chin et al. [14] classified students' question during laboratory activities in wonderment questions and basic information questions. In Barak and Rafaeli's research [15], quality of question was assessed both by educators and students: students assigned a value within a numerical scale, teachers evaluated on the basis of the cognitive level required to formulate the question, on the basis of a modification of the Bloom's et al. taxonomy of learning domains [16]. Bergman [8] evaluated students' questions through a pass/fail methodology.

In this work, the evaluation of student-generated questions is structured according to a multi-dimensional quantitative method: the PTLM (Pertinence, Terminology, Level, Multidisciplinary) model. Each question is assessed with the following binary criteria:

(a1) Pertinence: relevancy of the formulated questions to the assigned topic. It is evaluated whether the question has precise and logical connection to the topic.

(a2) Terminology: appropriateness of the words chosen by the student to formulate his questions. Syntax and vocabulary of technical terms used are evaluated.

(a3) Level: extent in approaching the topic. This facet of performance was included to take into account the level of detail of the student's question. It is evaluated whether the question is trivial or if it expresses a significant meaning reflecting the deepening of analysis involved in the reasoning.

(a4) Multidisciplinary: expansion of the range of subject areas included in the question. It evaluates the ability to connect knowledge from other areas to the topic.

For each dimension the examiner assigned a value 0 if the minimum level is not achieved, 1 if the question satisfies the criterion. The result of the evaluation for each question is a value between 0 and 4, successively normalised to 1 (Equation (1)).

$$C_j = \frac{\sum_{i=1}^4 a_i}{\sum_{i=1}^4 i} \quad (1)$$

The final grade for each student in the mixed reverse-question evaluation is computed as the average value of  $C_j$ ,  $j=1\dots3$ .

#### IV. RESULTS AND DISCUSSION

This section is dedicated to present and contrast result from the traditional evaluation and the mixed reverse-question evaluation of the thirty students of the Total Quality course of the MSc in Industrial Engineering at University of Guadalajara (Mexico), academic year 2008.

The first result of the experience was that the degree of interaction between students and educator significantly increased. Moodle sessions increased up to 37% with respect to the previous courses average. Students were enabled to generate questions within a formalized framework, which provided a further stimulus to express their doubts and perplexities. This incentive can be considered particularly significant in a limited interaction context as an asynchronous distance learning environment.

The numerical values resulting from phase (iv) of Almond's framework [10] are reported in Figure 2. Grades of the traditional evaluation and of the reverse-question evaluation are compared. All grades are normalised to 1.

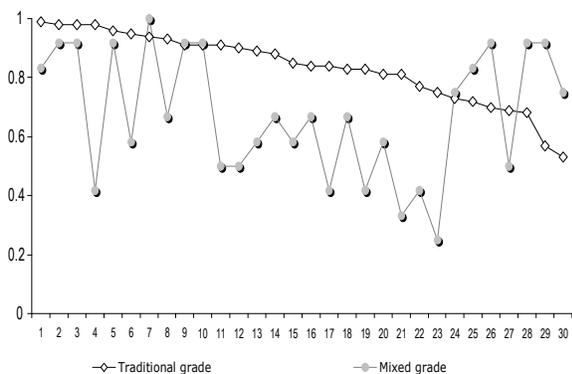


Figure 2: Traditional grade and mixed grade

From Figure 2 it is possible to identify three groups, distinguished by differences in the relative grade on the two evaluations. The first group students (from 1 to 10) performed well both in the traditional and the mixed reverse-question examination, the second group (11-23) performed in the traditional exam better than in the reverse-question, while for the last (24-30) group a better performance in the reverse-question was observed; observation 4, 6, 8 and 27 are exceptions.

In particular, comparing group 2 with group 3, an opposite trend between performance in traditional exam and

performance in the mixed reverse-question examination can be noticed. Besides, the grade in the mixed reverse-question examination of students which are low-performers in the traditional one (group 3) is higher in absolute value than the grade in the mixed reverse-question examination of students from group 2.

Figure 3 shows the percentages of accomplishment of the four criteria of the mixed reverse-question evaluation.

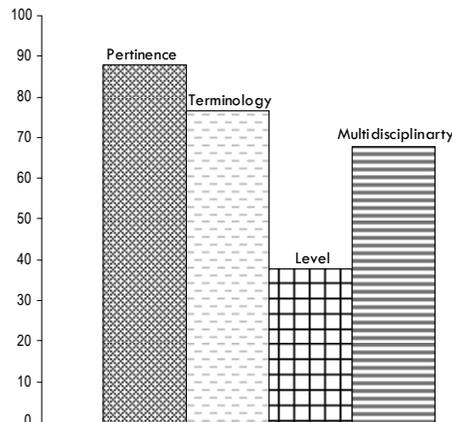


Figure 3: Percentages of accomplishment of the four evaluation criteria

Figure 3 shows that the level criterion was accomplished only by the 38% of questions, the pertinence by 88%, the terminology by 77% and the multidisciplinary criterion by 68%.

Jointly analysing Figure 2 and Figure 3, a significant difference can be noted between the average grade obtained by students in the traditional examination, 83%, and the percentage of accomplishment of the level criterion. The level criterion is assumed to reflect the deepening of analysis and the understanding of the topic. The relatively low average accomplishment of this criterion could be symptomatic of superficial understanding of the course topics by the majority of students. The difference between traditional tasks and student-generated questioning could be indicative that the learning strategies aimed at succeeding in a traditional examination are not always sufficient when an additional effort is required for creative conceptualisation, which is a fundamental component of problem solving capability. From the comparison between traditional exam and the mixed reverse-question evaluation it is possible to hypothesise that the majority of students involved in the experience has had difficulties in converting their knowledge in capacities and competences.

Results from the experience underline that the mixed reverse-question evaluation method provides the educator with a structured feedback tool, which can release information about students' achievements in terms of learning, abilities and capacities. Furthermore, the comparison between traditional grades and mixed reverse-question evaluation

grades are a concrete element for the educator to perform a structured strength/weakness analysis of the course.

From these reflections two main conclusions derive:

(1) Adopting a formative exam and formalising it within a structured framework capable of releasing quantitative information can represent a concrete tool to improve the learning system, as it provides information that is complementary to the classical feedback coming from a traditional exam. The effort for developing, improving and validating innovative tools is a priority for the scientific community. From this perspective, improving the Observation Matrix was reconfirmed as one of the main objectives of the experiment. In this experiment the educator provided a list of suggestions and observations, whose synthesis is presented in Appendix.

(2) The results of the experiment confirm the need for improvement of the higher education system auspicated by the European Community. There is the need to re-think and restructure the student-educator interaction and to enrich the teaching methodologies, in order to balance theoretical knowledge and practical capabilities and competences. Adopting mixed formative-summative evaluation methods can facilitate the conjoint development of the three learning outcomes established by the European Community: knowledge, skills and competences.

#### V. CONCLUSIONS AND FUTURE RESEARCH

This paper presented the experiment of a mixed summative-formative evaluation in an asynchronous distance higher education context. In the experiment, performed in the distance course of Total Quality Management, Industrial Engineering, University of Guadalajara (Mexico), student-generated questions were assessed through an ad-hoc quantitative tool, specifically designed for the purpose: the four-criterion Observation Matrix. The experience showed: (1) that it is possible to improve the higher education evaluation process and formalise students' skill in a more thoroughly way than with traditional evaluation; (2) how the student-educator interaction can be significantly increased through the mixed reverse-question evaluation method.

The limitations of this study also represent opportunities for future research in the formative-summative assessment field. For example, use of language and writing style could be further elements of student-generated question evaluation, regardless the use of technical terms. Furthermore, as answering the students' doubts resulted more interesting and helpful for teacher and students than mere numerical (summative) evaluation of the criterion fulfilment, this could be formally included in the reverse-question evaluation model. Another point regards the improvement of Observation Matrix. It would be desirable to have intermediate score between 0 and 1. By doing so, the evaluation of fulfilment of criteria could be more informative than the pass/fail scoring.

#### ACKNOWLEDGMENT

The authors wish to thank the students that participated to the experiment for their commitment and their precious feedback.

The research was accomplished while the authors were at the University of Seville as visiting researchers and PhD students, funded by the Italian Ministry of University and Research and the Mexican Council of Science and Technology.

#### REFERENCES

- [1] Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education*, 5(1), 7–74.
- [2] Koch, A., & Eckstein, S.G. (1991). Improvement of reading comprehension of physics texts by students' question formulation. *International Journal of Science Education*, 13(4), 473–485.
- [3] King, A. (1992). Facilitating elaborative learning through guided student-generated questioning. *Educational Psychologist*, 27(1), 111–126.
- [4] Foos, P.V., Mora, J.J., & Tkacz, S. (1994). Student Study Techniques and the Generation Effect. *Journal of Educational Psychology* 86(4), 567–576.
- [5] Maskill, R., & Pedrosa De Jesus, H. (1997). Pupils' questions, alternative frameworks and the design of science teaching. *International Journal of Science Education*, 19(7), 781–799.
- [6] Biddulph, F., Symington, & D. Osborne, R. (1986). The place of children's questions in primary science education. *Research in Science and Technological Education*, 4(1), 77–88.
- [7] Dori, Y.J., & Herscovitz, O. (1999). Question-posing capability as an alternative evaluation method: analysis of an environmental case study. *Journal of Research in Science Teaching*, 36(4), 411–430.
- [8] Bergman, B. (2005) Interactive assessment. A new element in the examination process. Proceedings of 8th QMOD, July 2005, Palermo, Italy, 959–966.
- [9] Ciancimino E., & Cannella S. (2008). The Alternative Assessment Experiment. *The International Journal of Learning*, 15(2), 295–300.
- [10] Almond, R., Steinberg, L., & Mislevy, R. (2002). Enhancing the Design and Delivery of Assessment Systems: A Four Process Architecture. *The Journal of Technology, Learning, and Assessment*, 1(5), 1–63.
- [11] Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119–144.
- [12] Keeling, E.L., Polacek, K.M., & Ingram, E.L. (2009). A statistical analysis of student questions in a cell biology laboratory. *CBE Life Sciences Education*, 8(2), 131–139.
- [13] Marbach-Ad, G., & Sokolove, P.G. (2000). Can undergraduate biology students learn to ask higher level questions? *Journal of Research in Science Teaching*, 37(8), 854–870.
- [14] Chin, C. (2001). Learning in science: What do students' questions tell us about their thinking? *Education Journal*, 29(2), 85–103.
- [15] Barak, M., & Rafaeli, S. (2004). On-line question-posing and peer-assessment as means for web-based knowledge sharing in learning. *International Journal of Human Computer Studies*, 61(1), 84–103.
- [16] Bloom, B.S., Engelhart, M.B., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook 1: The Cognitive Domain*. Longmans Green, New York.

#### BIOGRAPHY

SALVATORE CANNELLA is PhD student in Management Engineering at University of Palermo, Italy, and at University of Seville, Spain (co-tutoring agreement).

ELENA CIANCIMINO is PhD student in Management Engineering at University of Palermo, Italy, and at University of Seville, Spain (co-tutoring agreement).

MÓNICA LÓPEZ-CAMPOS is a PhD student at the University of Seville, Spain.

