JOURNAL OF E-LEARNING AND KNOWLEDGE SOCIETY
Vol. 17, No. 2 (2021), pp. 45-55

From traditional exams to closed-ended quizzes: an exploration towards an effective assessment in mathematics at university level

Giovannina Albano[a,1], Agnese Ilaria Telloni[b]

[a] *University of Salerno (Italy)*

[b] *Università Politecnica delle Marche (Italy)*

(submitted: 11/1/2021; accepted: 27/10/2021; published: 22/11/2021)

Abstract

The pandemic emergency has almost forced the transition from face-to-face to remote evaluation. Starting from the results of the research in Mathematics Education, this exploratory work focuses on how to design effective closed-ended questions of different types, capable of reliably assessing mathematical learning outcomes, especially in terms of the involved competencies. We also investigate how to aggregate the questions into Moodle quizzes able to effectively replace the traditional open written exam. We propose a three-dimensional theoretical model, which takes into account the various types of questions, expected learning outcomes, and mathematical arguments, to shed light on the problems of validity, reliability, balance, and correctness of closed-ended quizzes. We discuss the results of the first implementation of the model within a Linear Algebra course for engineering freshmen.

KEYWORDS: Closed-ended Quiz, Assessment, University, Mathematics, Moodle.

DOI
https://doi.org/10.20368/1971-8829/1135433

CITE AS
Albano, G., Telloni, A.I. (2021). From traditional exams to closed-ended quizzes: an exploration towards an effective assessment in mathematics at university level. *Journal of e-Learning and Knowledge Society*, *17*(2), 45-55.
https://doi.org/10.20368/1971-8829/1135433

1. Introduction

During the last term, we were engaged as teachers of a course concerning Linear Algebra, for Computer Science Engineering freshmen of University of Salerno.

Because of the pandemic emergency, we needed to move from face-to-face assessment to distance assessment. This was the occasion to deepen some issues related to computer-based assessment, which have already been treated in some of our previous research. We are familiar with the Moodle platform, which for some years we used to engage students in collaborative activities through the "Workshops", aimed at promoting a critical attitude in the study of mathematics and at fostering a formative (self-) assessment along the course. Moreover, we made different kinds of resources available, such as files containing notes, video, and interactive paths for the development of strategic thinking, tailored to the individual learning needs (Telloni, 2020). We had also exploited the opportunities offered by the activity "Quiz" of the Moodle platform, focusing our attention on the automatic formative (self-) assessment as a tool of learning (Albano & Ferrari, 2008, 2013; Albano, Pierri & Sabena, 2020). The pandemic gave rise to the need of effectively using Quiz as a summative assessment tool. The exams at distance offered us the opportunity to deeply reflect on this design.

In this paper, we focus on how to design an effective Moodle quiz, including different types of questions, that can effectively replace the traditional written open-ended exam. The starting point of our research consists of the following working hypothesis: the validity, reliability, balance, and fairness of the traditional assessment consisting of written open-ended questions, integrating operational knowledge and relational knowledge, and addressing the use of representations in various semiotic systems (Skemp, 1976; Duval, 1996).

We are aware that closed-ended questions have some limitations since they do not enable to fully assess the construction of a text, the design of a problem solving process or the argumentative competency (Ferrari, 2019;

[1] corresponding author - email: galbano@unisa.it

Garuti & Martignone, 2019; Trinchero, 2006). However, we believe that the mentioned weaknesses could be contained and largely overcome through a careful design of the questions and the whole quiz.

According to previous research, the design of questions should take into account that the student, in order to give the correct answer, should activate the desired competencies, such as the focused reading of the text, the modeling of the problematic situation, and the coordination of different semiotic systems (Niss & Hogart, 2019). Moreover, the formulation of the questions should discourage improper strategies, such as ruling out items or recurring external resources. Finally, the decay of questions/items should be taken into account, to avoid choosing items by heart. In this respect, it is necessary to continuously vary the questions, hence, to construct a large database, with an attentive choice of the distractors and the systematic use of the option "none of the other answers", which should be the right option in a significant number of cases. This is aimed to avoid students reaching the correct answers by remembering seen procedures (Darlington, 2014).

This paper focuses on two research questions:

RQ1: how to construct appropriate closed-ended questions able to effectively assess mathematical learning outcomes, according to standards shared by the Academic Community?

RQ2: how to aggregate closed-ended questions in order to construct appropriate quizzes that can replace traditional exams?

This paper is a first exploratory study, referring to the authors' experience during the pandemic, starting from research results in Mathematics Education.

2. Materials and methods

2.1 What should be assessed?

The first issue we addressed was to establish what should be assessed in the specific domain of Linear Algebra. This concerns the *validity* issue, that is the object and the aim of the assessment (Iannone, 2020b). Our perspective aims at the assessment of the competencies, mainly in terms of comprehension and handling of mathematical objects/concepts, their mutual relations, and the processes connected with the problem solving activity. These aspects are deeply taken into account in the traditional exams.

For each item, give appropriate justification.

Let $f\colon \mathrm{R}^4 \to \mathrm{R}^3$ the homomorphism defined by

$$f(x, y, z, t) = (x + 2y - z - 2t, x + y - 2z, -2x - y + 5z + t)$$

a) Find the dimension and a basis B of Ker(f).
b) Establish if f is surjective.
c) Determine if $(1 - 2,0) \in$ Im(f).

Figure 1 - Example of a problem of the traditional exam.

The problem in Figure 1, taken from a traditional exam, addresses some specific educational goals, referring to conceptual knowledge (e.g., definition of the kernel of a homomorphism), calculation skills (e.g. solve a linear system), and mathematical competencies (e.g. mathematical thinking, mathematical problem handling, mathematical communication).

More in general, within the traditional open-ended written exam, the knowledge of several mathematical concepts and calculations skills are addressed, together with some specific mathematical competencies (Niss & Hogart, 2019), such as *mathematical thinking*, *mathematical problem handling,* and *mathematical communication*. It is worth noting that the traditional exam does not call for the *modelling competency* nor the competency about the *use of aids and tools* because they are out of the course's scope and the foreseen exam's modalities.

We started our research by constructing closed questions that covered the same competencies as the traditional exam, giving rise to a list of typical questions. Later, we exploited the identified mathematical competencies in order to create further questions not strictly linked to previous open-ended exams. Indeed, we immediately realized that an exact translation of the traditional written exam into a closed-ended quiz was not possible nor desirable. Instead, the types of questions in the Quiz offered us the opportunity of addressing different aspects of the students' learning, such as their capability of understanding mathematical reasoning, connecting various elements of knowledge, interpreting the meanings of the results of procedures. These features are in line with the suggestions to construct a suitable and effective test for mathematics (Iannone, 2020a), including for instance 'why' questions such as the following.

Does it exist a surjective application $f\colon R^4 \to R^3$ such that

Ker$f = \{(x, y, z, t) \in R^4\colon 2x + 4y + 2z = 4z = 0\}$?

Choose one or more options

□ Yes, because the kernel of f contains only the null vector.

□ Yes, because dim(Ker f)=2.

□ No, because dim(Ker f)=2.

□ No, because Ker f is a vector subspace of R^4 .

Figure 2 - Example of closed "why" question.

The question in Figure 2 and the open question in Figure 1 share knowledge and skills (surjectivity's condition, rank-nullity theorem, and calculation of dimension of the kernel of f), but they differ for the competencies they address. Indeed, the student is asked to connect these elements of knowledge and skills in different settings: in the open question they come into play one at a time in a sequential way, whilst in the closed question they must be recalled and connected at the same time. The question in Figure 2 asks for something more than item b in Figure 1, related to the issue of the existence of a mathematical object (see, for example, Dubinski, 1997): asking if an application is surjective is different than asking if there is at least a surjective application (in the latter case, the mathematical thinking competency comes into play).

More in general, it is worthwhile to note that in the open-ended questions everyone can follow the preferred solving process to reach the answer, whilst the close-ended questions require to be able to identify, follow and evaluate the underlying reasoning of the proposed solutions. This is especially the case of Linear Algebra, where multiple solving strategies are allowed to solve a problem (Newton, Star & Lynch, 2010).

In a nutshell, we had evidence that some technical limitations became opportunities from the didactical point of view.

2.2 How to assess mathematical learning outcomes?

According to our choice to use the Moodle Quiz, also due to institutional constraints, we analysed the characteristics of the available question types in order to identify the better fitting between types and competencies to be assessed.

The available literature about computer-based assessment focuses mainly on the multiple-choice question (see for example Watson et al., 2017), but we intended to exploit all the available types of questions in order to capture different abilities and test the students' behaviors. Inspired by the work of Scalise and Gifford (2006), we made an a priori analysis of the potential of different question types with respect to the devised mathematical learning outcomes to be assessed (knowledge, skills, competencies). This gave rise to the graph in Figure 3, where the horizontal dimension collects the core elements of the identified learning outcomes, and the vertical dimension displays the question types. In particular, the core elements are differently colored according to the learning outcomes they refer to: the blue labels regard the knowledge of terms and concepts, the red labels concern the calculation skills, the green labels are linked to the understanding of terms and concepts, the orange labels concern the problem solving capability and the purple labels refer to the capability of reasoning mathematically and explaining the motivations of procedures. The analysis allowed us to highlight the matching between question types and learning outcomes.

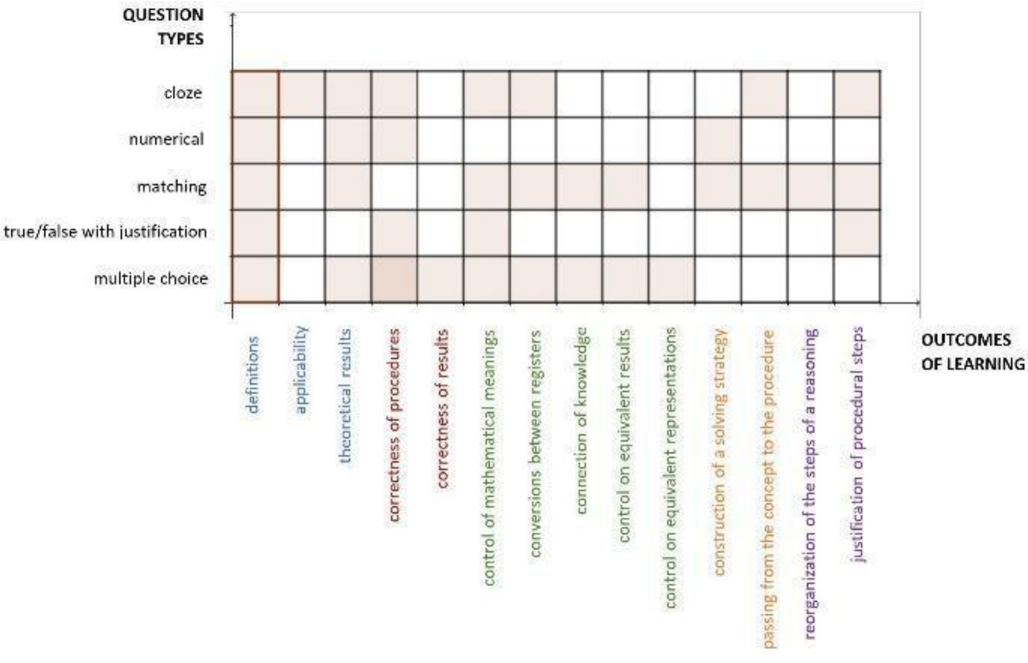

Figure 3 - The matching between question types and learning outcomes.

2.3 How to construct a quiz? The *fairness* and *balance* issues

Some issues about the distribution of the questions with respect to different parameters (mathematical content, didactical goal, question type, time) emerged when we needed to construct the quiz. Moreover, the will of realizing strong randomization of the questions determined a further difficulty.

We immediately grasped the need for a theoretical model, enabling us to simultaneously take into account topics, question types and learning outcomes. So, we extended the graph in Figure 3 by adding a further dimension concerning the topics addressed by the questions, giving rise to the three-dimensional model shown in Figure 4.

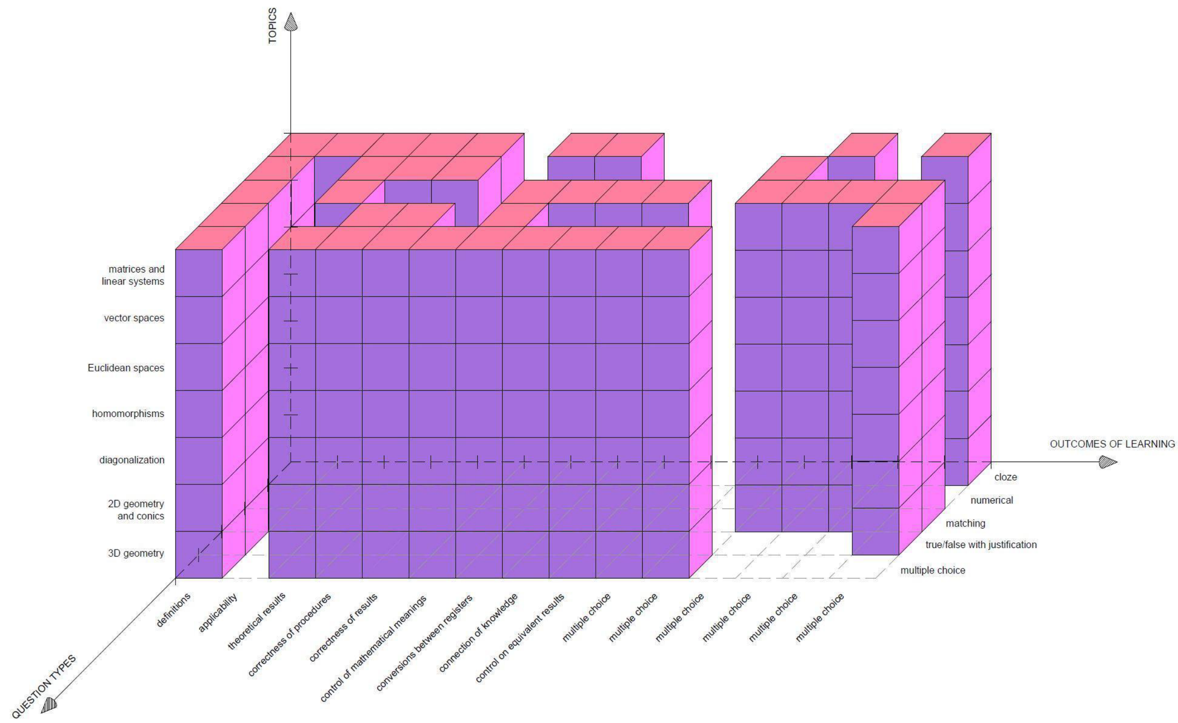

Figure 4 - The three-dimensional model.

We designed a mixed quiz, including open-ended questions and closed questions, in order to monitor the *reliability* of the quiz, that is the outcomes of the assessment in terms of grading (Iannone, 2020b). This was done according to two different aspects: on the one hand, we monitored the reliability of the whole quiz as a tool of assessment and, on the other hand, we looked at the reliability of the closed-ended questions. This is in tune with the split-half method (Chakrabarrty, 2013). Indeed, based on our hypothesis about the reliability of the traditional open-ended exam, we could a posteriori gain information about the reliability of the whole quiz as well as of the closed questions, by comparing the average scores. The length of the quiz and the wide spectrum of the addressed topics constitute further features of the quiz which foster the reliability (Livingston, 2018).

The open-ended questions in the quiz are similar to those of the traditional exam, even if the communication of the mathematical contents is more challenging for students. Actually, we chose to set the technical constraint of enabling only textual answers: this way, we could assess the students' capability of justifying the problem solving process in verbal language, either from the theoretical side and the procedural side. More in general, these kinds of questions allow us to assess the conceptual engagement of students with mathematical topics by means of their communications skills (Iannone, 2020a). This aspect is in tune with the definition of *competence*, including the capability of using appropriate linguistic resources with respect to specific functions and aims, as suggested by the Council of the European Union (https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018H0604(01)&from=EN).

In constructing the quiz, we needed to take into account technical constraints and didactical needs: indeed, by using the Moodle Quiz, we could not simply randomize along all the three dimensions of the model in Figure 4 (learning outcomes, question types, topics). This induced us to create equivalence classes of quizzes. So, considering the expected number of students for the exam, we created 5 templates of the quiz, so that around 20 students were delivered the same template. All of them share the same structure, where one of the dimensions of the model, that is the question type, is fixed as follows: every quiz contains 2 questions with numerical answers, 3 true-false with justification questions (one is about theoretical issues), 3 multiple-choice questions (one is about theoretical issues), 3 matching questions and 1 cloze question. This choice of giving more importance to the dimension of the model related to the question type is linked to the fairness of the quiz and to the time needed to perform the quiz: we considered that in order to create "equal" quizzes, we needed to include the same number of questions of a specific type. In each template of the quiz, we chose the value of the further dimensions of the model to be set for each question type, i.e., the topic and the learning outcome. So, for example, the 2 numerical questions of our quiz 1 concern vector spaces and address the knowledge of terms/concepts; on the other hand, the 2 numerical questions of our quiz 4 concern matrix and linear systems and address the calculation abilities. Table 1 shows an example of a template.

TEMPLATE 1	TOPIC	LEARNING OUTCOME
NUMERICAL	Vector spaces	Conceptual Knowledge
TRUE/FALSE WITH JUSTIFICATION	Matrices and linear systems	Calculation skills
MULTIPLE CHOICE	Homomorphism	Understanding of conceptual knowledge
MATCHING	2D geometry and conics	Problem-solving capabilities
CLOZE	Diagonalization	Understanding problem-solving process

Table 1 - The matching between question types and learning outcomes.

Corresponding to each cell of the tridimensional model, we created a large database of different questions. This choice underlies the idea of considering to be equivalent the questions associated with the same cell.

2.4 Sample of questions

In this section, we provide some examples of significant closed-ended questions of different types.

Let us start with a matching question (Figure 5), focused on 3D geometry and addressing problem-solving capabilities.

Consider, in the Euclidean space, the Cartesian equation of the plane *ax+by+cz+d*=0. Associate the mathematical conditions corresponding to each given case:

1. Plane which is perpendicular to *x+y+z*=1 and parallel to the *z*-axis.
2. Plan containing the origin of the axes and perpendicular to *2x+2y+2z-1*=0.
3. Plane containing the point (0,0,1).
4. Plane containing the origin of the axes and parallel to the *z*-axis.

a = d = 0
a + b = 0
c + d = 0
a + b + c = d = 0

Figure 5 - A matching question on 3D geometry.

In order to successfully answer the question, the student should proceed in a reverse way with respect to classical open-ended questions addressing the same knowledge and calculation skills. In fact, the student should not construct a plane according to some given conditions (parallel to a plane, orthogonal to a line, containing a point, etc.), but he should recognize suitable characteristics on the coefficients of the equation of a generic plane, corresponding to the given conditions. The correct performance depends on a connected knowledge and good control of the geometric and algebraic languages.

Also Figure 6 shows a question about 3D geometry. It is a cloze question, where the items to be chosen are in square brackets, addressing the understanding problem-solving process.

In the Euclidean 3-dimensional space, the linear system $\begin{cases} x+3y-8=0 \\ 2y+z-7=0, \end{cases}$

represents [a line/a plane], having as [normal/parallel] vector equal to v=[(3,-1,-2)/(-3,-1,2)/(-3,1-2)]. The equations of the linear system represent [orthogonal lines/parallel lines/parallel planes/orthogonal planes], respectively, to the planes z=0 and [y=0/x=0]. Moreover, the locus $\begin{cases} x=-1-6t, \\ y=2+2t, \\ z=-4t, \end{cases}$ represented by the system [intersects/ is parallel to] the locus represented by since the vector of it, u, [is/is not] parallel to v.

Figure 6 - A cloze question on 3D geometry.

In order to give the correct answers, students should have some knowledge of definitions (i.e., normal/parallel vector) and calculation skills (parallel vector of a line); moreover, they should be able to switch between different semiotic registers.

The close question in Figure 7 concerns diagonalization and addresses understanding of conceptual knowledge.

Consider the real matrix A= $\begin{pmatrix} 0 & 1 & -4 \\ 1 & 2 & 1 \\ -1 & 1 & 0 \end{pmatrix}$. Since A is [*symmetric/not symmetric*], then it [*is/is not*] orthogonally diagonalizable. Hence it is [*possible/not possible*] to deduce that A is also [*diagonalizable/not diagonalizable*]. In this case, it is [*possible/not possible*] to calculate the matrix P which orthogonally diagonalizes A by identifying a [*orthonormal basis/basis*] of all the eigenspaces of A. The matrix P can be constructed by putting as [*columns/rows*] the vectors of the obtained bases and P turns out to be [*an orthogonal/symmetric/diagonal*] matrix.

Figure 7 - A cloze question on the diagonalization of matrices.

The question investigated if the student grasped the relation between diagonalization and orthogonal diagonalization, as well as the conceptual existence of the diagonalization matrix P (not only the procedure to construct it).

The following two questions address conceptual knowledge, that is students' understanding of theorems.

The multiple-choice question in Figure 8 proposes items concerning the statement and the proof of Cramer's Theorem. Items 2 and 5 require the student to have gone into depth and to be able to explain the steps of the proof in detail.

Consider Cramer's Theorem. Choose the true items (one or more options):

□ The hypotheses of the Theorem include a linear system with complete matrix of order n

□ The proof exploits the definition of inverse matrix

□ In the thesis of the Theorem the *n* solutions of the linear system are provided

□ None of the other items

□ In the proof the property of the identity matrix of being the identity element for the matrix multiplication is exploited

Figure 8 - A multiple-choice question on Cramer's Theorem.

The matching question in Figure 9 asks the student to identify which items concern hypotheses and which ones concern the thesis of Steinitz's Lemma.

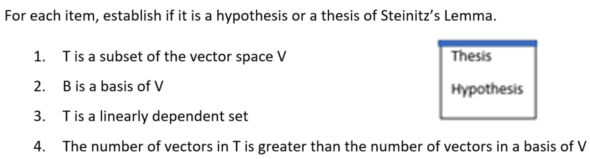

Figure 9 - A matching question on Steinitz's Lemma.

Usually students learn by heart theorems' statements (and often also their proof). This kind of question can affect students' learning style, forcing them to distinguish between hypotheses and thesis, and to be aware of the objects and of the meaning of variables at stake. For instance, item 4 refers to the relationship 'm>n' appearing in the statement, which is crucial.

The last two examples of questions go beyond the traditional written exams. This kind of issue is usually investigated during the traditional oral exam, but often this level of deepness is neglected, especially in the case of large numbers of students.

2.5 Some implementation details

The questions and the quiz have been implemented using the Quiz module in the platform Moodle.

We exploited some of the standard closed-ended question types: numerical, matching, cloze, and multiple-choice. This last type has been customized as follows:

a) all the questions appear as multiple-answer questions, both in the case of a single correct answer and in the case of several correct answers. This has been a didactical choice that forces the students to assess each item and aims to avoid them use some inappropriate strategies;
b) a new category, called 'true/false with justification', has been added. It is a multiple-choice question, foreseeing four answers like 'true because.../false because...'; also in this case there can be more than one correct answer, differing only in the justification. So, identifying the correct answers means to identify the correct pairs (true/false, justification);
c) the two previous categories have been replicated as new categories, specifying that they concern exclusively the assessment of comprehension of theoretical issues (definition, theorem statements and proofs, theoretical characterizations of a concept from different viewpoints).

We also used the "essay" question type for implementing the open-ended questions.

Moreover, since Moodle allows the teacher to add tags to a question, we exploited this facility in order to associate each created question with a topic and a learning outcome. So we created:

a) the tags corresponding to the macro-topics of the course: Matrices and linear systems, vector spaces, Euclidean spaces, homomorphism, eigenvalues and eigenvectors, 2D geometry and conics, 3D geometry;
b) the tags corresponding to the learning outcomes (see Figure 3).

So two tags have been added to each closed-ended question. Concerning the open-ended ones, we chose to split them into two classes of equivalences, by means of the two tags: algebra and geometry, depending on the content.

The setting up of each quiz has been guided by the template for what concerns the closed-ended part. Random questions have been added, filtering the question bank by tag.

Concerning the quiz layout, the order in which questions appear to the student as well as the order in which the items of the answers to a question appear have been also randomized.

Finally, a lockdown browser has been integrated in Moodle and it is activated when students access the quiz.

3. Results

In this section, we report the results of the first exams' session, just at the end of the course, which was performed by 96 students. The quiz was composed of 13 closed-ended questions and 2 open-ended questions. The best mark for each closed-question has been set to 1, whilst the score of each open-ended question was set to 8.5. In order to pass the exam, it is not sufficient to perform correctly in only one of the two parts of the quiz (close and open). Taking into account the issue of cheating, the team of the department teachers agreed to give around 4 minutes for each closed-ended question and around 20 minutes for each open-ended question, for a total of 90 minutes.

First of all, we will give details about the reliability of the quiz. We will handle the issue at two different levels: comparison with the traditional exams and reliability of the closed-ended questions.

Concerning the first level, we looked at the same session of the last year and we note that the overall percentage of successful students in this first attempt is slightly higher.

We notice that all and only the students who get at least 60% of the maximum open-ended part score, get a global sufficient score as well. This suggests that the mixed exam and the traditional open-ended exam have similar discrimination potential, as desired.

Concerning the second level, we compared the results of the closed-ended part and open-ended part within the quiz. We observed that the scores of the former are generally slightly better than those of the latter. We compared the percentage of the score obtained in the closed-ended questions with respect to 13 (the maximum score for this part of the quiz), and the percentage of the score obtained in the open-ended questions with respect to 17 (the maximum score for this part). The mean value of the former is 64,8%, whilst the mean value of the latter is 60,4 %, with a gap of 4,4% in favour of the closed-ended part. Going more in depth, focusing on the students who get a sufficient score only to the closed-ended questions, we note that they did not pass the quiz and they are 16% of the total students.

In the following, in order to deepen the analysis, we will focus on the data of quiz 1.

From the quantitative point of view, the diagram in Figure 10 shows the percentage of the score obtained in quiz 1, with respect to total questions (blu bar), closed-ended (orange bar) and open-ended questions (grey bar). The horizontal green line indicates the sufficient score.

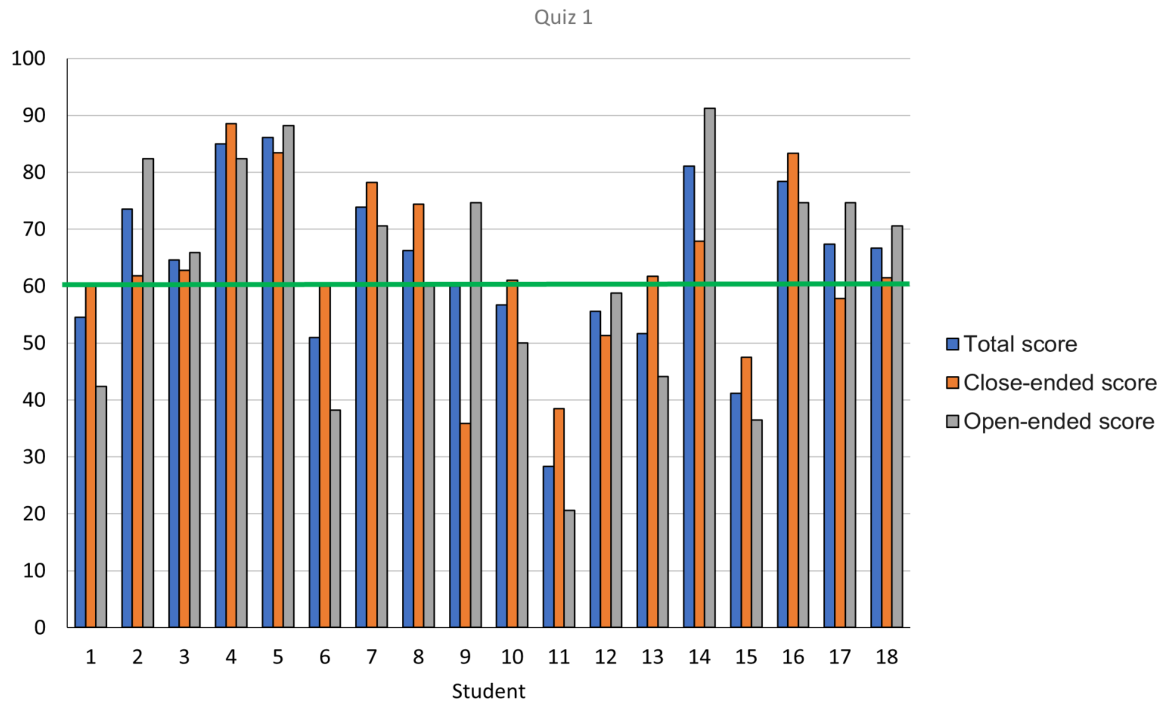

Figure 10 - Data of quiz 1 (whole quiz, closed-ended part, open-ended part).

Taking advantage of the statistics report provided by Moodle, we focus on the facility index, which is the percentage of the students that answered the question correctly[2]. The average facility index of the closed-ended questions is around 62%, whilst the one referred to the open-ended questions is 57%. This suggests that the two types of questions are comparable (as desired). The facility index of the closed-ended questions varies from 31,6% to 93,9%, whilst on the open-ended ones from 44,6% to 69,5%. Although some closed-ended questions require higher competencies (such as connections, as stated previously), it is not surprising that the open-ended questions appear more difficult on average, since the student is required to compose from blank, without any given clue.

There are 11 students who passed the exam.

It is worth noting that three students got the sufficiency to the closed-ended questions, but they did not to the open-ended ones neither to the total quiz. From now we refer to them as 'critical students'.

To deepen how to improve the closed-ended questions, we go into qualitative details of the questions which give evidence of students' difficulties, paying particular attention to the three critical students.

The questions can be grouped according to the kind of learning outcomes they address:

a) *questions concerning the meaning of the rank of a matrix with respect to the context of use*.

Let us see the question in Figure 11. Its facility index is 50%, but all the three critical students selected only wrong items.

[2] https://docs.moodle.org/310/en/Quiz_statistics_report

Let us consider the homomorphism represented by the matrix

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \\ 3 & -3 & 1 \\ 4 & -1 & 5 \end{pmatrix}$$

Choose the correct items:

□ The rows of A are linearly dependent

□ The domain and the codomain of the homomorphism have the same dimension

□ The dimension of Im f is at most 4

□ The homomorphism is surjective

□ The columns of A form a basis of Im f

Figure 11 - A question concerning the rank of a matrix.

From a procedural point of view, the question involves the computation of the rank of the matrix A, but choosing the correct items requires connecting the concept of rank with various other concepts in different fields, such as the linear dependence/independence of vectors, the basis of a vector space, the kernel and the image of a homomorphism, a matrix representation of a homomorphism.

Similarly, the question in Figure 12 has a facility index equal to 53,3%, and two of the critical students selected some correct items and some incorrect ones, whilst the other critical student did not answer at all. The question concerns the calculation of a matrix representation and its rank; therefore, most of the concepts involved are the same as the previous question. However, the two items require different connections to be answered.

Let us consider the homomorphism defined by

$f(x, y, z, t) = (2x - y, x + y - 2z, z + t)$.

Choose the correct items:

□ f is injective

□ The matrix associated to f has 3 rows

□ Im f always can be represented through Cartesian equations

□ The orthogonal subspace of Ker f has the same dimension of Im f

□ The matrix associated to f is square

Figure 12 - A question concerning homomorphisms.

b) *theoretical questions, which can be split into*:

i) questions requiring the concepts' knowledge.

Answering the question in Figure 13 just requires knowing the definition of the counterimage of a vector in a homomorphism. This question has the lowest facility index. It is worth noting that all the three critical students skipped this question, as well as most of the other students. This can be ascribed to two facts: 1) the presence of a parameter (even if it appears to describe a set of infinity elements); 2) a partial interpretation of the symbol f^{-1}. The latter seems to depend on the students' difficulty in distinguishing, managing, and coordinating the meanings of inverse function and counterimage.

Let f be an endomorphism of R^2 such that

$f^{-1}(1,2) = (k + 1, 3 - k), k \in R$.

Choose the correct items:

□ dim(Ker$f$)+dim(Imf)=4

□ $(1,3) \in$ Ker$f$

□ $(1,2) \in$ Imf

□ f is an isomorphism

□ Ker$f$={0}

Figure 13 - A question requiring the concepts' knowledge.

Figure 14 shows a question that has often been left blank and whose facility index is equal to 58,33%. The question is about the definition of eigenvectors, although some items can be analyzed independently (e.g. item 4).

Let $v_1 = (1,1,0), v_2 = (0,1,1)$ and $v_3 = (0,0,1)$ be eigenvectors of the homomorphism $f: R^3 \to R^3$ associated to the eigenvalues 1, 1, 2, respectively.

Choose the correct items:

□ $f(v_1) \neq v_1$

□ $f(v_1)$ is the null vector

□ $\forall\, v' \in R^3, f(v') = (x, y, x - y + 2z)$

□ v_1, v_2 and v_3 form a basis of R^3

□ It is not possible to calculate the linear extension $f(x, y, z)$

□ None of the other options

Figure 14 - A question about the definition of eigenvector.

The outcomes of both the above two questions suggest the students' difficulties in managing definitions, which seems to be consistent with a rote and procedural approach to Linear Algebra.

ii) questions requiring the knowledge of theorems.

The question in Figure 15 is based on the theorem about the existence and uniqueness of solutions of a linear system. It concerns the understanding of how the constraints of existence and uniqueness are related to the rank of the matrices associated with the system, rather than to their dimensions. The facility index of the question is equal to 60%. Two of the critical students failed and one skipped the question. The difficulty can be ascribed to the control of theoretical issues: answering without performing calculations often constitutes an obstacle for students preferring procedural approaches.

Given a solvable linear system of *m* equations and *n* variables (*m*≠*n*), does it always have more than one solution?

□ Yes, since the coefficient matrix is not square.

□ No, since the rank of the coefficient matrix and the rank of the complete matrix could be the same and equal to n.

□ Yes, since *m* and *n* are different.

□ No, because, by hypothesis, it follows that the number of variables is equal to the rank of the coefficient matrix and to the rank of the complete matrix.

Figure 15 - A question about the existence and uniqueness of solutions of a linear system.

c) *questions requiring the management of parameters, connected to the issue of quantifiers*.

The question in fig 16. concerns a square linear system, with non-constant coefficients, depending on the real parameter k. It is clear that the system has a unique solution for all k real, except a finite number of values (corresponding to the zeroes of the determinant of the coefficient matrix). The difficulty of this question can be ascribed to the fact that assessing the correctness of the various proposed items requires managing the relationship between the given values and the universal quantifiers. We underline that the question in Figure 16 foresees two correct answers, corresponding to the first and the third items. The facility index is equal to 33,4% while the percentage of the students who gave the correct answers is 22%, as well as the percentage of the students who skipped the question. All the other students selected only one item, no matter if it was correct or not.

This may suggest the doubt that they assumed the uniqueness of the correct answer.

4. Discussion and conclusion

In this paper we address the problem of moving from traditional assessment by written open-ended exams to computer-based assessment by closed-ended quizzes, using Moodle. The choice of the platform has been due also to institutional constraints. While being aware of the limitations of a closed-ended assessment, we faced the issue of constructing questions and assembling quizzes which assessed the same competencies as a traditional open-ended exam. We are also aware of the existence of other computer-aided assessment tools based on computer algebra systems (e.g. STACK, see Sangwin, 2013), which offer the possibility of handling and recognizing (equivalent) algebraic expressions. However, assuming the importance of verbal language in mathematics learning (Ferrari, 2020), we took the perspective of fully exploiting all available tools. In this respect, we chose to require students to answer in verbal language to the open-ended questions to force their production of written mathematics.

Let us consider the linear system

$$\begin{cases} x + kz = \ 1 - k \\ y + kz = \qquad 0, \quad k \in R. \\ x + ky + 2z = 0 \end{cases}$$

Does it always have a unique solution?

Choose one or more options.

□ No, because for $k = -2$ the rank of the coefficient matrix is not equal to the number of variables.

□ Yes, because the determinant of the coefficient matrix does not depend on k.

□ No, because for $k = -2$ the rank of the coefficient matrix is different from the rank of the complete matrix.

□ Yes, because for $k = 1$ the rank of the coefficient matrix is equal to the number of variables.

Figure 16 - A question concerning a linear system depending on a real parameter.

Various issues emerged, concerning the design of both significant questions and the whole quiz. The change of assessment posed some issues. One of the most relevant is about the possible discrepancy between the way of teaching and the way of assessing. To prevent this, we started with an a priori analysis of the traditional exam. It allowed us to highlight what learning outcomes had been assessed, that means what the students were used to focus on for the assessment, according to the course's teaching style. Further a priori analyses have been devoted to the learning outcomes assessed by the traditional oral exam. They mainly concerned the students' capability of connecting various pieces of knowledge, managing different meanings of the same concept with respect to different contexts, recognizing the theoretical underpinnings to operational procedures, investigating definition or proof comprehension of theoretical results. The analysis made was related to the potential of the various question types offered by Moodle quizzes. This has brought to the emergence of a graph (Figure 3), that highlights the matching between learning outcomes and question types. We implemented the graph, associating each question to a cell of the graph, by means of the Moodle question types and the tagging facility (using learning outcomes as tags). A further tag has been added to the question, specifying the topic the question refers to. The use of the pair of tags (learning outcomes, topic) within a question type gives rise to equivalence classes of questions.

Besides the questions, we also addressed the issue of building the whole quiz, which is a critical issue, in our opinion, not sufficiently addressed in the current literature. In this phase we needed to further develop the graph and design the three-dimensional model in Figure 4. It takes into account the mathematical topics, in addition to the learning outcomes and the question types. This brought into play the introduction of equivalence classes on the quizzes. We chose to define two quizzes as equivalent if they contain the same number of different question types and aim to assess globally the same set of learning outcomes, whilst the topics can vary. In a few words, two quizzes are equivalent if they correspond to the same block of the three-dimensional model.

The results of the first round of exams show that the students have not been affected by the new assessment method, and this seems to confirm the consistency between teaching and assessment. This is true also in the case of questions which were negatively evaluated, such as the example in Figure 16, requiring the handling of parameters and quantifiers. The same type of difficulty was also found in open-ended questions, such as the question in Figure 17.

Solve in a clear, complete and efficient way the following task, explaining step by step the reasoning and justifying it with appropriate theoretical definitions or results.

Let $f: R^3 \to R^3$ be the homomorphism defined by

$$f(x,y,z) = (x + hy - z, -y + hz, hx - y + z), \qquad h \in R.$$

a) Find the dimension and a basis of $(Imf)^{\perp}$ when f is not injective.
b) Calculate, when it is possible, the inverse matrix of the matrix A associated to f with respect to the canonical bases.
c) State the theorem of invertibility of matrices and explain the necessity of the invertibility condition.

Figure 17 - An open-ended question involving a real parameter.

Many students did not correctly solve item b), requiring the handling of the parameter in its generality, and they computed the inverse of A by assigning a fixed value to h.

A further remark concerns the issue of traditional closed-book setting of exams, difficult to realize in the current situation of remote assessment, due to the pandemic (Iannone, 2020a). In our setting, students were required to perform the quiz using a proctoring software, being controlled through a camera. They were also prevented from consulting any knowledge source (books, notes, etc). However, we expected that the questions we designed, requiring connections between different elements of knowledge, changes of semiotic registers and the control on mathematical meanings, could limit the danger of cheating. In this respect, during this term, we are using closed-ended quizzes as informal in-itinere tests, whose positive mark gives the student some 'bonus' for the final exam. The students perform these quizzes at home, without any control, in an 'open-book' setting. The obtained outcomes are comparable to the ones obtained in the 'close-book' setting, shown in this paper. This shed light on the fact that the well-design of the questions can bring us towards open-book exams. Actually, the quizzes submitted during this term contain new questions, which have been designed based on the on-going analysis of the previous exams' outcomes.

What comes out from this first exploratory study suggest further research questions to be investigated, such as the following:

- looking at the students who had better mark to closed-ended questions than open-ended ones, if and how to define students' profile in mathematics;
- how to exploit the results of closed-ended questions in order to define a hierarchy of question types, learning outcomes and topics according to the students' difficulties made evident from their success or not in correctly answering;
- if and how to use only close-ended questions, especially for what concerns communication skills and activation of linguistic resources;
- how to exploit the item analysis of previous quizzes (in terms of facility index of the single question) in order to a priori evaluate the difficulty level of the quiz and construct fair quizzes;
- carrying out a deepen analysis at two levels (ongoing): at question level, concerning their facility/difficulty and their discrimination potential; at quiz level, concerning the reliability. The analysis of these aspects requires a much bigger question bank as well as much more data coming from quiz sessions.

References

Albano, G., and Ferrari, P.L. (2008), Integrating Technology and Research in Mathematics Education: The Case of E-Learning, In F.J. García-Peñalvo (Ed.), Advances in E-Learning: Experiences and Methodologies (pp.132-148). Hershey, NY: InformationScienceReference.

Albano, G. and Ferrari, P.L. (2013), Mathematics education and e-learning: meaningful use of the quiz-modules, In E. Faggiano & A. Montone (Eds.). ICTMT 11 Conference Proceedings (pp.53-58).

Albano G., Pi,erri, A. & Sabena, C. (2020), Enhancing formative assessment practices in undergraduate courses by means of online workshops, In: Proceedings of the 14th International Conference on Technology in Mathematics Teaching – ICTMT 14 DuEPublico, Duisburg-Essen Publications Online (pp. 155-162).

Chakrabarrty, S.N. (2013) Best Split – half and Maximum Reliability, Journal of Research & Method in Education 3 (1), 1-8.

Darlington, E. (2014), Contrasts in mathematical challenges in A-level Mathematics and Further Mathematics, and undergraduate mathematics examinations. Teaching Mathematics and its Applications, 3, 213–229.

Dubinsky, E. (1997), On Learning Quantification, Journal of Computers in Mathematics and Science Teaching. 16 (213), 335-362.

Duval, R. (2006), A cognitive analysis of problems of comprehension in a learning of mathematics, Educational Studies in Mathematics, 61, 103-131.

Ferrari, P.L. (2019), Argomentare a scuola: intrecci fra matematica e lingua, L'insegnamento della matematica e delle scienze integrate, vol.42 A-B, n.5, 611-625.

Ferrari, P.L. (2020). Educazione matematica, lingua, linguaggi. Costruire, condividere e comunicare matematica in classe. UTET Università.

Garuti, R. & Martignone, F. (2019), Assessment and argumentation: an analysis of mathematics standardized items, In: U.T. Jankvist, M. Van den Heuvel-Panhuizen & M. Veldhuis (Eds.). Proceedings of CERME 11 (pp. 4075-4082) Utrecht: Freudenthal Group & Institute, Utrecht University and ERME.

Iannone, P. (2020a), Assessing Mathematics at University: Covid-19 and beyond. London Mathematical Society, Newsletter, Issue 490, september 2020, 34-40.

Iannone, P. (2020b), Assessment of mathematics in the digital age: the case of university mathematics. In: A. Donevska-Todorova, E. Faggiano, J. Trgalova, Z. Lavicza, R. Weinhandl, et al. (Eds.), Proceedings of the Tenth ERME Topic Conference (ETC 10) on Mathematics Education in the Digital Age (MEDA).

Livingston, S. A. (2018). Test reliability—Basic concepts (Research Memorandum No. RM-18-01). Princeton, NJ: Educational Testing Service.

Niss, M., and Højgaard, T. (2019), Mathematical competencies revisited, Educational Studies in Mathematics, 102, 9–28.

Sangwing, C. (2013), Computer Aided Assessment of Mathematics, Oxford University Press.

Scalise, K. and Gifford, B. (2006), Computer-Based Assessment in E-Learning: A Framework for Constructing “Intermediate Constraint” Questions and Tasks for Technology Platforms, Journal of Technology, Learning, and Assessment, 4.

Skemp, R. (1976), Relational understanding and instrumental understanding. Mathematics Teaching, 77, 20–26.

Telloni, A.I. (2021), Design of individualized digital activities fostering strategic planning in linear algebra, In: G. Mele, P. Magnaghi-Delfino & T. Norando (Eds.) Faces of geometry II. Lectures Notes in Networks and Systems, 172, pp. 287-298. Springer International Publishing.

Trinchero, R. (2006), Valutare l’apprendimento nell’e-learning, Erickson.

Watson, K., Wawro, M., Zandieh, M & Kerrigan, S. (2017), Knowledge about student understanding of eigentheory: Information gained from multiple choice extended assessment, Proceedings of the 20th Annual Conference on Research in Undergraduate Mathematics Education, pp. 311-325.