Problem-Solving Multiple-Response Tests: Guessing is not a Favourable Strategy

Panos Photopoulos and Dimos Triantis
Department of Electrical and Electronic Engineering, University of West Attica, Athens, Greece
Email: {pphotopoulos, triantis}@uniwa.gr

***Abstract*—This publication proposes an objective examination format for assessing students' ability to solve problems in Analogue Electronics. It suggests that a problem asking more than one question can take the form of a Problem-Based Multiple-Response (PBMR) item. Depending on the inter-relationship between the questions addressed in the stem, this paper identifies four types of items and suggests a scoring rule for each one of them. Issues related to the time slot given per item are also discussed. Examination results show that both the facility (F.I) and the discrimination (D.I) indices improve, and much guesswork is discarded whenever the scoring of an item considers the inter-relationship between the questions addressed in the stem. PBMR items addressing two or more fully inter-related questions are more suitable for testing the students' ability to solve problems when an objective examination is desirable. Such problem items give a more realistic picture of the student's actual knowledge.**

***Index Terms*—objective test, electronics, problem-solving, guessing, reliability, discrimination index**

I. Introduction

In Physics and Engineering education, students learn concepts, theories, and models and, more importantly, how to use them to solve problems [1]. Teachers use problem-solving as a tool to teach and assess students' learning [2]-[5]. Traditionally, the final written exams assess students' learning asking them to solve a set of problems. However, budgetary cuts imposed on public universities after 2010 and the increasing number of students [6] make universities focus on cost-effective solutions [7]. Computer-assisted objective tests are an efficient way to assess students' learning, replacing constructed-response tests. The recent health crisis increased the interest in e-learning and innovative e-assessment [8]-[10], making the design of reliable computer-assisted problem-based objective tests a priority. Although single-answer multiple-choice tests effectively assess knowledge and comprehension [11], it is debatable whether they can gauge students' problem-solving abilities [9], [10], [12], [13]. In single answer multiple-choice exams designed to assess the students' abilities to solve problems, the stem gives a linguistic representation of the problem, usually accompanied by a figure. The students solve the problem on paper and then select the correct answer among the given options. However, in the case of a five-option multiple-choice item, the student has a 20% chance to get full marks by pure guessing. Partial but genuine knowledge demonstrated during the solution is not communicated to the examiner and, therefore, not rewarded [6]. At the same time, miscalculations can be disastrous [9]. Both the examiner and the examinee are losing control [6], [12]. Negative marking discourages guessing, and it is a commonly used intervention to improve tests' reliability and validity. However, negative marking disadvantages risk-averse students and introduces a significant bias against female exam takers [14]. For these reasons, some researchers have proposed assessment methods that avoid this penalty [3], [15].

Assessments must be reliable and valid [16], [17], i.e. the scores must reflect the examinees' knowledge. In multiple-choice tests, reliability improves by increasing the number and the quality of the items [17]. When the number of items in a test is large, the measurement error due to guessing is small. Still, objective tests assessing students' ability to solve problems demand much writing besides extensive and complicated thinking [18], [19] before selecting the answer. Consequently, the necessary time slot per item is much greater than standard multiple-choice tests [6]. Because of time limits and fatigue, a problem test must include a limited number of items; therefore, attention must focus on improving item quality and scoring. This publication discusses various types of objective test items for assessing students' ability to solve problems in a first-year course in Analogue Electronics. Each item represents a problem addressing two or more questions. More sophisticated question types require more complicated scoring rules that allow students to obtain marks for partial solutions. Looking into the inter-relationship between the questions addressed in each problem, we identified four scoring alternatives. The scoring rules avoid negative marking, reward partial knowledge and identify answers given by guessing. Short problem-based multiple-response tests can effectively discard the influence of the lack factor on exam scores and make guessing a non-favourable strategy. PBMR items require breaking the problem into two or more questions and identifying the inter-relationship between the questions. Carefully designed items make the reward of partial knowledge attainable because they identify guessing and give marks for answering less than all the questions.

Manuscript received October 25, 2021; revised January 18, 2022.

doi: 10.18178/ijlt.8.1.64-70

II. Materials and Methods

A. Problem-Based Multiple-Response Tests

Problems in Analogue Electronics describe the operating conditions of circuits and ask the solver to calculate one or several physical quantities, usually currents and voltages. The solution requires the translation from the initially given representation to intermediate ones and then to a final representation, leading to the calculation of the unknown quantities [20], [21]. Problem-solving in first-year Analogue Electronics also requires the application of appropriate approximate models [22], which will give results close enough to the actual behaviour of the given circuit [23], [24]. The perplexity of the solution allows the examiner to break a single question statement into many sub-questions and track the solver's effort. However, the coherence of the physical theories underpinning the applied models implies an internal inter-relationship between the problem questions.

B. Scoring of PBMR Items

The following sections describe four possible inter-relationships between the questions stated in a problem item. Each inter-relationship requires a different scoring rule to ensure that students' grades accurately reflect their actual knowledge.

1) A set of simultaneous questions

Consider the case where an objective test item asks the students to find the solution of a 2x2 system of linear equations. The item includes four options for the first unknown and an equal number of options for the second. However, there is a 1/16 probability for a student to select the correct pair of numbers by pure guessing. The students who select the correct answer for the x value but a wrong answer for y should get half the marks or nothing?

In this problem, knowledge of Cramer's rule allows the simultaneous calculation of the two unknown numbers. Calculating y does not require some extra knowledge compared to x. Adopting a subjective interpretation of probabilities, one may say that the probability to calculate x, given some knowledge K, is equal to the probability to calculate y, given the same knowledge or

$$P(find\ x|K) = P(find\ y|K)$$

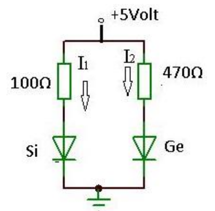
In this sense, the questions 'calculate x' and 'calculate y' are simultaneous. For this type of item, it is justified to give full marks to those students who submit correct values for both x and y and zero marks to all the others. "All or nothing" or dichotomous is the proper scoring for this type of question.

Dichotomous scoring is also suitable for marking two-tier multiple-choice items. A two-tier multiple-choice item addresses two questions: the first is factual, and the second is reasoning-based, explaining the choice made in the first question [25]. The demand for complete knowledge of both the factual and the explanatory question justifies the application of dichotomous scoring [26].

Dichotomous scoring is also justified in some problems in Analogue Electronics, stating two or more questions. In Example 1.1, knowing how to calculate I_2 is equivalent to calculating I_1. Similarly, to say that a student knows the current and voltage relationships when a BJT operates in the saturation region requires knowledge of all three conditions (Example 1.2). In another example, a student can extract information from a specification sheet when s/he can identify all the values required (e.g. the Zener voltage and the minimum and maximum current values). Dichotomous scoring increases the discrimination index of the individual items [26]. However, its usage must be adequately justified to ensure fair treatment of the students and prevent demotivation.

Example 1.1:

For the circuit given below, calculate I_1 and I_2. Select the correct values.



I_1=60mA	I_2=10mA
I_1=43mA	I_2=15mA
I_1=24mA	I_2=24mA
I_1=10mA	I_2=43mA

Example 1.2

A n-p-n BJT operates in saturation.
Select the correct propositions.

$V_{CE} \leq 0{,}2V$	$V_{BE} < 0{,}7V$	$I_B = \frac{I_C}{\beta}$
$V_{CE} \geq 0{,}2V$	$V_{BE} > 0{,}7V$	$I_B < \frac{I_C}{\beta}$
$V_{CE} = 0{,}2V$	$V_{BE} = 0{,}7V$	$I_B > \frac{I_C}{\beta}$

2) A set of fully inter-related questions

In the case of items that state entirely inter-related questions, answering must proceed in a specific sequence. The students cannot answer question i+1 unless they have correctly answered the ith question, i.e., the probability of answering question i+1 is zero, given that the answer to the ith question is wrong.

$$P(\,Q_{i+1}|\overline{Q}_i) = 0,\ \text{i=1,2, …}$$

For this type of item, finding the correct answer in one question necessitates the correct numerical answer to the previous one. The questions track students' effort, while the answers reveal authentic knowledge and identify guessing. The inter-relationship between the questions allows a more sophisticated and fair scoring. Table I shows the possible answer patterns in the case of a 3-question item together with the proper scoring rule. This scoring rule applies to Examples 2.1 and 2.2, considering that each one of them is worth ten marks. The first question counts for two marks, the second 3 and the third 5.

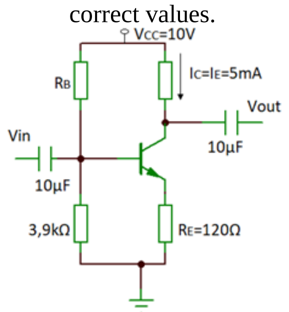
TABLE I. SCORING RULE FOR FULLY INTER-RELATED QUESTIONS

Examples 2.1, 2.2, 2.3				
Q1 (2marks)	Q2 (3marks)	Q3 (5marks)	Naïve scoring	Informed scoring
✓	✓	✓	10	10
✓	✓	✗	5	5
✓	✗	✓	7	2
✓	✗	✗	2	2
✗	✓	✓	8	0
✗	✓	✗	3	0
✗	✗	✓	5	0
✗	✗	✗	0	0
✗ : wrong selection, ✓: correct selection				

"Naïve scoring" refers to allocating marks by counting the correct answers without considering the inter-relationship between the answers. This type of scoring differs from polytomous scoring, which applies whenever the questions are independent of each other, i.e., the answer to one question does not depend on the answer to another. "Informed scoring" considers not only the correct answers selected but also the interrelationship between the answers, as described in Table I.

Example 2.1

In the following circuit, $I_E=I_C=5mA$. Calculate V_E, V_B, and R_B. Select the correct values.



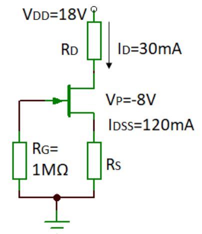
$V_E=1{,}2V$	$V_B=1{,}9V$	$R_B=10k\Omega$
$V_E=1V$	$V_B=0{,}7V$	$R_B=25{,}6k\Omega$
$V_E=0{,}6V$	$V_B=1{,}3V$	$R_B=26{,}1k\Omega$

For example, selecting the correct answer for questions 2 and 3 after making a wrong selection for question 1, is a clear case of guessing. In this case, the student will get zero marks. A student who selects the correct answer for the first two questions will get the same marks under naïve and informed marking. Informed marking does not eliminate guesswork, and some guessing is still probable. In Examples 2.1 and 2.2, there is a 1/3 probability for a student to select the correct answer to the first question by pure guessing. There is also a 1/9 probability to select the correct answer of the first two questions by guessing. However, as seen in Table II, the percentage of students who answered at least two questions was always much higher than this probability. PBMR items addressing fully inter-related questions convey a clear message to the students that authentic solutions are valued, and the selection of numbers is not a favourable strategy. Breaking the problem into three questions rewards the well-prepared students for answering less than all the questions. As discussed in Section III, this type of PBMR items and the associated scoring rule successfully identify and discard a large amount of guesswork, rewarding authentic effort.

Example 2.2

In the following circuit, the drain current is 30mA and $R_D=120\Omega$. Calculate V_{GS}, V_S, and V_{DS}. Select the correct answers.

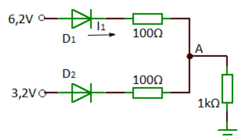
$V_{GS}=0V$	$R_S=600\Omega$	$V_{DS}=10{,}4V$
$V_{GS}=-8V$	$R_S=267\Omega$	$V_{DS}=6{,}4V$
$V_{GS}=-4V$	$R_S=133\Omega$	$V_{DS}=1{,}3V$



Example 2. 3

D_1 and D_2 are identical Si diodes. Calculate I_1. Then find V_A. Is D_2 ON or OFF?

$I_1=6mA$	$V_A=3V$	D_2 is ON
$I_1=5mA$	$V_A=4V$	D_2 is OFF
$I_1=4mA$	$V_A=5V$	

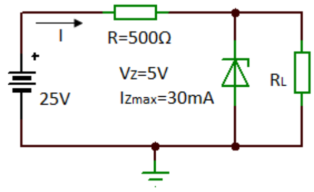


3) *A set of partially inter-related questions*

Example 3

Calculate the current I. Find the maximum (R_{max}) and minimum (R_{min}) value of R_L, for the diode to operate safely in the Zener region.

$I=50mA$	$R_{max}=500\Omega$	$R_{min}=125\Omega$
$I=40mA$	$R_{max}=1k\Omega$	$R_{min}=150\Omega$
$I=30mA$	$R_{max}=1{,}2k\Omega$	$R_{min}=200\Omega$



In this type of multiple-response item, only a subset of all the questions is inter-related. In Example 3, only questions 1 and 2 are inter-related, while answering question 3 is independent of answering the first two. The students cannot identify the correct answer for question 2 unless they know the answer to question 1, or

$$P\left(Q_{j+1} \middle| \overline{Q_j} \right) = 0, \text{ for some } j<i$$

The examiner cannot identify whether the correct selection to the independent question is the result of the student's original work or guessing; therefore, the student will get full marks for this question.

TABLE II. SCORING RULE FOR PARTIALLY INTER-RELATED QUESTIONS

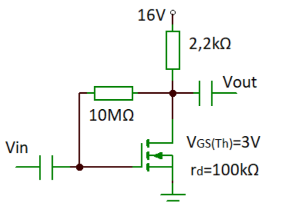
Q1 (2marks)	Q2 (3marks)	Q3 (5marks)	Naïve scoring	Informed scoring
✓	✓	✓	10	10
✓	✓	✗	5	5
✓	✗	✓	7	7
✓	✗	✗	2	2
✗	✓	✓	8	5
✗	✓	✗	3	0
✗	✗	✓	5	5
✗	✗	✗	0	0
✗ : wrong selection, ✓: correct selection				
Only Q_1 & Q_2 are inter-related.				

4) A set of completely independent questions

For this type of item, the questions are independent, and the solver may answer the questions in any sequence. Usually, items of this type can take the form of two or more single answer multiple-choice questions. In Example 4, a solver may know how to calculate Z_{in} without knowing how to calculate Z_{out} or the other way around. The absence of an inter-relationship between the questions implies a scoring rule that rewards the student for any correct answer.

Example 4

For the circuit given below, find Z_{in} and Z_{out} given that g_m=2,22mS



Zin=1,2MΩ Zout=2,2MΩ

Zin=1,7MΩ Zout=2,6MΩ

Zin=2,2MΩ Zout=2,9MΩ

III. DISMANTLING GUESSWORK & REWARDING EFFORT

This section describes some of the findings from the introduction of PBMR tests in the Department of Electrical and Electronic Engineering, University of West Attica, for final examinations and phase examinations in Analogue Electronics. A PBMR test may include items of any four item types described above, with the scoring rule varying accordingly. Items stating fully interrelated questions are the optimum type to assess the ability of the students to solve problems.

From the examiner's perspective, items of this type track the students' effort and make their knowledge transparent. Meanwhile, the associated scoring rule identifies and discards much guessing. From the student's perspective, these items make guessing ineffective and reward partial knowledge rendering negative marking unnecessary.

The duration of a test is another significant parameter. If not properly adjusted, it can affect students' performance and distort statistics [16] by forcing them to select answers randomly. If a student does not have enough time to solve a problem, ticking at random is a non-loss strategy [6]. When the duration of the exams is appropriately tuned, the examination results reflect much of the student's actual knowledge more reliably than single-answer multiple-choice tests. In the case of an item with three fully inter-related questions, selecting answers at random will have little effect on the student's overall score. On the other hand, random selections distort the item's statistics, affecting both the facility and discrimination indices.

Table III shows the students' performance in a PBMR test taken remotely during the COVID-19 pandemic. Following OECD's suggestions, we set stricter time limits to prevent dishonest behaviours [27]. The test consisted of eight items answered in four consecutive subtests (A, B, C, and D) of 15 minutes each. Subtest A comprised two items, subtest B three items, subtest C two items, and subtest D only one item. Every next subtest commenced immediately after the end of the previous, and there were no brakes in-between. Each item addressed three questions. In 6 out of the eight test items, the questions were fully inter-related (I_1, I_3, I_4, I_6, I_7, I_8), and for the other two (I_2, I_5), partially inter-related (shaded columns in Table III). Column I_1 shows that 33,6% of the students answered all the questions correctly for the first item. For item 2 (I_2), a percentage equal to 54,9% submitted correct selections for the first two questions.

TABLE III. STUDENTS' PERFORMANCE IN A PBMR TEST

Percentages of answers given										
Q1	Q2	Q3	I_1(%)	I_2(%)	I_3(%)	I_4(%)	I_5(%)	I_6(%)	I_7(%)	I_8(%)
✓	✓	✓	33,6	12,3	63,9	67,2	12,3	63,9	62,3	18,0
✓	✓	✗	7,4	54,9	15,6	11,5	4,9	9,0	5,7	45,1
✓	✗	✓	3,3	3,3	4,9	0,0	2,5	4,9	4,1	3,3
✓	✗	✗	7,4	9,8	4,1	8,2	9,8	8,2	2,5	9,0
✗	✓	✓	15,6	1,6	3,3	0,8	2,5	3,3	13,1	1,6
✗	✓	✗	10,7	2,5	2,5	5,7	8,2	1,6	3,3	3,3
✗	✗	✓	7,4	1,6	0,0	0,0	4,1	0,8	4,1	7,4
✗	✗	✗	14,8	13,9	5,7	6,6	55,7	8,2	4,9	12,3
I_1, I_3, I_4, I_6, I_7, I_8: fully inter-related, I_2, I_5: partially inter-related										

TABLE IV. TIME STATISTICS

		Subtests (min)			
		(A)	(B)	(C)	(D)
Latency	Tot	11.9	13.0	10.4	11.1
	H	11.7	13.6	10.7	12.1
	M	12.0	13.5	10.7	11.9
	L	11.1	11.9	9.8	9.0

The time statistics shown in Table IV provide information regarding students' behaviour during the test and the difficulty of the items. We compared the average time to submission (latency) for High, Medium and Low performers. Latency was approximately the same for High and Medium performers. For the Low performers, the mean latency was 89% of the respective value of the High and Medium performers. These results show that, more likely, Low performers do not solve the problems and randomly select answers. The results show that a ~7minutes time frame is necessary for answering this type of item.

Allowing longer time did not result in better performance. The time given for item 8 was 15 minutes, twice the time given for items I_1, I_2, I_6, I_7 (subtests A and C); still, the average performance was lower.

In item 5, students had the lowest performance (F.I=0,3). This item belonged together with items I_3 and I_4 to subtest B. While students performed well in items I_3 and I_4 (F.I equal to 0,7 and D.I 0,5 and 0,7 respectively), their performance was much lower in I_5. The statistics of items similar to I_5 recorded in previous tests confirm that the recorded low-performance was due to the short time given to subtest C. The number of students who submitted their answers at the last minute of the subtest confirms this explanation. For subtests A and D, nearly 50% of the students submitted their answers during the last minute. This percentage was even lower for subtest C, 25%, indicating an adequate time slot. However, 85% of the examinees submitted their answers during the last minute for subtest B. The results show that the estimated time frame for the specific PBMR items is not less than 7-8mins/item. The examiners adjusted the marks to ensure that no student was penalized because of the short time given.

Carefully designed PBMR items eliminate a lot of the mark inflation caused by guessing. Fig. 1 shows the effect of informed scoring after removing the effect of guessing from students' grades. The upper graph of Fig. 1 compares the grades of the students when scoring is “naïve” (grey dots) and “informed” (black dots). In the horizontal axis, each number represents an individual student. The graph shows that informed scoring eliminates a lot of the noise caused by guessing.

The picture becomes richer by adding information from the lower part of Fig. 1. This graph shows the potential effect of guessing on students' final scores. The horizontal axis corresponds to the Informed scores of the students, and the vertical axis shows the difference between naïve and informed scores for each examinee. The scattered symbols show that guessing could increase the overall mark of some students' scores up to 2,8 marks, having a more significant effect at the low portion of the scale. The impact of guessing is significant because of the small number of items in the test.

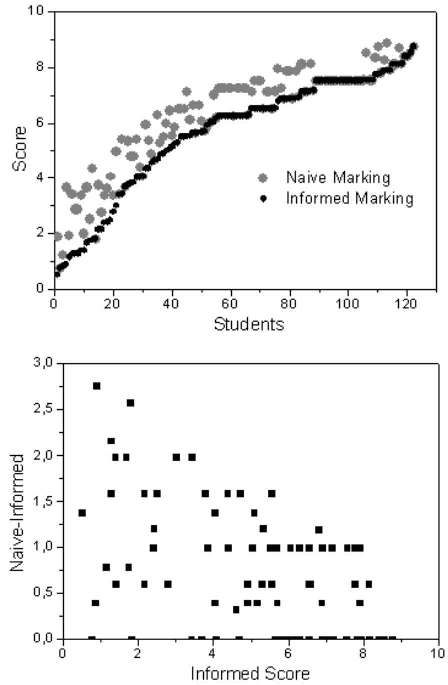


Figure 1. Visualization of the effect of naïve and informed scoring on students’ grades.

Informed scoring modified the grades of the PBMR items and improved the facility and discrimination indices. For item I_1, the 0,8-facility factor calculated after naïve scoring decreased to 0,6 after applying informed scoring. At the same time, the discrimination index increased from 0,5 to 0,7. As expected, informed scoring did not improve item statistics whenever the time slot to answer the questions was not adequate.

IV. CONCLUSION

Students often decide what to study depending on the content and type of the assessment. PBMR items combine problem-solving with the merits of objective-type examinations testing the higher levels of cognitive abilities in Bloom's taxonomy [11], [28]. The inter-relationship between the PBMR questions dictates a scoring rule that makes negative marking unnecessary, rewards partial knowledge, and corrects mark inflation caused by guessing. PBMR items track the student's effort more effectively than conventional multiple-choice tests. They point to the examinees that the assessment values original solutions and not the selection of number-answers. PBMR items are suitable for assessing engineering students' competence to solve problems and enjoy the students' preference compared to constructed response exams [6]. Further research is needed to compare student performance on these two types of assessment in an environment where ample time is given to the students to unfold their abilities.

However, three-question PBMR items leave a substantial amount of students' work and effort uncommunicated to the examiner. Future research should focus on designing PBMR items addressing many more questions. Such items will lead the students through consecutive steps, asking them to select numbers, schematics or equations from a set of appropriate options. PBMR items of this type will make a lot of the paper and pencil calculations unnecessary. Published research on multiple representations [5], [19]-[21] and appropriate approximations [23], [24] in solving problems provide sufficient background for designing more sophisticated PBMR items.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

P. Photopoulos and D. Triantis conceptualized the study, wrote and edited the full text. P. Photopoulos collected the data and edited the figures. Finally, both the authors approved the final version.

References

[1] A. N. Safitri, R. Sari, and S. Wahyuni, "The influences of mathematics ability toward physics learning in senior high school based on an authentic assessment system," *International Journal of Learning and Teaching*, vol. 3, no. 1, pp. 11-14, 2017.

[2] J. L. Docktor, J. Dornfeld, E. Frodermann, K. Heller, L. Hsu, K. A. Jackson, A. Mason, Q. X. Ryan, and J. Yang, "Assessing student written problem solutions: A problem-solving rubric with application to introductory physics," *Phys. Rev. Phys. Educ. Res.*, vol. 12, pp. 1-18, 2016.

[3] Wasis, Kumaidi, Bastari, Mundilarto, and A. Wintarti, "Analytical weighting scoring for physics multiple correct items to improve the accuracy of students' Ability assessment," *Eurasian Journal of Educational Research*, vol. 18, no. 76, pp. 187-202, 2018

[4] H. Fan, W. Chen, J. Zhang, Y. Li, X. Ye, and Q. Feng, "Innovation and practice of electronic circuits," *International Journal of Information and Education Technology*, vol. 9 no. 1, pp. 51-55, 2019.

[5] N. Johnson-Glauch, D. S. Choi, and G. Herman, "How engineering students use domain knowledge when problem-solving using different visual representations," *J. Eng. Educ.*, vol. 109, pp. 443–469, 2020.

[6] P. Photopoulos, C. Tsonos, I. Stavrakas, and D. Triantis "Preference for multiple choice and constructed response exams for engineering students with and without learning difficulties," *in* in *Proc. Proceedings of the 13th International Conference on Computer Supported Education-CSEDU*, vol. 1, pp. 220-231, 2021.

[7] E. Sorensen, "Implementation and student perceptions of e-assessment in a chemical engineering module," *European Journal of Engineering Education*, vol. 38, no. 2, pp. 172–185, 2013.

[8] S. G. Farrag, "Innovative assessment practice to improve teaching and learning in civil engineering," *International Journal of Learning and Teaching*, vol. 6, no. 2, 2020.

[9] O. E. Teo and L. E. Pueh, "Challenges for conducting the online assessment for a large class in engineering mechanics," *Advances in Engineering Education*, vol. 8, no. 4 pp. 1-5, 2020.

[10] R. Babo, L. Babo, J. Suhonen, and M. Tukiainen, "E-assessment with multiple-choice questions: A 5-year study of students' opinions and experience," *Journal of Information Technology Education: Innovations in Practice*, vol. 19, pp. 1-29, 2020.

[11] S. R. Sobral "Bloom's taxonomy to improve teaching-learning in introduction to programming," *International Journal of Information and Education Technology*, vol. 11, no. 3, pp. 148-153, 2021.

[12] S. Haniya, A. O. Tzirides, K. Georgiadou, M. Montebello, M. Kalantzis, and B. Cope "Assessment innovation in higher education by integrating learning analytics," *International Journal of Learning and Teaching*, vol. 6, no. 1, pp. 53-57, 2020.

[13] R. Elmas, G. Bodner, B. Aydogdu, and Y. Saban, "The inclusion of science process skills in multiple choice questions: Are we getting any better?" *European Journal of Science and Mathematics Education*, vol. 6, no. 1, pp. 13-23, 2018.

[14] J. Kacprzyk, M. Parsons, P. B. Maguire, and G. S. Stewart, "Examining gender effects in different types of undergraduate science assessment," *Irish Educational Studies*, vol. 38 no. 4, pp. 467-480, 2019.

[15] D. Triantis and E. Ventouras, "Enhancing electronic examinations through advanced multiple-choice questionnaires," in *Virtual Learning Environments: Concepts, Methodologies, Tools and Applications*, D. Triantis and E. Ventouras, Eds., IGI Global, 2012, pp. 1645-1664.

[16] L. Mirbahai and J. W. Adie "Applying the utility index to review single best answer questions in medical education assessment," *Arch. Epid. Pub. Health*, vol. 2, 2020.

[17] S. H. Ali, P. A. Carr, and K. G. Ruit, "Validity and reliability of scores obtained on multiple-choice questions: Why functioning distractors matter," *Journal of the Scholarship of Teaching and Learning*, vol. 16, no. 1, pp. 1–14, 2016.

[18] G. Duffy, S. Sorby, and B. Bowe, "An investigation of the role of spatial ability in representing and solving word problems among engineering students," *J. Eng. Educ.*, vol. 109, pp. 424– 442, 2020.

[19] P. Klein, A. Müller, and J. Kuhn "Assessment of representational competence in kinematics," *Physical Review Physics Education Research*, vol. 13, 2017.

[20] N. Munfaridah, L. Avraamidou, and M. Goedhart, "Preservice physics teachers' development of physics identities: The role of multiple representations," *Res. Sci. Educ*, 2021.

[21] S. V. D. Eynde, P. V. Kampen, W. V. Dooren, and M. D. Cock, "Translating between graphs and equations: The influence of context, direction of translation, and function type," *Physical Review Physics Education Research*, vol. 15 no. 2, 2019.

[22] A. Miron, E. Trotskovsy, and A. C. Cziker, "Experienced students' errors in electrical engineering," in *Proc. IEEE Frontiers in Education Conference*, 2020.

[23] E. Trotskovsky and N. Sabag, "The problem of non-linearity: An engineering students' misconception," *International Journal of Information and Education Technology*, vol. 9, no. 6, pp. 449-452, 2019.

[24] E. Trotskovsky, I. Raveh, and N. Sabag, "Mathematical vs. engineering understanding: engineering educators' perspective," in *Proc. IEEE Global Engineering Education Conference (EDUCON)*, Santa Cruz de Tenerife, Canary Islands, Spain, April 2018, pp. 649-653.

[25] A. Gero, Y. Stav, I. Wertheim, and A. Epstein, "Two-tier multiple-choice questions as a means of increasing discrimination: Case study of a basic electric circuits course," *Global Journal of Engineering Education*, vol. 21, no. 2, pp. 139–144, 2019.

[26] A. Gero and Y. Stav, "Summative assessment based on two-tier multiple-choice questions: Item discrimination and engineering students' and teachers' attitudes," *International Journal of Engineering Education*, vol. 37, no. 3, pp. 830–840, 2021.

[27] OECD. (2020). Remote online exams in higher education during the COVID-19 crisis. [Online]. Available: http://www.oecd.org/education/remote-online-exams-in-higher-education-during-the-covid-19-crisis-f53e2177-en.htm

[28] D. J. Bremner, J. Le Kernec, F. Fioranelli, V. H. M. Dale, and P. Rattadilok, "The use of multiple-choice questions in 3rd-year electronic engineering assessment: A case study," in *Proc. IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 2018, pp. 887-892.

Panos Photopoulos holds a Bachelor's degree in Physics from the National & Kapodistrian University of Athens, a Master's in Philosophy of Science from King's College London and a PhD in Nanoelectronics from the National Technical University of Athens. He is an Assistant Professor at the Department of Electrical & Electronic Engineering, University of West Attica, Greece. His research interests include

semiconducting and metallic nanoparticles' optical and electrical properties and the application of e-learning methodologies in Engineering studies.

Dimos Triantis studied at the National & Kapodistrian University of Athens. He received a BSc in Physics (1975), an MSc degree in Electronics (1980) and a PhD in Solid State Physics (1983). He is a Professor at the Department of Electrical & Electronic Engineering, University of West Attica, Greece. His research interests mainly focus on materials and electronic devices' electrical properties and the detection and study of mechanically stimulated weak electric signals and acoustic emissions. His research work also includes new technologies and methodologies in education and student assessment.