

# PEER LEARNING FOR LARGE SIZE PHYSICS LECTURES IN HIGHER EDUCATION: YES, WE CAN

Matteo Bozzi<sup>1</sup>, Juliana Elisa Raffaghelli<sup>2</sup>, Maurizio Zani<sup>1</sup>

<sup>1</sup>*Politecnico di Milano, Department of Physics (ITALY)*

<sup>2</sup>*Universitat Oberta de Catalunya, Department of Psychology and Educational Sciences Studies (SPAIN)*

## Abstract

The academic course on Mechanics and Electromagnetism “Fisica Sperimentale A+B” at Politecnico di Milano was selected to investigate peer learning effects in large size lectures. It was attended by both freshmen attending the first year of Chemical Engineering and those studying Materials and Nanotechnology Engineering during the first term of the academic year 2017-2018. After arranging the students in three different sections on the basis of an alphabetical order, the research team modified the design and pedagogical method of one of these sections. In addition to traditional lectures, the students of this section took periodical tests based on multiple choice quizzes whose provision was implemented by using the on-line portal Socrative and by allowing the students to use their own electronic devices. Immediately after each quiz and without getting any feedback on the accuracy, freshmen would discuss the quizzes in small groups for few minutes. At the end of this debate, they retook the same test.

Thanks to the significant bulk of data collected, the findings are likely to show that the average score achieved by the students in the same test, normalised so that the highest value is equal to 10, increases considerably as a consequence of peer discussion: from 5.1 to 6.7 in Mechanics, corresponding to a growth by 32%, and from 3.6 to 5.6 in Electromagnetism, corresponding to an increase by 55%. Moreover, a sensible decrease in the incorrect answers rate emerges in relation to each quiz: on average, 43% and 30% in Mechanics and Electromagnetism respectively. Furthermore, significant and comparable reductions of this rate occur as regards quizzes characterised by both low and high percentage of incorrect answers in quizzes which precede peer learning and peer evaluation. Finally, the percent decrease in incorrect answers given by female students is higher than the male percent reduction in 90% of the Mechanics quizzes and lower in the remaining 10%; conversely, they are both equal to 50% with relation to Electromagnetism quizzes.

Keywords: Large size lectures, Physics, Peer learning, BYOD.

## 1 INTRODUCTION

In the context of the modernisation of Higher Education, traditional lectures are considered insufficient to cover learners’ needs [1] and the efficacy of more participatory and active methods has been pointed out [2]. The synergistic use of both traditional lectures and educational technologies supporting self-assessment, peer learning and peer evaluation alike, has been proved effective. In this context, the high number of students attending massive lectures may pose a crucial problem connected with disengagement, demotivation and final failure; these issues are an object of increasing research and interest from educational researchers and practitioners alike [3]. In the literature, crowded lectures lead to fewer interactions and pedagogical support from teachers, not to mention disorientation and troubled access to course activities and resources for learning [4]. Frequent solutions are highly directive and transmissive pedagogical approaches, where the lecturer is forced to deliver content. Less often, due to the workload, lecturers introduce solutions that aim at active learning, or that encompass the adoption of tools for self-monitoring learning [5]. Within this context, peer learning has become an issue of relevance, due to the fact that the teacher presence is amplified by peer interactions [6, 7]. Nevertheless, it seems there is a lack of evidence for the combination of factors that encourage interactions among peers in specific learning situations like those provided by large size lectures, which still requires specific research [8]. Overall, in this article we built on the confirmed effectiveness of peer learning in contexts characterised by synergistic use of both traditional lectures and educational technologies [9], deepening on the case of peer discussions combined with Students’ Response Systems (SRS) in the specific context of Physics’ teaching and learning.

## 2 BACKGROUND

Peer learning and peer assessment have been investigated in the last 40 years with consolidated conclusions for the issue. Good peer interactions are always based on sound students' preparation, along with socialisation with the teaching method and with the assessment model and instruments. Moreover, peer interactions lead to effective learning when teachers give feedback both on the specific learning results and on the quality of the peer learning process [10]. In the case of Physics, it is worth considering the Peer Instruction (PI) strategy which was introduced in 1991 by Physics Professor Eric Mazur [11]. This technique has proved to be effective in enhancing the undergraduate student's learning outcomes in Physics both in the theoretical reasoning and in the quantitative problem solving [9]. PI is based on the idea, mainly studied in Maths teaching but applicable to Physics, that expert teachers could fail in uncovering the difficult connections between factual knowledge and conceptual problems in sciences, due to their own high conceptual skills that would generate what has been called the "expert blind spot" [12]. To this regard, the students' struggle to apply factual knowledge to conceptual problems requires modelled explanations that are often based on more "visible" reasoning and procedures, such as verbalised step by step explanations. In Mazur's technique, multiple choice conceptual questions are posed at key parts of the lecture. If the majority of the students' responses are incorrect they are asked to turn to their neighbour to convince them of their answer. After being adopted by physics teachers [13], PI has been widely employed in many other disciplines, such as undergraduate Science, Technology, Engineering and Maths (STEM) [14]. In order to become fully effective, this technique requires students' preparation before class, when their knowledge is tested. The part of peer instruction is hence accomplished in class, including several forms of feedback (from the teacher, from peers and through self-guided activities).

More recently, in the context of experimental massive learning environments like those provided by MOOCs, peer learning and particularly peer assessment have been implemented as a strategy to overcome the logistic problem of providing feedback to a high number of students [15, 16]. While the peer learning and peer assessment techniques have proved to be effective, there is a number of considerations to take into account when implementing them. For example, the need to adopt software supporting casual assignment of assessment tasks and objects of assignment; students' training in the assessment task; structured tools to perform assessment like rubrics; training in the assessment activity with overall teacher feedback; worked examples and best practices preparing the students to assess their peers [16]. MOOC's expertise has been insightful for face-to-face education and for blended learning, encompassing teaching innovations based on peer learning.

Hence, educational technologies have brought about important innovations supporting the forms of peer interactions and teacher feedback. To this regard, a valuable contribution to new forms of feedback has been provided by the research on SRS. Again, the topic is not new to the research, with contributions that go back to the 70s relating the adoption of interactive technologies to facilitate teacher feedback. In the last ten years we have assisted to an explosion of techniques such as personal response system, audience response system, electronic voting system, which allow each student to send their response to the teacher, who can see these data in a class dashboard and share, with the students [17, 18, 19]. These types of tools have been finally grouped under the name of "Students' Response System", which typically consists of software to create and manage questions, and a combination of three hardware components: interactive remote controls, a receiver unit and a classroom computer with a projection system. The students might use handheld devices with unique ID called "clickers" to give their responses by pressing the clicker buttons. More recently the massive presence of smartphones amongst students has led to the adoption of systems where the clicker is replaced by the student's own device, aligning with the Bring Your Own Device (BYOD) strategy. While this strategy is clearly effective at reducing costs, and also encompasses good motivational effects for the students that can control their own learning in a highly flexible and personalised way, there is a possible negative impact associated to connection costs and the potential distraction effect of other app notifications in the smartphone while performing the learning tasks [20]. In any case, recent studies have proved that the BYOD strategy is preferred to clickers [21, 22]. The experimental activity in class has been supported by the increasing presence of online applications supporting students' interactions and particularly forms of response systems, some of them using gamified approaches [23]. In our analysis of the background, there is little research on the approaches that combine peer learning after the SRS activity, as form of specific feedback. This issue motivated our experimental approach in the field of undergraduate Physics teaching. Moreover, taking into consideration the importance of teaching methods that address the problem of females' engagement and participation in the STEM field, we conducted an analysis of the specific effects of our experimental activities on the female students.

### 3 METHODOLOGY

In the light of the posed research problem, the authors designed a quasi-experiment for the academic year 2017-2018 which aimed at analysing the effectiveness of peer learning activities supported by a SRS as a teaching method addressing the challenge of feedback in a large size class. This quasi-experimental design encompasses a protocol to administrate an experimental treatment (in this case the adoption of a peer learning strategy mediated by a SRS) in a natural environment, i.e. an ongoing class activity. While experimental design encompasses manipulation that might not align with the social and cultural goals of educational institutions, a quasi-experimental design would be a more feasible option in complex educational settings [24, 25] in that it encompasses control and experimental groups that are not randomly selected but available on the basis of the course's enrolments.

The quasi-experimental activity was undertaken within the academic course on Mechanics and Electromagnetism "Fisica Sperimentale A+B". This Physics programme is normally provided to both freshmen attending the first year of Chemical Engineering and those studying Materials and Nanotechnology Engineering during the first term of the academic year at Politecnico di Milano (North-West Italy). The students enrolled in the academic year 2017-2018 amounted to a total of 646; nevertheless, some freshmen did not attend the lessons regularly and the average number of students participating reduced over time. After arranging the students in three different sections on the basis of an alphabetical order, one of these sections (section 1 with 206 students) participated in the research. The results of this section were compared to the other two sections (section 2 and section 3, with 179 and 261 students respectively); as a consequence, one treated group and two control and comparator groups were established.

The experimental treatment was based on an educational design where, in addition to the traditional lectures, the students of section 1 took periodical tests based on multiple choice questions; the topic of the first and the second test was Mechanics (each test consisted in five multiple choice questions), whereas Electromagnetism was investigated in the third and fourth test (each one consisted of four multiple choice questions). Freshmen took these quizzes by using the SRS Socrative, whose use in the implementation of interactive teaching has been widely highlighted [18, 26]. Taking into account the documented positive results of BYOD [18, 27], the students were allowed to use their own electronic devices, like smartphones, tablets and laptops. They were required to log in on Socrative with the same ID as the one they use at university, which would have allowed us to determine their gender. Nonetheless, some students logged in using a pseudonym, like "Gdff" or "Y", making it impossible for us to identify their gender. As a consequence, the sum of female and male registered students was lower than the total number of freshmen who sat the tests.

The time allotted to each quiz was two minutes. The students earned one point in the event of the correct answer and 0 points in all other cases; the total score achieved by a student in each test was normalised so that the highest value was equal to 10. Immediately after each quiz and without getting any feedback on the accuracy by the other students and the teacher, freshmen were asked to work with four fellow students sitting next to them so as to discuss the quizzes in small groups for three minutes. At the end of this debate, they retook the same quiz. Finally, the correct solution of each quiz and the percentage of answers ascribed to each possible option were shown to the students.

In order to evaluate the trial carried out with section 1, both the treated group and the control groups involved in the research took an identical test during the first lesson of the course (Pre Test) which consisted of twelve multiple choice questions, six of which based on Mechanics topics and the other six on Electromagnetism subjects. Towards the end of the course, the same set of quizzes was administered again to the students of the three sections (Post Test). On both occasions, neither peer learning nor peer evaluation were put into action.

The innovative teaching elements which characterised the course attended by the freshmen of section 1 aimed at investigating the potential impact of self-assessment, peer learning and peer evaluation alike on students' learning in Physics, as their positive effect on education has been thoroughly investigated and described in relevant works [8, 9]. However, considering that all the quizzes were taken by about 150 students, a significant bulk of data was collected and quantitatively processed, which constitutes an element of novelty in the literature on Physics teaching. Table 1 introduces a synthesis of the research design.

Table 1. Research design

Independent variables and co-variates	Dependent variables	Instruments for data collection	Statistics
Use of SRS [Presence/Absence]	Overall students' performance	General Pre and Post Test Control Test	Descriptive Statistics Asymptotic Kruskal-Wallis Test (non parametric) Effect size
Use of Peer Instruction [Presence/Absence]	Specific Students' Performance	Topics under Peer-Instruction Quizzes	Descriptive Statistics Average Normalised Gain Wilcoxon Test
Gender (co-variate)	Gender influence on overall and specific performance	Variable extraction from Students' ID	Rank repeated measures ANCOVA (non parametric)

## 4 RESULTS

Firstly, we analysed the overall results of the experimental method, comparing the several sections (students' groups) engaged in the activity. As we mentioned before, one section received the experimental treatment (SRS + peer learning activity) and the other two sections participated in traditional lectures. Table 2 shows the descriptive score statistics for the different sections, with a focus on gender. Overall, it was possible to appreciate the positive difference between the experimental section (S1) and the two control sections (S2 and S3). This result led to further exploration through an inferential, non parametric test focused on the effects of the treatment across section. The test applied was the Kruskal-Wallis H, which result (chi-squared 37.58, N = 196, df = 2, p-value  $\leq 0.01$ ) supported the rejection of the null hypothesis where the sample medians were equal. It could be inferred that there was very small probability that the differences across sections' medians in the analysed post-tests scores could be ascribable to randomness, and the overall teaching method influenced the positive results.

Table 2. Descriptive score statistics for the several sections in Pre and Post Test, with a focus on gender

Gender	Section	Median		Mean		Standard Deviation	
		Pre Test	Post Test	Pre Test	Post Test	Pre Test	Post Test
F	S1	3.33	7.50	3.35	6.45	1.49	1.67
F	S2	4.17	5.00	4.11	5.22	1.85	1.63
F	S3	5.00	6.67	4.34	5.71	2.13	2.21
M	S1	4.17	6.67	3.71	6.82	1.42	1.31
M	S2	4.17	5.00	4.47	5.08	1.60	1.62
M	S3	5.00	6.67	4.33	6.63	1.63	1.86

Regarding the overall results for gender, diversified situations were observed across the sections. While in sections 1 and 2 males obtained better mean score in the Pre Test, the females' scores in the section 3 were slightly better. Moreover, the medians in sections 2 and 3 had very similar values. However, it can be noted that in the Post Test the females mean score in the section 2 improved to align males' performance, whereas in the section 3 the males' performance clearly overcome females' scores. The data shows variability and a general pattern of improvement across different sections, from the initial to the final situation, in spite of the gender constant (better initial and final performance of males). The assumption that gender as co-variate could influence the different performances was explored through the rank repeated measures ANCOVA (ANalysis of CO-Variance test). The result was not significant at the cut-off point 0.05 ( $F = 0.0176$ , N = 196, df = 1, p-value  $\geq 0.05$ ). Therefore, the conclusion was that the medians of females and males were influenced equally according to their baselines, with no significant difference across groups.

In the following paragraphs we will discuss the peer-learning activities and results in the light of the diversified male-female results. Given the available data, gathered before the peer activity and related

to Mechanics (Mc) and Electromagnetism (Em) tests, one could highlight that the level of the students' knowledge of these two branches of Physics is rather dissimilar. On the one hand, the average normalised score achieved by the freshmen in Mc Pre Test is 5.07 and the standard deviation is 2.28, which is close to a sufficient result; on the other hand, the findings in Electromagnetism are far from a threshold of acceptable sufficiency, being the average normalised score equal to 3.61 and the standard deviation 2.42. The gender analysis points out that no notable differences characterise female and male performance in relation to Electromagnetism: indeed, their averaged normalised scores are similar, as summarised in Table 3. On the contrary, the level of knowledge of Mechanics shown by male students is significantly higher than the female's.

Table 3. Data about Mc and Em Pre tests

	Gender	Students' number	Mean $\langle s_{pre} \rangle$	Standard deviation
<b>Mc Pre Test</b>		166	5.07	2.28
	F	42	4.49	2.23
	M	98	5.36	2.27
<b>Em Pre Test</b>		144	3.61	2.42
	F	29	3.83	2.50
	M	66	3.67	2.39

After the peer activity sessions, the students' results in the same trials considerably grow, as shown in Table 4. The percentage increase ( $\Delta = \langle s_{post} \rangle - \langle s_{pre} \rangle$ ) in the Mechanics test is equal to 32.20%, with a more pronounced effect on female's value (44.19%) than male's findings (27.60%), and even bigger in Electromagnetism, which is equal to 55.48% (49.28% for female's findings and 59.03% for male's findings).

Table 4. Data about Mc and Em Post tests

	Gender	Students' number	Mean $\langle s_{post} \rangle$	Standard deviation	Increase $\Delta$ (%)	Gain $\langle g \rangle$
<b>Mc Post Test</b>		166	6.70	2.44	32.20	0.33
	F	42	6.48	2.51	44.19	0.36
	M	98	6.84	2.40	27.60	0.32
<b>Em Post Test</b>		144	5.61	2.66	55.48	0.31
	F	29	5.72	2.37	49.28	0.31
	M	66	5.84	2.87	59.03	0.34

However, the average normalised gain  $\langle g \rangle$  has been considered as an appropriate quantitative benchmark to evaluate the peer activity effectiveness [28, 29], since 1998 [30]. Taking into account the fact that the maximum normalised value achievable in our tests is 10, it can be defined by the following mathematical relationship:

$$\langle g \rangle = \frac{\langle s_{post} \rangle - \langle s_{pre} \rangle}{10 - \langle s_{pre} \rangle}$$

where  $\langle s_{post} \rangle$  is the average normalised score achieved by the students in a test administered after the peer activity and  $\langle s_{pre} \rangle$  is the average normalised value achieved by the freshmen in the same test administered before the peer activity.

If  $\langle g \rangle$  takes on values between 0 and 0.30, the gain is low though positive; it is considered medium in the range from 0.30 to 0.70, whereas it is high with values between 0.70 and 1 [30]. It should be emphasised that a traditional course design generally produced positive but low  $\langle g \rangle$  values [29, 30]. By  $\langle g \rangle$  data in Table 2, it could be argued that the average normalised gain is in the medium range for both Mechanics and Electromagnetism and for both males and females.

If one considers the incorrect answers rate shift as a function of each single quiz in both tests, it might be highlighted that this parameter is always negative, as shown in Figures 1 and 2.

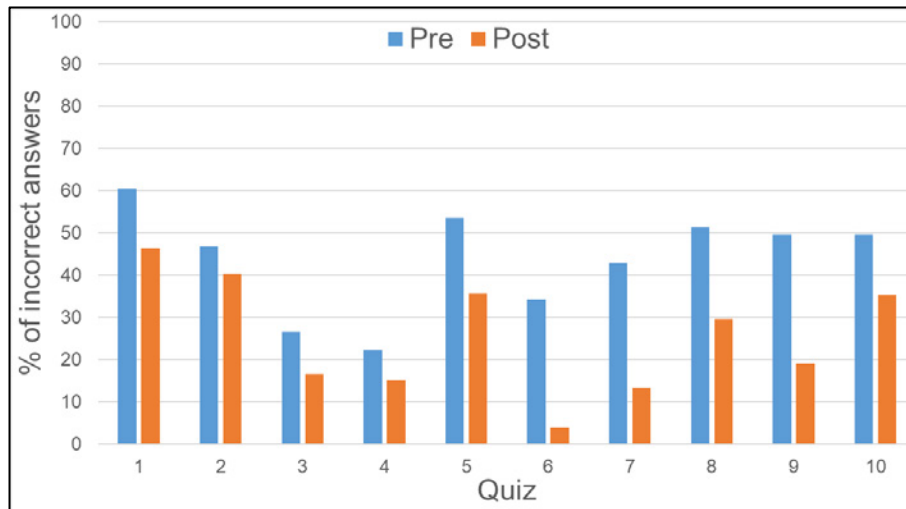


Figure 1. Percentage of incorrect answers for each single Mc Pre and Post quiz

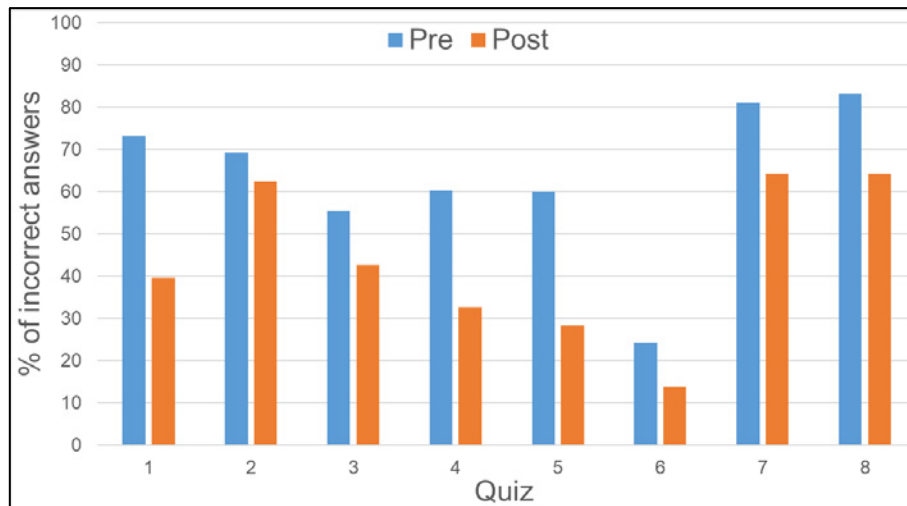


Figure 2. Percentage of incorrect answers for each single Em Pre and Post quiz

One could easily demonstrate that the incorrect answers rate shift, defined by the following mathematical relationship

$$\Delta N_{wrong}(\%) = \frac{N_{wrong\ post}(\%) - N_{wrong\ pre}(\%)}{N_{wrong\ pre}(\%)} \cdot 100$$

is opposite to a “new” parameter representing the average normalised gain ratio related to a single quiz, which might be defined as:

$$\langle g_{quiz} \rangle (\%) = \frac{N_{right\ post} - N_{right\ pre}}{N_{tot} - N_{right\ pre}} \cdot 100 = -\Delta N_{wrong}(\%)$$

where  $N_{right\ post}$  is the number of correct answers given to that quiz by the students in a test provided after the peer activity, while  $N_{right\ pre}$  is the number of correct answers to the same question given by the freshmen in the same test administered before the peer activity and  $N_{tot}$  is the total number of answers. On average, the gain is  $\langle g_{Mc} \rangle = 43.12\%$  in Mechanics and  $\langle g_{Em} \rangle = 30.44\%$  in Electromagnetism.

If the level of difficulty of a single quiz is classified by the number of correct answers given by the students in a test administered before the peer activity (easy question if the correct answers are more

than 70%, difficult if they are less than 35%, regular otherwise) [4, 5], it ought to be highlighted that the average normalised gain ratio related to a single quiz is essentially independent of the level of difficulty of the quiz. Indeed, comparable variations in  $\langle g_{quiz} \rangle$ , both high and low, occur if the question is classified as easy as well as regular or difficult, as shown in Table 5.

Table 5. Difficulty of each quiz and average normalized gain related to each quiz

Quiz		$N_{wrong\ pre}$ (%)	Difficulty	$\langle g_{quiz} \rangle$ (%)		$N_{wrong\ pre}$ (%)	Difficulty	$\langle g_{quiz} \rangle$ (%)
1	Mc	60.55	Regular	23.48	Em	73.27	Difficult	45.95
2		46.79	Regular	13.73		69.31	Difficult	10.00
3		26.61	Easy	37.93		55.45	Regular	23.21
4		22.32	Easy	32.00		60.40	Regular	45.90
5		53.57	Regular	33.33		60.00	Regular	43.48
6		34.29	Regular	88.89		24.21	Easy	20.78
7		42.86	Regular	68.89		81.05	Difficult	22.78
8		51.43	Regular	42.59		83.16	Difficult	31.39
9		49.52	Regular	61.54				
10		49.52	Regular	28.85				

Finally, the average normalised gain ratio related to a single quiz given by female students is higher than the male percent reduction in 90% of the Mechanics quizzes and lower in the remaining 10% (i.e. quiz #8); conversely, they are both equal to 50% with relation to Electromagnetism quizzes, as shown in Table 6.

Table 6. Average normalised gain ratio related to each quiz

Quiz		$\langle g_{quiz} \rangle$ (%)	$\langle g_{quiz} \rangle$ (%)	$\langle g_{quiz} \rangle$ (%)		$\langle g_{quiz} \rangle$ (%)	$\langle g_{quiz} \rangle$ (%)	$\langle g_{quiz} \rangle$ (%)
			<b>F</b>	<b>M</b>			<b>F</b>	<b>M</b>
1	Mc	23.48	32.43	19.18	Em	45.95	52.94	50.00
2		13.73	15.79	15.38		10.00	12.50	21.21
3		37.93	83.33	25.00		23.21	0.00	17.86
4		32.00	50.00	20.00		45.90	64.29	44.44
5		33.33	46.67	26.47		43.48	100.00	41.67
6		88.89	100.00	85.00		20.78	5.26	28.21
7		68.89	100.00	60.71		22.78	38.89	20.93
8		42.59	36.36	43.75		31.39	31.46	34.00
9		61.54	60.00	58.33				
10		28.85	17.65	14.29				

In order to further verify the assumptions made by the item analysis, the comparisons between the Pre and Post peer learning activities were made adopting the Wilcoxon signed-rank test. The values obtained for the two topics (Mechanics and Electromagnetism) were significant at the cut-off value of  $< 0.01$  (Mechanics:  $F = 27.99$ ,  $N = 50$ ,  $df = 1$ ,  $p\text{-value} \leq 0.0001$ ; Electromagnetism:  $F = 10.85$ ,  $N = 50$ ,  $df = 1$ ,  $p\text{-value} \leq 0.001$ ); it was possible to reject the null hypothesis and infer that peer learning as an experimental activity influenced positively the students' learning as measured in the quizzes.

## 5 CONCLUSIONS

Our research corroborates the effectiveness of PI as teaching strategy in accordance with the scientific literature on education [28, 29]. In our research, the effect has been observed in a large size lecture which imposes specific constraints to the teacher-student's communication and relationship, which might entangle lower learning effectiveness. However, our research has shown an overall positive effect of PI, paving the way for the adoption of this method as strategy in large size lectures. Taking into account the values of the averaged normalised gain in relation to both Mechanics ( $\langle g_{Mc} \rangle$ ) and Electromagnetism ( $\langle g_{Em} \rangle$ ) tests, it might be argued that peer activity allows the students to improve and deepen their knowledge of studied topics, regardless of the different branches of Physics.

For the first time, as far as we know, considering the average normalised gain related to a single quiz ( $\langle g_{quiz} \rangle$ ) has revealed that peer activities are likely to be successful in improving the results of female students in Mechanics. Indeed, their reduction of the incorrect answers rate shift registered between the test administered before and after the peer activity is nine times out of ten higher than male students' one. By contrast, differences in the values emerging with relation to Electromagnetism seem to be fairly insignificant: the freshmen's reduction of the incorrect answers rate shift is five times out of ten lower than the freshwomen's. Furthermore, contrary to some reports in several previous papers [7, 8, 28], the analysis of the values that refer to the average normalised gain related to a single quiz emphasises that the effectiveness of PI is independent of the level of difficulty of the single quiz, with relation to both the branch of Physics investigated and the student's gender involved. In other words, it could be highlighted that comparable variations of the incorrect answers rate shift occur if the quiz is characterised by a low percentage of incorrect answers before the peer activity (i.e. quizzes #3 and #4 in Mechanics, quiz #6 in Electromagnetism) as well as by a high rate of incorrect answers (i.e. quizzes #5 and #8 in Mechanics, quiz #4 in Electromagnetism). As regards the latter aspect, this would seem to indicate that the performance improvement resulting from peer activity is unlikely to be exclusively caused by the presence, in each of the small groups, of freshmen and freshwomen who know the correct answer and who are endowed with the charisma and personality necessary to convince their fellow students to embrace their ideas. They are far too few, indeed. Conversely, the performance improvement could be attributed to the cooperative learning fostered by the peer activity, which might trigger the use of personal resources and knowledge by the students.

In order to further deepen these issues in the context of large size lectures, the procedure of engaging small groups involved in peer activities could be improved, increasing the randomness in the selection of group members. Moreover, it could be interesting to create at the same time all-male and all-women groups as well as mixed teams in relation with the investigation of the peer activity.

## REFERENCES

- [1] High Level Group on the Modernisation of Higher Education, "Report to the European Commission on New modes of learning and teaching in Higher Education". Luxembourg, 2014.
- [2] D. Laurillard, *Rethinking University Teaching: A Framework for the Effective Use of Educational Technology*. Routledge, 1993.
- [3] M. Brown, A. Calkins, and G. Siemens, "The Current and Future State of Higher Education," EDUCAUSE, 2012. [Online]. Available: <https://library.educause.edu/resources/2012/11/the-current-and-future-state-of-higher-education>. [Accessed: 27-Jul-2018].
- [4] C. Allendoerfer, D. Wilson, M. Plett, R. A. Bates, T. F. Smith, and N. M. Veilleux, "Student perceptions of faculty support: Do class size or institution type matter?," in *ASEE Annual Conference and Exposition, Conference Proceedings*, 2016.
- [5] J. A. Mangram, M. Haddix, M. K. Ochanji, and J. Masingila, "Active Learning Strategies for Complementing the Lecture Teaching Methods in Large Classes in Higher Education," *J. Instr. Res.*, vol. 4, pp. 57-68, 2015.
- [6] K. J. Topping, "Peer Assessment," *Theory Pract.*, vol. 48, no. 1, pp. 20-27, 2009.
- [7] D. Dancer, K. Morrison, and G. Tarr, "Measuring the effects of peer learning on students' academic achievement in first-year business statistics," *Stud. High. Educ.*, vol. 40, no. 10, pp. 1808-1828, 2015.
- [8] D. Nulty, "Peer and self-assessment in the first year of university", *Assessment & Evaluation in Higher Education*, vol. 36, no 5, pp. 493-507, 2011.



- [9] C. H. Crouch and E. Mazur, "Peer Instruction: Ten years of experience and results," *Am. J. Phys.*, 69, pp. 970-977, 2001.
- [10] D. R. Arendale, "Postsecondary Peer Cooperative Learning Programs: Annotated Bibliography 2017", Online Submiss., 2017.
- [11] E. Mazur, "Peer Instruction: A User's Manual", Prentice-Hall, Upper Saddle River, NJ, 1997.
- [12] M. J. Nathan and A. Petrosino, "Expert Blind Spot Among Preservice Teachers," *Am. Educ. Res. J.*, vol. 40, no. 4, pp. 905–928, 2003.
- [13] C. Henderson et al., "Use of research-based instructional strategies in introductory physics: where do faculty leave the innovation-decision process?," *Physical Review Special Topics - Physics Education Research*, vol. 8, pp. 1-15, 2012.
- [14] M. Borrego and C. Henderson, "Increasing in the Use of Evidence-Based Teaching in STEM Higher Education: A comparison of Eight Change Strategies", *Journal of Engineering Education*, vol. 103, no. 2, pp. 220-252, 2014.
- [15] H. K. Suen, "Peer assessment for massive open online courses (MOOCs)," *Int. Rev. Res. Open Distance Learn.*, vol. 15, no. 3, pp. 312-327, 2014.
- [16] W. Admiraal, B. Huisman, and M. Van de Ven, "Self- and Peer Assessment in Massive Open Online Courses," *Int. J. High. Educ.*, vol. 3, no. 3, p. 119-128, 2014.
- [17] M. Barber and D. Njus, "Clicker Evolution: Seeking Intelligent Design," *CBE—Life Sci. Educ.*, vol. 6, no. 1, pp. 18, 2007.
- [18] J. E. Caldwell, "Clickers in the large classroom: Current research and best-practice tips", *CBE-Life Sci Educ*, vol. 6 (1), pp. 9-20, 2007.
- [19] R. H. Kay and A. LeSage, "Examining the benefits and challenges of using audience response systems: A review of the literature," *Comput. Educ.*, vol. 53, no. 3, pp. 819-827, 2009.
- [20] J. Traxer, "Students and mobile devices," *ALT-J*, vol. 18, no. 2, pp. 149-160, 2010.
- [21] A. I. Wang, "The wear out effect of a game-based student response system," *Comput. Educ.*, vol. 82, pp. 217–227, 2015.
- [22] L. Katz, M. C. Hallam, M. M. Duvall, and Z. Polsky, "Considerations for using personal Wi-Fi enabled devices as 'clickers' in a large university class," *Act. Learn. High. Educ.*, vol. 18, no. 1, pp. 25-35, 2017.
- [23] B. Andrés, R. Sanchis, and R. Poler, "Quiz game applications to review the concepts learnt in class: An application at the University context," in *INTED proceedings*, 2015, pp. 5654-5662.
- [24] L. Cohen, L. Manion, and K. Morrison, *Research Methods in Education*, vol. 55, no. 4. London: Routledge, 2007.
- [25] M. Hammersley, "Methodological Paradigms in Educational Research," 2012.
- [26] M. F. Tretinjak et al., "Interactive Teaching with Socrative", *MIPRO 2015*, pp. 848-851, 25-29 2015.
- [27] D. Méndez-Coca and J. Slisko, "Software Socrative and smartphone as tools for implementation of basic processes of active physics learning in classroom: An initial feasibility study with prospective teachers", *EJPE*, vol. 4 (2), pp. 17-24, 2013.
- [28] A. P. Fagen, C. H. Crouch and E. Mazur, "Peer instruction: Results from a range of classrooms", *The Physics Teacher*, vol. 40, no 4, pp. 206-209, 2002.
- [29] D. Suppapyttaiaporn et al., "The effectiveness of peer instruction and structured inquiry on conceptual understanding of force and motion: a case study from Thailand", *Research in Science & Technological Education*, vol. 28, no 1, pp. 63-79, 2010.
- [30] R. R. Hake, "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses", *Am. J. Phys.*, vol. 66, pp. 64-74, 1998.