

## PEER LEARNING IN LARGE SIZE FORMAT: IS IT EFFECTIVE IN BOLSTERING PROBLEM-SOLVING AND METACOGNITIVE SKILLS?

**Matteo Bozzi<sup>1</sup>, Roberto Mazzola, Maurizio Zani**

*Department of Physics, Politecnico di Milano, Milan, Italy*

**Abstract.** *Active methods tend to bolster university students' learning. Previous studies highlight that they generally improve learners' outcomes in concept inventories focused on conceptual understanding. In the context of a physics course offered to some Italian Engineering freshmen in academic year 2021-2022, a quasi-experiment was implemented to investigate the effectiveness of a new pedagogical approach, based on the synergistic use of different educational practices and tools, and strengthened by the use of technology. Its effectiveness was measured investigating the freshmen' rate of success in their physics course final exam. This study highlights that this methodology is effective in increasing students' learning in the context of large size formats and a conservative estimate quantifies in 9.4% the size of this effect.*

**Keywords:** *Peer learning, Peer discussion, Student response system, Technology, STEM*

### 1. INTRODUCTION

The transition from secondary to tertiary education is a challenging step which requires students to take a true quality leap in terms of effort, determination and perseverance. In science, technology, engineering and mathematics (STEM) faculties introductory physics, mathematics and chemistry courses are a stumbling block to undergraduates, who face the risk of modest results in their

---

<sup>1</sup> Corresponding author (matteo.bozzi@polimi.it)

final examination or outright failure (Bozzi et al., 2019, p. 2241; Medicine et al., 2016; Örnek et al., 2007; Redish et al., 2006).

Koch (2017) defines high-enrolment, foundation courses, characterised by low success rate, gateway courses. They often constitute barriers which learners need to overcome if they want to achieve graduation (Gasiewski et al., 2012; Medicine et al., 2016). A negative learning experience in these gateway courses, indeed, represents one of the principal reasons for dropping out (Seymour & Hunter, 2019).

Since the end of the 20<sup>th</sup> century it has been pointed out that the adoption of active methodologies in STEM courses bolsters students' learning (Hake, 1998). These findings have been confirmed by both a meta-analysis which examined 225 studies involving a wide range of active instructional practices in different disciplines (Freeman et al., 2014), and a number of recent studies (Balta et al., 2017; Bauer et al., 2022; Bozzi et al., 2021). In a recent book, Hattie (2023) provides a synthesis of more than 2,100 meta-analyses covering about 130,000 studies which engaged over 200 million students at all levels of education, from primary school to tertiary level. The authors categorise about 350 factors related to learning outcomes from very positive effects to very negative influences and the average effect size taken into account (Cohen's  $d$ ) is .40.

Hattie's investigations (Hattie, 2008, 2023; Hattie et al., 2015) show that some widespread active methodologies have a positive effect size, although lower or equal to the mean value. For instance, Cohen's  $d$  ( $d$ ) for problem-based learning and for cooperative learning is respectively .12 and .40. On the contrary, such factors as classroom discussion ( $d = .82$ ), positive peer influences ( $d = .53$ ) and peer tutoring ( $d = .53$ ) show a higher effect size than the average value. With specific reference to the teaching of physics, Student-Centred Activities for Large Enrolment Undergraduate Programs (Beichner, 2007; K. Foote et al., 2016; K. T. Foote et al., 2014), Problem-Based Learning (Duch, 1996) and Peer Instruction (Crouch & Mazur, 2001; Dancy et al., 2016; Korpas et al., 2019; Rudolph et al., 2014; A. K. Wood et al., 2016; Zhang et al., 2017) are among the most noticeable pedagogical strategies designed and implemented in numerous universities.

Educational research allows arguing that students' conceptual understanding thrives when active methods are employed in STEM courses

(Freeman et al., 2014; Hake, 1998; Wieman, 2014). Furthermore, these interactive-engagement methods foster learners' interest, self-efficacy (Dou et al., 2018), metacognitive regulation (De Backer et al., 2015), motivation and engagement (Schunk & Zimmerman, 2012), which improve their academic outcomes (Delaney & Royal, 2017; Schunk & Zimmerman, 2012; Williams et al., 2019).

Freeman et al.'s (2014) seminal investigation highlights, however, that active learning improves learners' outcomes in concept inventories more than in final course examinations. Moreover, it has been empirically demonstrated that these educational approaches foster a reduction in the achievement gap which could afflict under-represented groups in STEM faculties (Seymour & Hunter, 2019; Theobald et al., 2020).

Nevertheless, the use of active methods in the gateway courses of STEM faculties remains sparse even despite an increased awareness of the need to implement them (Rasmussen et al., 2019; Seymour & Hunter, 2019; Stains et al., 2018). Instructors' dearth of time, their need to cover completely the syllabus content of their academic course, limited resources available, and a bland departmental as well as institutional support hamper their spread (Deslauriers et al., 2019; Silverthorn et al., 2006). A further barrier which may hamper the adoption of active learning is large size classes (Apkarian et al., 2021; Hornsby & Osman, 2014; Palfreyman et al., 2011; Scott, 1995; Shadle et al., 2017; Sturtevant & Wheeler, 2019). It is no wonder, thus, that studies on active methods undertaken in Europe frequently involve small or medium size formats (Leinonen et al., 2017; Schiltz et al., 2019; A. K. Wood et al., 2014).

Although the Italian context is even more complicated by the hegemonic role played by traditional academic lectures, the effects of innovative teaching methods like peer learning (PL) on students' conceptual understanding has already been investigated in the context of Italian STEM large size classes (Bozzi et al., 2018, 2021; Commission's Expert Group on Quality Investment in Education and Training et al., 2022; Directorate-General for Education, 2022; Raffaghelli et al., 2018).

In this study we wish to further investigate the issue of PL effectiveness in large size formats to examine whether this innovative educational methodology, already illustrated in previous studies (Bozzi et al., 2021, 2023), may contribute to improve not only the learners' conceptual understanding, but more generally

their learning strategies and bolster different skills, like problem-solving and metacognitive skills.

## **2. METHOD**

### **2.1 REASERCH DESIGN**

In the Italian context, undergraduates' achievements in introductory physics, mathematics and chemistry courses are generally assessed through instructor written exams which focus on problem-solving skills rather than conceptual understanding. Although this type of final examination is probably not the most favourable to evaluate the effectiveness of active instructional practices implemented in gateway courses, this is a widespread option adopted by many Italian universities, departments and instructors to measure learners' achievements. Thus, a wider adoption of innovative teaching methods requires taking up the challenge of demonstrating that they are effective in improving undergraduates' results in final exams based also on skills different from conceptual understanding. As a result, in the present study we identified the pass rate of a physics course as the dependent variable, comparing freshmen who attended and not attended the peer learning (PL) sessions and evaluating the amount of attendance.

In this regard, an innovative teaching methodology was implemented in the context of an introductory physics course attended by some Politecnico di Milano freshmen in academic year 2021-2022. Its main characteristics were the synergistic use of different educational practices and tools, like concept tests with double vote within the "Bring Your Own Device" (BYOD) framework, students response system (SRS), discussions among peers and explanations provided by an expert, along with the integration of active methods enhanced by the use of technology, traditional lectures, and exercise sessions in large size formats. A research question was formulated:

(RQ) Does this integrated teaching methodology improve the students' final examination pass rate in a physics course?

### **2.2 RESEARCH CONTEXT AND PARTICIPANTS**

In the first term of the academic year, Politecnico di Milano freshmen enrolled on both Chemical Engineering and Materials and Nanotechnology Engineering are required to attend an introductory physics course called "Experimental

Physics A+B” (“Fisica Sperimentale A+B” in Italian language). In this gateway course students address a number of mechanics and electromagnetism topics and are granted ten university educational credits if they pass the final exam.

In academic year 2021-2022 we conducted a quasi-experiment which involved 202 first year students enrolled on “Experimental Physics A+B”. In terms of gender, 62 of them identified themselves as females and the others 140 as males. According to a student survey, conducted by our university at the end of the term, the attendance at this academic course was high. Only 5.4% of the participants attended less than 25% of the activities, namely lectures and exercise sessions. Moreover, only 11.4% of the learners attended less than 50% of this course.

Moreover, all of them accepted to participate in the study. As the students were numerous, the size of the class may be regarded as large.

### **2.3 DESIGN OF THE TEACHING/LEARNING EXPERIENCE**

The overall duration of “Experimental Physics A+B” was 100 hours, from September 2021 to December 2021. Traditional academic lectures within a theoretical framework covered 57 hours, 40 hours consisted in exercise sessions aimed at fostering learners’ problem-solving skills, and 3 hours were devoted to PL activities which were divided into 7 sessions focused on mechanics (4 sessions) and electromagnetism topics (3 sessions). Offered to the learners on average every ten days, the PL sessions were included within seven traditional lectures and the setting was the same large lecture classroom. They were conducted by the same instructor who gave classic academic lectures and their attendance was not mandatory in the same way as every lecture or exercise session. Moreover, the students did not know when the PL sessions would take place beforehand.

As illustrated in Bozzi et al. (2023), each PL session developed over four steps, the former three were repeated three times and the latter only once at the end of the session. In every PL session a questionnaire of three multiple choice items given in a sequence was administered to the learners through Socrative, an online portal whose effective use as a SRS has been extensively investigated (Balta et al., 2018; Gómez-Espina et al., 2019; Guarascio et al., 2017; Lim, 2017; Wood & Shirazi, 2020). Since the BYOD approach has been documented to be successful in terms of students’ engagement and outcomes (Afreen, 2014;

Blasco-Arcas et al., 2013; Caldwell, 2007; Hattie, 2008, 2023; Mayer et al., 2009; McLean, 2016), we asked participants to answer every question by using their own smartphones, tablets or laptops. Each item provided four possible alternatives, one correct and three incorrect, and the learners had 1.50 min to give their answer.

Before describing the design of the PL session in detail, it is worth noting that these items were constructed ad hoc by the researchers to highlight some widespread misunderstandings or conceptual errors which freshmen often make about some topics discussed in our academic course. They are frequently related to how these issues were addressed in their physics lectures at high school. These closed-ended questions were generally based on the researchers' professional background and experience in teaching physics and their awareness of how these topics were proposed in some popular Italian high school physics textbooks. Moreover, they were not selected from or inspired by published concept inventories or research-based textbooks (see Appendix 1 for one of these items).

As aforementioned, each PL session developed over four steps repeated three times. The first step began when the instructor showed to the participants the first item and gave them 1.50 min to answer it. After this first question was completed, and without any feedback on the answer, a peer discussion phase was developed in small groups (3–4 people) for few minutes (3–5 min). This was the second step. For ease of arrangement, participants were instructed to work in groups with students sitting next to them. At the end of this discussion, they answered the same question individually again. The time allotted to this third step was again 1.50 min. This way, the students debated the issue with their classmates in the small group and soon afterwards could ponder personally on it and draw their own conclusions.

This sequence was repeated in succession twice more so that the freshmen could address also the second and the third item of the questionnaire. Finally, there was one last step, i.e. the explanation from an expert, which lasted between 3 and 5 minutes. The instructor gave the correct answer to every item in the concept test and explained why the other alternatives were incorrect. Moreover, the teacher showed the percentage of participants who had chosen the different options in each question. These different percentages were shown only at the end and not after the first and the third step, not to influence the peer discussion phase and the possible contribute of each participant to it.

In conclusion, we would like to emphasise again that PL should not be intended as peer discussion alone, considering that our overall innovative active method combined a plurality of educational tools and instructional practices, among others peer discussion.

## **2.4 MODELLING OF THE PROBLEM AND INDEPENDENT VARIABLES**

To assess the effectiveness of our innovative teaching method, we considered the freshmen' rate of success in their physics course final exam during the first examination session (dependent variable), which took place between January 2022 and February 2022. The final examination was an instructor written exam. It lasted for two hours and consisted in four problems, two focused on mechanics topics and the other two on electromagnetism subjects (see Appendix 2 for an example of these problems).

This pass rate was conceivably influenced by multiple factors, like students' motivation, engagement, interest, initial knowledge in Physics and, we hope, our integrated educational approach. To evaluate their role and quantify their weight on the pass rate, two independent variables, were defined, namely "Attendance at the PL sessions" (V1) and "Exposure dose to the PL sessions" (V2).

V1 was a dichotomous variable which aggregated the learners on the basis of their attendance or non-attendance at the PL sessions. These data were collected during the same PL sessions. The criterium adopted was 50%, hence freshmen who had attended more than 50% of the PL sessions, from 4 to 7 sessions, were considered to have been exposed to PL and constituted the first group (G1). Those who were classified as not exposed to PL (between 0 and 3 sessions attended) were assigned to the second group (G2). It should be emphasised that factors like students' motivation, engagement and interest could be substantially assumed constant into G1 and G2, while they showed a discontinuity in the transition from G2 to G1. It seems reasonable, indeed, to assume that learners who attended nought PL sessions rather than one or, symmetrically, six sessions rather than seven did not have a truly different motivation, engagement and interest. Therefore, V1 was sensitive to these factors but, at the same time, was not extremely sensitive to exposure to the PL methodology due to the fact that being a dichotomous variable, it condensed the

eight possible exposure doses, which were in themselves few, into just two categories (exposed vs non-exposed).

V2 was an ordinal variable which aggregated the freshmen on the basis of the number of PL sessions which they had attended and it ranged from 0 to 7. It was blatantly sensitive to exposure to PL methodology, but it modelled the other factors worse than V1. Finally, on account of the high course attendance rates (see paragraph 2.2), V1 and V2 may not be extremely sensitive to the students general attendance at traditional lectures and exercise sessions.

### **3. RESULTS**

The data collected were analysed through the statistical opensource software R (version 4.2.2) in the integrated development environment RStudio (<https://rstudio.com/> accessed on 8 November 2023). The statistical significance was beforehand established at level  $\alpha = .05$ .

#### **3.1 IMPACT OF THE STUDENTS' INITIAL KNOWLEDGE IN PHYSICS**

A first investigation stage consisted in determining the initial level of knowledge in physics of all the students enrolled on “Experimental Physics A+B” and involved in this study at the beginning of the course. For this purpose, an initial questionnaire on some physics topics was administered to the participants during one of the first lectures of the academic course (see Appendix 3 for an example of these problems).

This questionnaire consisted of 18 multiple choice items focused on some mechanics (6), thermodynamics (6) and electromagnetism (6) subjects. Not taken from published concept inventories or research-based textbooks, it was substantially a concept test without any numeric problems. Its items, which were different from the questions offered during the subsequent PL sessions, allowed to test or check some possible freshmen’s misunderstandings, conceptual errors or misconceptions. The learners gained one point for each correct answer and nought points in all other cases, therefore the maximum possible score achieved by the participants was 18 and the minimum 0. The questionnaire was administered online and freshmen answered it using their own electronic devices (BYOD approach).

Through this instrument we could compare the initial level of knowledge in physics of G1 and G2. Table 1 shows some descriptive score statistics about the findings in this initial questionnaire regarding the two cohorts.

**Table 1: Descriptive score statistics for G1 and G2 in the initial questionnaire.**

Cohort	Number of students	Average percentage of correct answers	Mean score (max. 18)	Standard error
G1	63	30.78	5.54	.24
G2	41	30.78	5.54	.27

To check the normality of the population distributions a Shapiro-Francia test was implemented. Given that the populations were not normally distributed, we concluded that the most appropriate inferential tests to compare the different mean scores in the initial questionnaire were a robust ANOVA as well as the Mann-Whitney U-test. We conducted both these tests, which highlighted that the difference between the mean score of the two cohorts was not statistically different, i.e. the findings of the two groups were equivalent. Table 2 summarises the outcomes of these statistical tests.

**Table 2: Robust ANOVA and Mann-Whitney U-test results for G1 and G2.**

Test	Parameter	df	p-value	Effect size
Robust ANOVA	F = 0.1089	1; 53.36	.74	.0924
Mann-Whitney U	Z = 0.18563		.85	(negligible)

### 3.2 IMPACT OF THE STUDENTS' MOTIVATION, ENGAGEMENT AND INTEREST

To evaluate the possible impact of the learners' participation and engagement, we compared the pass rate of G1 and G2 students in their physics course final exam. To turn it round, we analysed if there was an association between the dependent variable "Rate of success in physics course final exam" and the independent variable V1. Although motivation and interest should be investigated through specific questionnaires and could not be completely

identified with participation, V1 was related to them. As a consequence, this analysis may provide a first indication of their possible impact.

Descriptive statistics appeared to highlight possible major differences between the pass rate of the two cohorts of freshmen. In G1, the learners who passed the exam were many more than anticipated if there were not an association between the pass rate and the attendance at the PL sessions, and the learners who failed were largely fewer than expected. The opposite happened in G2. Tables 3 and 4 are the contingency tables and highlight respectively the observed and the expected distribution, i.e. the number of learners who passed or failed the examination and the number of learners who may be expected to pass or fail it if there were no association between the two variables.

**Table 3: Observed distribution of students who passed or failed the final exam for cohorts G1 and G2.**

<b>Cohort</b>	<b>Pass number</b>	<b>Failure number</b>
G1	64	33
G2	24	81

**Table 4: Expected distribution of students who may have passed or failed the final exam for cohorts G1 and G2.**

<b>Cohort</b>	<b>Pass number</b>	<b>Failure number</b>
G1	42.26	54.74
G2	45.74	59.26

It should be pointed out that all the students enrolled on “Experimental Physics A+B”, ( $n = 202$  as shown in table 3) were considered in our analysis. The freshmen who took the physics course final exam in the first session were 161 (79.7%), aligned with the previous academic years. In fact, the average exam attendance rate had been equal to 79.0% between the academic year 2014-2015 and 2020-2021. As the present study focuses on evaluating if our methodology might determine an increase in the number of students who passed the exam within the beginning of the second academic term, the fact that some students failed the exam or did not take it was considered equivalent.

To investigate the possible association between the outcome of the exam and the students’ attendance of the PL activities, we implemented a Pearson’s

chi-square test. It highlighted a statistically highly significant association between the two variables ( $\chi^2 = 38.13533$ ,  $df = 1$ ,  $p\text{-value} = 6.60 \times 10^{-10} \ll .05$ ). Moreover, the intensity of the effect was evaluated through both odds ratio (OR) and Cramer's V (V), which were respectively equal to 6.545 (large or strong (Rosenthal, 1996)) and .434 (relatively strong (Rea & Parker, 2014)). Thus, we might argue that not only was there a statistically highly significant association between the outcome of the final exam and V1, but the strength of this association was extremely significant and appreciable.

### 3.3 IMPACT OF THE INNOVATIVE INSTRUCTIONAL PRACTICE

To assess the possible impact of our innovative educational methodology, we compared the pass rate of students who attended a different number of PL sessions in their physics course final exam. To turn it round, we analysed if there was an association between the dependent variable "Rate of success in physics course final exam" and the independent variable V2. Moreover, the pass proportion was calculated with reference to all the possible exposure doses to PL ( $P_i$ ).  $P_i$  was defined as the ratio between the number of students who both attended  $i$  PL sessions and passed the final exam and the total number of learners who attended  $i$  PL sessions. The  $P_i$  expected in case of no association between the dependent variable and V2 was .436. Table 5 shows the descriptive score statistics regarding students who were differently exposed to PL and their results in the final exam.

**Table 5: Descriptive score statistics regarding students differently exposed to the PL and their results in the final exam.**

	V2								
	0	1	2	3	4	5	6	7	Total number
Pass number	9	6	5	4	10	17	12	25	88
Failure number	39	12	18	12	9	9	10	5	114
Total number	48	18	23	16	19	26	22	30	202
$P_i$	.188	.333	.217	.250	.526	.654	.545	.833	.436

Descriptive statistics appeared to highlight possible major differences between the pass rate of the diverse cohorts of freshmen. It should be emphasised that the lowest  $P_i$  regarded students who had never attended PL sessions, while the highest  $P_i$  concerned learners who had participated in all the PL activities. Furthermore, every cohort previously assigned to G1 (from 4 to 7 PL sessions attended) showed a  $P_i$  superior to .436, which was the value expected in event of no association between the two variables studied. Conversely, each group which had been classified in G2 had a  $P_i$  inferior to .436.

Employing a Pearson's chi-square test ( $\chi^2 = 45.53499$ ,  $df = 7$ ,  $p\text{-value} = 1.08 \times 10^{-7} < .05$ ) we could argue that there was a statistically highly significant difference between at least two of these pass proportions. Moreover, Cramer's V effect size was extremely significant ( $V = .475$ , relatively strong (Rea & Parker, 2014)).

To analyse if the dependent variable "Rate of success in physics course final exam" depended on exposure to our instructional method, which was only one of the factors that might influence V2, we investigated if there was a linear trend among the pass proportions. In our model all the other possible factors, for instance students' motivation, engagement and interest, levelled out among the cohorts of students corresponding to G1 and G2, thus a linear trend might suggest an association between the exposure to PL and the rate of success. From this perspective and in analogy with a variety of carcinogenicity studies, toxicological risk assessment and clinical trials based on dose-response relationships, we carried out a dose-response experiment as we were interested in investigating a possible dose-related trend, hypothesising an increasing trend in pass proportions (response) with increasing attendance to the PL sessions (dose).

As the Cochran-Armitage (Armitage, 1955; Cochran, 1954) is the most frequently adopted trend test for binomial proportions (Chen et al., 1997; Neuhaus & Hothorn, 1999), it was implemented for this purpose and three different cases were analysed. Firstly, we considered the eight pass proportions related to every possible number of PL sessions attended by the participants (from nought to seven). Secondly, we carried out the test considering only the pass proportions corresponding to nought-three PL sessions attended (G2) and finally we investigated the possible linear trend intra G1. Table 6 summarises the Cochran-Armitage test findings.

**Table 6: Cochran-Armitage test findings.**

Case	Parameter	dim	p-value	Linear trend
From 0 PL to 7 PL	$Z = 6.2588$	8	$3.88 \cdot 10^{-10}$ ***	Yes
From 0 PL to 3 PL (G2)	$Z = 0.49276$	4	.6222	No
From 4 PL to 7 PL (G1)	$Z = 2.0095$	4	.04448 *	Yes

\*  $p < .05$ , \*\*\*  $p < .001$

The Cochran-Armitage test highlighted an overall linear trend among the pass proportions, but, at the same time, it pointed out that this trend concerned only G1.

#### **4. DISCUSSION**

To evaluate if the integrated teaching methodology which we developed was effective in improving the participants' final examination pass rate of the academic course "Experimental Physics A+B", a quasi-experimental research study was conducted. We preliminarily investigated possible differences among the freshmen involved in this study in terms of their initial knowledge of Physics. Although the students had not been randomly assigned to G1 and G2, it was demonstrated that their knowledge of Physics was equivalent at the beginning of the academic course. It should be emphasised that their mean score in the initial questionnaire (5.54, corresponding to 30.78% of correct answers) appeared to be much lower than the maximum score potentially achievable, i.e. 18. It implies that all the learners appeared to be slightly weak in Physics.

Interestingly, a distinctive peculiarity of this study was that not only both G1 and G2 had the same instructor, but they attended the same lectures and exercise sessions at the same time. As a consequence, notwithstanding that the teacher is recognised to play a paramount role in the students' learning process (Hattie, 2012), the possible positive or negative impact of the lecturer was substantially equivalent for every group and each learner.

The Pearson's chi-square test pointed out a statistically highly significant association between the rate of success in the physics course final exam and the independent variable "Attendance of the PL sessions". Furthermore, the size of the effect might be classified as large or relatively strong with reference to odds ratio and Cramer's V respectively. However, this independent variable was influenced by many factors, like students' motivation, engagement, interest, and

involvement in our innovative instructional practice (“Exposure to the PL methodology”). As a consequence, it is not possible to argue that the learners’ increased success rate was the result of our teaching method. Rather, it depended on all these factors and their synergistic action. Furthermore, if it was not possible to quantify the contribute of each of them to an increase in the freshmen’s success rate in their final exam, likewise it was not possible to evaluate how much every factor contributed to the odds ratio and Cramer’s V overall values. However, since the independent variable V1 was not very sensitive to the factor “Exposure to the PL methodology”, we assumed very conservatively that the intensity of the effect was entirely related to all the other factors and considered the calculated odds ratio and Cramer’s V as a baseline.

In a further step, a statistically highly significant correlation was found between the rate of success in the physics course final exam and the second independent variable V2 “Exposure dose to the PL sessions”. Like V1, V2 did not purely depend on the students’ exposure to the PL methodology, but it was also influenced by all the other aforementioned factors. However, it ought to be remembered that in our model these further factors, like engagement, interest and motivation, should be considered substantially constant within G1 and G2 and they may show a discontinuity, a sudden increase in the transition from G2 to G1 owing to the binary nature of V1.

Therefore, with regard to the influence of these factors if the “Exposure to the methodology” was ineffective or its effects were negligible, the model would forecast that they equally affect the rate of success in the final exam of learners who respectively had attended from nought to three PL sessions (G2) and from four to seven PL sessions (G1). Thus, the  $P_i$  values should show a discontinuity in the transition from three and four PL sessions attended, but they ought to be flat simultaneously within G2 and G1. However, the  $P_i$  values were flat within G2, but not within G1, where a linear upward trend emerged. As a result, we might conclude that these further factors could not determine the observed upward trend in the  $P_i$ , while this trend is consistent with our modelling of the factor “Exposure to the methodology”. Thus, our instructional practice appeared to be effective in fostering participants’ learning.

However, it is reasonable to assume that there is an efficacy threshold of this instructional practice, i.e. our innovative educational method may be generally effective only if a student attends a minimum number of PL sessions. The findings of the Cochran-Armitage tests involving G1 and G2 confirmed this hypothesis and showed that this threshold corresponded to four PL sessions.

If instructors would like to try out our PL methodology in their academic courses, they should be conscious that it needs regularity in its administration to

the students. With regard to time, however, the efficacy threshold found corresponded to only about 1.5%-2% of the overall duration of our physics course, a modest investment in time. Moreover, PL sessions were included within traditional lectures which were not based on physics education research. The setting was the same large lecture classroom and the PL sessions were conducted by the same lecturer who gave classic academic lectures. Furthermore, only three items per session were used and an instructor could decide to employ concept questions already available in research-based textbooks.

A conservative quantitative assessment of the PL effectiveness in increasing the students' success rate in their final exam could be done by comparing the intensity of the effect when V1 and V2 were considered. Cramer's V was respectively equal to .434 and .475, corresponding to a percentage increase of 9.4%. On balance, our innovative teaching methodology might improve the freshmen's success rate by about 9.4% and this appears to be a cautious estimate.

This result is not immediately comparable with the findings achieved in Hattie's studies (Hattie, 2008, 2023) due to the fact that this percentage cannot be expressed in terms of Cohen's d. However, according to Freeman et al. (2014) the students who attend academic courses consisting in traditional lectures are 1.5 times more likely to fail their final exams than students attending courses with active learning, as the risk ratio for failure (RR) is 1.5. In our study, if the freshmen who had attended 0 PL sessions are assumed as a baseline on account of their being the only learners who were not at all exposed to PL methodology, the RR with reference to respectively all the other students (from 1 to 7 PL sessions attended), the other freshmen of G2 (between 1 and 3 PL sessions attended) and the G1 students (from 4 to 7 PL sessions attended) is respectively  $1.67 > 1.5$  (+11.3%),  $1.10 < 1.5$  (-26.7%) and  $2.74 >> 1.5$  (+82.7%). A prudent conclusion is that learners who attend only traditional lectures and exercise sessions are 1.67 times more likely to fail their final exams than freshmen exposed to our PL methodology and this value is 11.3% higher than the average RR that they would have if other active methods are considered.

It is, however, worth highlighting that although there seems to be a correlation between the innovative teaching method that we propose in the present paper and our participants' rate of success, it is not possible to establish a causal effect between them, considering that other factors, like the participants' general course attendance, motivation, engagement, and interest may have contributed to their success rate.

## 5. CONCLUSIONS

In the context of a physics course offered to some Politecnico di Milano Engineering freshmen across the first term of academic year 2021-2022, a quasi-experiment was implemented to investigate the effectiveness of a new pedagogical practice, based on the integration of PL activities strengthened by the use of technology, traditional lectures and exercise sessions. Importantly, it has been highlighted that this innovative educational methodology is effective in increasing the students' learning in the context of large size formats and a conservative estimate quantifies the size of this effect as 9.4%. Furthermore, it appears on average more effective than other active methods in improving the students' rate of success in their final exam when this final exam consists in some problems to solve.

As the physics course final exam focused on problems which required a clear conceptual understanding of the topics involved, we argue that our PL methodology bolsters problem-solving skills and conceptual understanding. It should be emphasised that active learning generally tends to improve learners' outcomes in concept inventories focused on conceptual understanding more than it tends to boost the students' findings in instructor-written course examinations (Freeman et al., 2014). It is worth pointing out that we did not measure specifically the improvement of freshmen conceptual understanding through an ad-hoc concept test in this study. However, we had already highlighted the effectiveness of our PL methodology in increasing student conceptual understanding in physics in a previous research (Bozzi et al., 2021; Directorate-General for Education, 2022, p. 48).

Finally, it has been confirmed the hypothesis that there is an efficacy threshold of the PL educational method. Student learning begins to improve already after four PL sessions, corresponding to about 1.5 h - 2 h out of 100 h which is the overall duration of our physics course. It would seem a low threshold and this might encourage instructors to experiment with this innovative educational methodology.

Lastly, it is worth emphasising that traditional lectures and exercise sessions offered during "Experimental Physics A+B" were not physics education research-based and the items employed during PL sessions were not inspired by published concept inventories or research-based textbooks. As a consequence, it is unlikely that the effectiveness of our innovative teaching methods is related to physics education research findings, like the activation of some productive students' resources during PL sessions. On the contrary, the positive results achieved may be due to the method which fosters the learners' metacognitive regulation and their self-regulated learning. The research on this issue has shown that metacognitive regulation improves the students'

performance (Meijer et al., 2006; Schraw & Moshman, 1995; Zimmerman & Schunk, 2011). Moreover, peer interaction is extremely effective in triggering regulatory skills, like planning, monitoring and evaluation (De Backer et al., 2015; Schraw & Moshman, 1995). Considering this perspective, it could be argued that our innovative teaching methodology is also effective in bolstering metacognitive skills.

## REFERENCES

Afreen, R. (2014). Bring Your Own Device (BYOD) in Higher Education: Opportunities and Challenges. *International Journal of Emerging Trends & Technology in Computer Science*, 3(1), 4.

Apkarian, N., Henderson, C., Stains, M., Raker, J., Johnson, E., & Dancy, M. (2021). What really impacts the use of active learning in undergraduate STEM education? Results from a national survey of chemistry, mathematics, and physics instructors. *PLoS ONE*, 16(2 February). Scopus. <https://doi.org/10.1371/journal.pone.0247544>

Armitage, P. (1955). Tests for Linear Trends in Proportions and Frequencies. *Biometrics*, 11(3), 375–386. <https://doi.org/10.2307/3001775>

Balta, N., Michinov, N., Balyimez, S., & Ayaz, M. F. (2017). A meta-analysis of the effect of Peer Instruction on learning gain: Identification of informational and cultural moderators. *International Journal of Educational Research*, 86, 66–77. <https://doi.org/10.1016/j.ijer.2017.08.009>

Balta, N., Perera-Rodríguez, V.-H., & Hervás-Gómez, C. (2018). Using socrative as an online homework platform to increase students' exam scores. *Education and Information Technologies*, 23(2), 837–850. <https://doi.org/10.1007/s10639-017-9638-6>

Bauer, T., Biehler, R., & Lankeit, E. (2022). ConcepTests in Undergraduate Real Analysis: Comparing Peer Discussion and Instructional Explanation Settings. *International Journal of Research in Undergraduate Mathematics Education*. <https://doi.org/10.1007/s40753-022-00167-y>

Beichner, R. J. (2007). The Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) Project. In *Research-Based Reform of University Physics* (Vol. 1, pp. 1–42). E. F. Redish and P. J. Cooney (eds). <https://www.per-central.org/items/detail.cfm?ID=4517&Attached=1>

Blasco-Arcas, L., Buil, I., Hernández-Ortega, B., & Sese, F. J. (2013). Using clickers in class. The role of interactivity, active collaborative learning

and engagement in learning performance. *Computers & Education*, 62, 102–110. <https://doi.org/10.1016/j.compedu.2012.10.019>

Bozzi, M., Balossi, B., Sieno, L. D., Ganzer, L., Gondoni, P., Genco, I., Minnai, C., Pini, A., Rezoagli, F., Zanoletti, M., & Zani, M. (2019). Securing freshmen's learning through a physics refresher course: A breakthrough experience at Politecnico di Milano. *ICERI2019 Proceedings*, 2237–2243. <https://library.iated.org/view/BOZZI2019SEC>

Bozzi, M., Mazzola, R., & Zani, M. (2023). Peer learning in higher education: An example of practices. *Il Nuovo Cimento C*, 46(6), 1–11. <https://doi.org/10.1393/ncc/i2023-23206-7>

Bozzi, M., Raffaghelli, J. E., & Zani, M. (2021). Peer Learning as a Key Component of an Integrated Teaching Method: Overcoming the Complexities of Physics Teaching in Large Size Classes. *Education Sciences*, 11(2), Article 2. <https://doi.org/10.3390/educsci11020067>

Bozzi, M., Raffaghelli, J., & Zani, M. (2018). Peer learning for large size physics lectures in higher education: Yes, we can. *ICERI2018 Proceedings*, 8739–8747. <https://library.iated.org/view/BOZZI2018PEE>

Caldwell, J. E. (2007). Clickers in the Large Classroom: Current Research and Best-Practice Tips. *CBE—Life Sciences Education*, 6(1), 9–20. <https://doi.org/10.1187/cbe.06-12-0205>

Chen, J. J., Kodell, R. L., & Pearce, B. A. (1997). Significance Levels of Randomization Trend Tests in the Event of Rare Occurrences. *Biometrical Journal*, 39(3), 327–337. <https://doi.org/10.1002/bimj.4710390307>

Cochran, W. G. (1954). Some Methods for Strengthening the Common  $\chi^2$  Tests. *Biometrics*, 10(4), 417–451. <https://doi.org/10.2307/3001616>

Commission's Expert Group on Quality Investment in Education and Training, Directorate-General for Education, Y., Fack, G., Agasisti, T., Bonal, X., De Witte, K., Dohmen, D., Haase, S., Hylen, J., McCoy, S., Neycheva, M., Pantea, M. C., Pastore, F., Pausits, A., Poder, K., Puukka, J., & Velissaratou, J. (2022). *Interim report of the Commission expert group on quality investment in education and training*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2766/37858>

Crouch, C. H., & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics*, 69(9), 970–977. Scopus.

Dancy, M., Henderson, C., & Turpen, C. (2016). How faculty learn about and implement research-based instructional strategies: The case of Peer

Instruction. *Physical Review Physics Education Research*, 12(1), 010110. <https://doi.org/10.1103/PhysRevPhysEducRes.12.010110>

De Backer, L., Van Keer, H., & Valcke, M. (2015). Promoting university students' metacognitive regulation through peer learning: The potential of reciprocal peer tutoring. *Higher Education*, 70(3), 469–486. <https://doi.org/10.1007/s10734-014-9849-3>

Delaney, M. L., & Royal, M. A. (2017). Breaking Engagement Apart: The Role of Intrinsic and Extrinsic Motivation in Engagement Strategies. *Industrial and Organizational Psychology*, 10(1), 127–140. <https://doi.org/10.1017/iop.2017.2>

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences of the United States of America*, 116(39), 19251–19257. Scopus. <https://doi.org/10.1073/pnas.1821936116>

Directorate-General for Education, Y. (2022). *Investing in our future: Quality investment in education and training*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2766/45896>

Dou, R., Brewe, E., Potvin, G., Zwolak, J. P., & Hazari, Z. (2018). Understanding the development of interest and self-efficacy in active-learning undergraduate physics courses. *International Journal of Science Education*, 40(13), 1587–1605. <https://doi.org/10.1080/09500693.2018.1488088>

Duch, B. J. (1996). Problem-Based Learning in Physics: The Power of Students Teaching Students. *Journal of College Science Teaching*, 15(5), 326–329.

Foote, K., Knaub, A., Henderson, C., Dancy, M., & Beichner, R. J. (2016). Enabling and challenging factors in institutional reform: The case of SCALE-UP. *Physical Review Physics Education Research*, 12(1), 010103. <https://doi.org/10.1103/PhysRevPhysEducRes.12.010103>

Foote, K. T., Neumeyer, X., Henderson, C., Dancy, M. H., & Beichner, R. J. (2014). Diffusion of research-based instructional strategies: The case of SCALE-UP. *International Journal of STEM Education*, 1(1), 10. <https://doi.org/10.1186/s40594-014-0010-8>

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the*

*National Academy of Sciences of the United States of America*, 111(23), 8410–8415. Scopus. <https://doi.org/10.1073/pnas.1319030111>

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From Gatekeeping to Engagement: A Multicontextual, Mixed Method Study of Student Academic Engagement in Introductory STEM Courses. *Research in Higher Education*, 53(2), 229–261. <https://doi.org/10.1007/s11162-011-9247-y>

Gómez-Espina, R., Rodriguez-Oroz, D., Chávez, M., Saavedra, C., & Bravo, M. J. (2019). Assessment of the Socrative platform as an interactive and didactic tool in the performance improvement of STEM university students. *Higher Learning Research Communications*, 9(2).

Guarascio, A. J., Nemecek, B. D., & Zimmerman, D. E. (2017). Evaluation of students' perceptions of the Socrative application versus a traditional student response system and its impact on classroom engagement. *Currents in Pharmacy Teaching and Learning*, 9(5), 808–812. <https://doi.org/10.1016/j.cptl.2017.05.011>

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>

Hattie, J. (2008). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement* (1st Edition). Routledge. <https://www.routledge.com/Visible-Learning-A-Synthesis-of-Over-800-Meta-Analyses-Relating-to-Achievement/Hattie/p/book/9780415476188>

Hattie, J. (2012). *Visible Learning For Teachers: Maximizing impact on achievement*. Routledge. <https://visible-learning.org/2013/01/visible-learning-for-teachers-book-review/>

Hattie, J. (2023). *Visible Learning: The Sequel: A Synthesis of Over 2,100 Meta-Analyses Relating to Achievement*. Routledge. <https://www.routledge.com/Visible-Learning-The-Sequel-A-Synthesis-of-Over-2100-Meta-Analyses-Relating/Hattie/p/book/9781032462035>

Hattie, J., Masters, D., & Birch, K. (2015). *Visible Learning into Action: International Case Studies of Impact* (1st Edition). Routledge. <https://www.routledge.com/Visible-Learning-into-Action-International-Case-Studies-of-Impact/Hattie-Masters-Birch/p/book/9781138642294>

Hornsby, D. J., & Osman, R. (2014). Massification in higher education: Large classes and student learning. *Higher Education*, 67(6), 711–719. Scopus. <https://doi.org/10.1007/s10734-014-9733-1>

Koch, A. K. (2017). It's About the Gateway Courses: Defining and Contextualizing the Issue. *New Directions for Higher Education*, 2017(180), 11–17. <https://doi.org/10.1002/he.20257>

Korpas, G. S., Kahrs, M. S., & Andersen, T. H. (2019). Peer learning and concept tests in physics. *Varietas Delectat... Complexity Is the New Normality*, 651–657. <https://researchportal.tuni.fi/en/publications/proceedings-of-the-sefi-47th-annual-conference-2019-varietas-dele>

Leinonen, R., Asikainen, M. A., & Hirvonen, P. E. (2017). Peer discussions in lecture-based tutorials in introductory physics. *Physical Review Physics Education Research*, 13(1). Scopus. <https://doi.org/10.1103/PhysRevPhysEducRes.13.010114>

Lim, W. N. (2017). Improving student engagement in higher education through mobile-based interactive teaching model using socrative. *2017 IEEE Global Engineering Education Conference (EDUCON)*, 404–412. <https://doi.org/10.1109/EDUCON.2017.7942879>

Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., Bulger, M., Campbell, J., Knight, A., & Zhang, H. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, 34(1), 51–57. <https://doi.org/10.1016/j.cedpsych.2008.04.002>

McLean, K. J. (2016). The Implementation of Bring Your Own Device (BYOD) in Primary [Elementary] Schools. *Frontiers in Psychology*, 7. <https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01739>

Medicine, N. A. of S., Engineering, and, Engineering, N. A. of, Affairs, P. and G., Workforce, B. on H. E. and, Education, D. of B. and S. S. and, Education, B. on S., & Degrees, C. on B. and O. in C. 2-Y. and 4-Y. S. (2016). *Barriers and Opportunities for 2-Year and 4-Year STEM Degrees: Systemic Change to Support Students' Diverse Pathways*. National Academies Press.

Meijer, J., Veenman, M. V. J., & van Hout-Wolters, B. H. A. M. (2006). Metacognitive activities in text-studying and problem-solving: Development of a taxonomy. *Educational Research and Evaluation*, 12(3), 209–237. <https://doi.org/10.1080/13803610500479991>

Neuhäuser, M., & Hothorn, L. A. (1999). An exact Cochran–Armitage test for trend when dose–response shapes are a priori unknown. *Computational Statistics & Data Analysis*, 30(4), 403–412. [https://doi.org/10.1016/S0167-9473\(98\)00091-7](https://doi.org/10.1016/S0167-9473(98)00091-7)

Örnek, F., Robinson, W. R., & Haugan, M. R. (2007). What Makes Physics Difficult? *Science Education International*, 18(3), 165–172.

Palfreyman, D., Tapper, T., & Thomas, S. (2011). Series editors' introduction: International studies in higher education. *Accountability in Higher Education: Global Perspectives on Trust and Power*, 1–270. Scopus. <https://doi.org/10.4324/9780203846162>

Raffaghelli, J., Ghislandi, P., Sancassani, S., Canal, L., Micciolo, R., Balossi, B., Bozzi, M., Di Sieno, L., Genco, I., Gondoni, P., Pini, A., & Zani, M. (2018). Integrating MOOCs in physics preliminary undergraduate education: Beyond large size lectures. *Educational Media International*, 55(4), 301–316. <https://doi.org/10.1080/09523987.2018.1547544>

Rasmussen, C., Apkarian, N., Hagman, J. E., Johnson, E., Larsen, S., & Bressoud, D. (2019). Brief Report: Characteristics of Precalculus Through Calculus 2 Programs: Insights From a National Census Survey. *Journal for Research in Mathematics Education*, 50(1), 98–111. <https://doi.org/10.5951/jresmetheduc.50.1.0098>

Rea, L. M., & Parker, R. A. (2014). *Designing and Conducting Survey Research: A Comprehensive Guide* (4th Edition). Jossey-Boss. <https://www.wiley.com/en-us/Designing+and+Conducting+Survey+Research%3A+A+Comprehensive+Guide%2C+4th+Edition-p-9781118767030>

Redish, E. F., Scherr, R. E., & Tuminaro, J. (2006). Reverse-engineering the solution of a 'simple' physics problem: Why learning physics is harder than it looks. *Physics Teacher*, 44(5), 293–300. Scopus. <https://doi.org/10.1119/1.2195401>

Rosenthal, J. A. (1996). Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research*, 21(4), 37–59. [https://doi.org/10.1300/J079v21n04\\_02](https://doi.org/10.1300/J079v21n04_02)

Rudolph, A. L., Lamine, B., Joyce, M., Vignolles, H., & Consiglio, D. (2014). Introduction of interactive learning into French university physics classrooms. *Physical Review Special Topics - Physics Education Research*, 10(1), 010103. <https://doi.org/10.1103/PhysRevSTPER.10.010103>

Schiltz, G., Feldman, G., & Vaterlaus, A. (2019). *Active-learning settings and physics lectures: A performance analysis*. 1286(1). Scopus. <https://doi.org/10.1088/1742-6596/1286/1/012019>

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7(4), 351–371. <https://doi.org/10.1007/BF02212307>

Schunk, D. H., & Zimmerman, B. J. (Eds.). (2012). *Motivation and Self-Regulated Learning: Theory, Research, and Applications*. Routledge.

Scott, P. (1995). *The Meanings of Mass Higher Education*. Open University Press.

Seymour, E., & Hunter, A.-B. (Eds.). (2019). *Talking about Leaving Revisited: Persistence, Relocation, and Loss in Undergraduate STEM Education*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-25304-2>

Shadle, S. E., Marker, A., & Earl, B. (2017). Faculty drivers and barriers: Laying the groundwork for undergraduate STEM education reform in academic departments. *International Journal of STEM Education*, 4(1), Article 1. <https://doi.org/10.1186/s40594-017-0062-7>

Silverthorn, D. U., Thorn, P. M., & Svinicki, M. D. (2006). It's difficult to change the way we teach: Lessons from the Integrative Themes in Physiology curriculum module project. *American Journal of Physiology - Advances in Physiology Education*, 30(4), 204–214. Scopus. <https://doi.org/10.1152/advan.00064.2006>

Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., Eagan, M. K., Jr., Esson, J. M., Knight, J. K., Laski, F. A., Levis-Fitzgerald, M., Lee, C. J., Lo, S. M., McDonnell, L. M., McKay, T. A., Michelotti, N., Musgrove, A., Palmer, M. S., Plank, K. M., ... Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science*, 359(6383), 1468–1470. Scopus. <https://doi.org/10.1126/science.aap8892>

Sturtevant, H., & Wheeler, L. (2019). The STEM Faculty Instructional Barriers and Identity Survey (FIBIS): Development and exploratory results. *International Journal of STEM Education*, 6(1), Article 1. <https://doi.org/10.1186/s40594-019-0185-0>

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., Dunster, G., Grummer, J. A., Hennessey, K., Hsiao, J., Iranon, N., Jones, L., Jordt, H., Keller, M., Lacey, M.

E., Littlefield, C. E., ... Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, 117(12), 6476–6483. <https://doi.org/10.1073/pnas.1916903117>

Wieman, C. E. (2014). Large-scale comparison of science teaching methods sends clear message. *Proceedings of the National Academy of Sciences*, 111(23), 8319–8320. <https://doi.org/10.1073/pnas.1407304111>

Williams, E. A., Zwolak, J. P., Dou, R., & Brewe, E. (2019). Linking engagement and performance: The social network analysis perspective. *Physical Review Physics Education Research*, 15(2), 020150. <https://doi.org/10.1103/PhysRevPhysEducRes.15.020150>

Wood, A. K., Galloway, R. K., Donnelly, R., & Hardy, J. (2016). Characterizing interactive engagement activities in a flipped introductory physics class. *Physical Review Physics Education Research*, 12(1), 010140. <https://doi.org/10.1103/PhysRevPhysEducRes.12.010140>

Wood, A. K., Galloway, R. K., Hardy, J., & Sinclair, C. M. (2014). Analyzing learning during peer instruction dialogues: A resource activation framework. *Physical Review Special Topics - Physics Education Research*, 10(2). Scopus. <https://doi.org/10.1103/PhysRevSTPER.10.020107>

Wood, R., & Shirazi, S. (2020). A systematic review of audience response systems for teaching and learning in higher education: The student experience. *Computers and Education*, 153. Scopus. <https://doi.org/10.1016/j.compedu.2020.103896>

Zhang, P., Ding, L., & Mazur, E. (2017). Peer Instruction in introductory physics: A method to bring about positive changes in students' attitudes and beliefs. *Physical Review Physics Education Research*, 13(1), 010104. <https://doi.org/10.1103/PhysRevPhysEducRes.13.010104>

Zimmerman, B. J., & Schunk, D. H. (2011). Self-regulated learning and performance: An introduction and an overview. In *Handbook of self-regulation of learning and performance* (pp. 1–12). Routledge/Taylor & Francis Group.

## APPENDIX 1

Example of an item administered to the students during a peer learning session.  
The correct answer is in bold.

An aeroplane runs through equal spaces in equal times

- A. only if the velocity vector is constant over time;
- B. only if its trajectory is straight;
- C. only if the tangent component of the acceleration is zero and the centripetal component is constant;
- D. only if the tangent component of the acceleration is zero.**

## APPENDIX 2

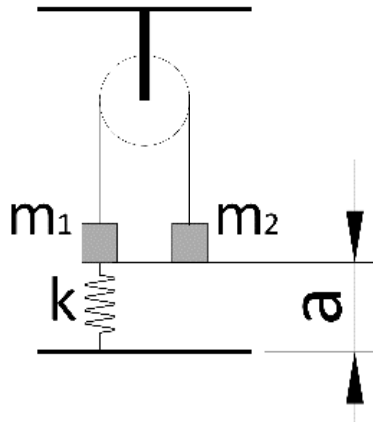
Example of a problem given to the students during their final exam (exercise 1, 21/01/2022)

Two blocks of masses  $m_1 = 1 \text{ kg}$  and  $m_2 = 2 \text{ kg}$  are attached to opposite ends of an ideal rope which can run through the throat of an ideal pulley, like in the figure. Block 1 ( $m_1$ ) is connected to the ground through an ideal vertical spring which has a relaxed length  $h_0 = 10 \text{ cm}$ .

a) Determine the spring constant in such a way that the two blocks are initially at the same height  $a = 15 \text{ cm}$  from the ground.

Block 2 ( $m_2$ ) is lowered vertically by  $5 \text{ cm}$  and then is released:

b) Calculate the acceleration of the two blocks when block 2 is released.



### APPENDIX 3

Example of an item administered to the students during the initial questionnaire. The correct answer is in bold.

The force of static friction exerted by a horizontal floor on a crate at rest on it

- A. has always a magnitude higher than the magnitude of the force of dynamic friction exerted by the horizontal floor on the crate if the latter is moving;
- B. has a magnitude directly proportional to the magnitude of the normal contact force exerted by the horizontal floor on the crate;
- C. is never null;
- D. is opposite to the sum of all other external forces lying on the horizontal plane.**