Article

# "FLIPPED ASSESSMENT": Proposal for a Self-Assessment Method to Improve Learning in the Field of Manufacturing Technologies 

José Díaz-Álvarez ${ }^{1,2, *} \times$ © , Antonio Díaz-Álvarez $^{1(1)}$, Ramiro Mantecón ${ }^{1,3} \mathbf{B}^{(D)}$ and María Henar Miguélez ${ }^{1}$<br>1 Department of Mechanical Engineering, University Carlos III of Madrid, Avda de la Universidad 30, 28911 Leganés, Madrid, Spain; andiaza@ing.uc3m.es (A.D.-Á.); ramiro.mantecon@uc3m.es (R.M.); mhmiguel@ing.uc3m.es (M.H.M.)<br>2 Institute of Innovation in Sustainable Engineering (IISE), College of Science and Engineering, University of Derby, Derby DE22 1GB, UK<br>3 Experimental Mechanics Laboratory, Mechanical Engineering Department, San Diego State University, 5500 Campanile Dr, San Diego, CA 92182, USA<br>* Correspondence: jodiaz@ing.uc3m.es

Citation: Díaz-Álvarez, J.; Díaz-Álvarez, A.; Mantecón, R.; Miguélez, M.H. "FLIPPED ASSESSMENT": Proposal for a Self-Assessment Method to Improve Learning in the Field of Manufacturing Technologies. Educ. Sci. 2023, 13, 831. https://doi.org/ 10.3390/educsci13080831

Academic Editor: James Albright
Received: 10 July 2023
Revised: 11 August 2023
Accepted: 11 August 2023
Published: 15 August 2023


Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).


#### Abstract

Striving toward goal completion and achieving objectives is one of the motors of personal advancement. The path to goal completion is fueled by many reasons, among which motivation stands out as one of the core impulses. Motivation acquires a particularly high relevance in learning, prompting educators to mind its substance when designing not only the material to be imparted but also the approach and the mechanisms to assess knowledge acquisition. The intrinsic nature of motivation might stem from self-realization, thriving in specific goals, or even exploring unknown ground. One of the main teacher-student interactions is the provision of adequate tools to achieve learning outcomes. One of the tools available to teachers is the exercise of extrinsic motivation. This paper proposes and assesses the initial implementation of a student-involved extrinsic motivation method. A pilot group in the Junior year of a Bachelor in Mechanical Engineering program was selected, in which the evaluation system was slightly modified with respect to the system that is normally used. The course selected for the study was a compulsory six European Credit Transfer and Accumulation System (ECTS) course covering production and manufacturing technology. Students were asked to partake in the drafting of questions to assess their own knowledge, hence indirectly increasing their motivation to learn the content. The tentative results obtained with the pilot group appear to be positive and relevant. Students showed a higher engagement during class and reported needing fewer hours of preparation at home ( $32 \%$ reduction). In addition, global satisfaction with the course was improved.


Keywords: self-assessment; flipped assessment; motivation

## 1. Introduction

Historically, assessment and evaluation methods have been a matter of debate and discussion with the aim of improving education and training. According to RodriguezEsteban et al. [1], the concept of evaluation has evolved to encompass not only a final act of proving an outcome but rather a process that is a constituent of the learning system. While not being a new proposal, there seems to be a recent focus on exploring the advantages of "self-assessment" methods. This term, however, has been granted many definitions and has been the topic of numerous reviews [2-5]. Judging from the literature, to establish a complete definition common to the various perspectives of researchers is no simple task. Some authors have defined it as a "wide variety of mechanisms and techniques through which students describe (i.e., assess) and possibly assign merit or worth to (i.e., evaluate) the qualities of their own learning processes and products" [6]. It has also been understood
as "process during which students engage in feedback seeking, evaluation, reflection and making judgements regarding the quality of their learning process and outcomes" [7].

The lack of an unambiguous definition notwithstanding, there does seem to be agreement on the advantageous nature of this practice for students. Exemplary benefits listed by pedagogues include increases in motivation and improved academic results.

The repercussions of self-assessment in subjective aspects, such as the motivation of the students, are difficult to gauge. Attempts to emphasize somewhat more quantitative outcomes have been made, focusing on students' achievement of learning objectives [8,9]. Involving students in their own assessment has been shown to improve their strategies used to prepare for different subjects and, consequently, lead to better grades [10]. Works in the literature showcase a robust relationship between self-efficacy-consequential to these improved preparation strategies-and the academic achievements of students, as presented by Bartimote-Aufflick et al. [11]. Although student-based self-assessment systems are reportedly beneficial in the long term, it is crucial for educators to make specific efforts to determine what criteria are deemed objective to serve as the basis for student selfevaluation. Additionally, self-assessment processes may trigger unsought feelings of intimidation or scrutiny in the students, which teachers must tackle by creating a conducive learning environment that counters those feelings with a sense of safety and support [12]. As mentioned above, there are studies indicating that students who better understand evaluation systems acquire greater learning skills [13]. A similar inference is made about using a peer assessment method with a common rubric among students [14]. It is worth noting that meta-analyses have highlighted that interventions involving self-assessment and explicit feedback on student performance had a significantly greater impact than those without explicit feedback, suggesting that the provision of external feedback is an important framework for successful self-assessment [15]. Furthermore, its impact on a student's ability to take responsibility for their own learning has also been tested [16]. Other studies explored a different involvement of the students in the evaluation using studentsproposed evaluation rubrics. This method also reported improved results achieved by the students [17].

On a parallel note, the outcome of implementing self-assessment methods is twofold. The improvement in student performance aside, benefits appear in the teacher performance as well, bettering their pedagogical skills [18].

### 1.1. Self-Assessment of Student Motivation

While, as stated above, the subjective aspects provoked in students are difficult to gauge, there are a number of published studies that approach the psychological consequences of engaging in self-assessment practices. In diverse educational environments, self-evaluation has been observed to generate a beneficial influence on student motivation. By allowing students to evaluate their own advancement and learning, they tend to exhibit increased enthusiasm and motivation to enhance their academic performance [19,20]. According to Zimmerman et al. [21], self-assessment can help students develop a sense of control over their own learning, which in turn leads to increased motivation and a greater willingness to take risks. Furthermore, self-assessment can help students identify their strengths and weaknesses, which can be used to inform their future learning goals and strategies [22].

In addition to promoting student motivation, self-assessment has also been linked to improved academic performance. A study by Andrade et al. [23] found that when students were given the opportunity to self-assess their work, they were able to identify and correct their own errors, leading to higher quality work and improved grades. Similarly, Black et al. [24] found that self-assessment can lead to improved academic achievement, particularly when it is incorporated into regular classroom practices. These findings suggest that self-assessment can be a valuable tool for promoting both student motivation and academic success.

A comparative study was conducted to determine the effectiveness of incorporating self-assessment and conventional methodologies to increase student academic motivation. The study proved that students who were engaged in self-evaluation displayed higher levels of motivation compared with their peers who relied on traditional methods [25]. In the context of COVID-19, the self-assessment benefit was also tested as a worthy tool to keep students involved, particularly when studying remotely [26].

In addition, there is a reciprocal relationship between student motivation and selfassessment. In other words, higher levels of student motivation can encourage the practice of self-assessment, while increased engagement in self-assessment can also enhance student motivation [27].

### 1.2. Student Motivation

Motivation, defined as "a process that requires students to perform physical or mental activities for achieving their goals" [28], is essential for a student's academic performance, with some authors even highlighting it as the most important aspect [29]. It is worth noting that studies have shown a direct effect of self-assessment on increasing intrinsic motivation (referring to motivation that originates within the individual), social development, and well-being [30], as well as an increase in extrinsic motivation (referring to motivation caused by external factors) due to self-assessment practices [31]. Other studies have concluded that the inclusion of self-assessment contributes to an improvement in both external and internal motivation.

Therefore, motivation is a critical factor in student learning and achievement. Recent research has highlighted the importance of intrinsic motivation, which is characterized by a genuine interest in and enjoyment of the learning process. Studies have shown that when students are intrinsically motivated, they are more likely to engage in profound learning strategies, such as elaboration and reflection, which can lead to improved academic performance [32]. Additionally, intrinsic motivation has been linked to positive outcomes such as greater creativity, persistence, and well-being [30].

On the other hand, extrinsic motivation, which is driven by external rewards or punishments, has been found to have a more limited impact on student learning. While extrinsic motivation can be effective in the short term, it often leads to a focus on achieving grades or other external rewards rather than on the learning process itself [33]. This can result in a lack of engagement and a decreased willingness to take risks or explore new ideas. Therefore, it is important for educators to focus on promoting intrinsic motivation in their students using strategies such as providing opportunities for autonomy, competence, and relatedness [32].

Furthermore, research has identified a clear correlation between the increase in motivation and the provision of good feedback to students in conjunction with self-assessment [34]. Timely, specific, and actionable feedback can assist students in identifying areas for improvement and setting goals for themselves, which can subsequently enhance their motivation to learn [35].

### 1.3. Research Questions

As mentioned in the preceding paragraphs, the practice of self-assessment is often accompanied by an increase in student motivation. One of the consequences of this heightened motivation is a higher success rate in student learning [36-38], which also can be found in meta-analyses supporting this finding [39,40]. In addition, some studies have even highlighted the measurement of student motivation as a variable for assessing the likelihood of success in their education [41].

This study explores a strategy for self-evaluation engagement, named "flipped assessment". It has been previously used in different works to refer exclusively to procedures where the role of the student has been changed into that of the teacher only in the field of evaluation [42,43]. The method used in this study addressed four research hypotheses. The first hypothesis concerned whether student self-assessment directly and positively affects
student intrinsic learning motivation. The second hypothesis concerned the extent to which it contributes to improving student perception of teaching and the subjects themselves. The third hypothesis explored how classroom dynamics could be improved, with an increase in student participation and interest in the subject using of self-assessment systems. Finally, the fourth hypothesis concerned the validity of asserting that self-assessment brings better academic results.

## 2. Materials and Methods

### 2.1. Scenario of this Study

The situation of this study is essential to understand its outcomes. This study was carried out for the first time in the 2022-2023 academic year during the first quarter of the year. The chosen subject is included in the bachelor's program in Mechanical Engineering offered by the Carlos III University of Madrid. It is a compulsory course taught during the junior year with a weight of 6 ECTS credits. Its contents cover the production and manufacturing technologies field with a theoretical-practical approach. Part of the course focuses on the theoretical basis of the field, while the remaining part is oriented toward solving application problems.

Some studies indicate that the impact of self-assessment practices varies across different school levels, being more relevant for older students [6]. This is consistent with the psychological aspects of these practices, which require a degree of maturity to have a profitable repercussion. Such a notion is of high relevance when pondering the analysis performed in this study. In this aspect, it is noteworthy to remember the course under study in this work is taught to students in their junior year.

The selected subject is taught by distributing material in two weekly face-to-face sessions. Theoretical concepts are presented and discussed in one of the sessions, where students are grouped in what is called the "Master Group", with a maximum of 120 students. In the remaining weekly session, students are grouped into the so-called "Small groups" with a maximum of 40 students, where the practices and problem-solving associated with the subject are carried out. In addition, the selected subject has the particularity of being taught in both English and Spanish.

The evaluation system utilized in the subject is a coupled strategy of continuous and summative evaluation. The continuous evaluation is distributed throughout the course, and it involves two tests and a project. Multiple-choice tests aim to gauge a student's grasp of the concepts discussed in class. The project, on the other hand, is carried out in groups of 4 students, evaluating their ability using a practical application of the material taught. In all, the weight of the continuous evaluation with respect to the total grade obtained in the subject is $45 \%$, equally divided among the three deliverables. In addition, the students must carry out different practical sessions throughout the course that are graded in an absolute manner (pass/fail). Complementary to the continuous evaluation is the final exam of the subject. It assesses a student's ability to perform problem resolution and his/her global understanding of the course, and it has a weight of $55 \%$ of the overall grade.

A student's evaluation of the subject is carried out in the final period of the course. This evaluation consists of a sort of student feedback and is completed prior to taking the final exam when the students already know the result of the continuous part of the assessment. The feedback is obtained using an anonymous and voluntary survey, in which the students have an opportunity to evaluate concepts related to the teaching and the organization of the subject. In the survey, they answer a series of questions on a Likert scale—from "completely agree" to "completely disagree"-covering organizational and effectiveness aspects of the course. Worded feedback is also encouraged in the form of comments to improve the performance of the teachers and the organization of the lectures.

The results of these surveys are used by the teachers and the Department to implement modifications to the strategies used to teach the subjects. The implementation of the selfassessment strategy, which will be addressed subsequently, stems from some insights gained using the surveys and comments from years prior to the realization of this work.

The motivation of the students seems to be lower in those groups with the course imparted in English as compared to those taught in Spanish. This difference is also visible in the results of the knowledge evaluations (both continuous and summative) of the subject.

A frequent comment in the student surveys is a feeling of dissonance between the assessment goals of the partial tests and the knowledge they acquired regarding the material studied. There appears to be a mismatch between teacher and student expectations of what reflects material knowledge. This feeling among students is recurrent regardless of the courses taken [2,3] and is one of the motivations for conducting this study.

### 2.2. Implementation of the Methodology

As discussed in the previous subsection, according to surveys and comments from years prior to the realization of this work, it was found that the students who attend the course in groups taught in English are not as motivated as those taught in Spanish, noting this in the academic results obtained in the subject, with a visible decrease in the grades obtained in the English groups.

The evaluation of the effects and consequences of the self-assessment strategy was performed on a pilot group, in which the evaluation system was modified. The expectation of the outcomes was focused on the student grades, concerning quantitative observations, motivation to learn, and general satisfaction with the subject and teachers, concerning subjective and qualitative aspects.

### 2.2.1. Quantitative Data

The new evaluation system stems from comments in the students' satisfaction surveys. As stated previously, a commonly reported feeling among students concerned the multiplechoice tests that were part of the continuous evaluation of the subject. The tests were made of twenty questions posed by the teachers involved in teaching the course. To address the dissonance between students and teachers regarding what questions reflect their knowledge, the new evaluation system incorporated a shared responsibility in the creation of the multiple-choice test. Out of the total twenty questions, only ten would be created by the teachers, while the remaining half were posed by the students. To select the questions that would be included in the final exam, students were asked to come up with four questions each, creating a large pool of possible items that were corrected by the teachers of the course. Finally, ten were included in the final exams. A fairness mechanism was implemented to mitigate possible unethical practices, such as students sharing sensitive information. Such a mechanism consisted in disregarding the students-posed questions that were correctly answered by more than $75 \%$ of the group-a success rate far superior to previous courses. This mechanism was only provisional, as the evaluation system is being probed and implemented, and it aimed to palliate the effects of a misplaced focus on grade achievement, the fruit of inappropriate motivation techniques [27].

The pilot group chosen for the implementation of the new evaluation system was the group that was taught the course in English. This group consisted of students including both Spanish nationals and visiting exchange students, who presented a worse history in terms of student motivation, satisfaction with the subject and teachers, and obtained grades, in general. In the academic year of interest, the group totaled 29 students. The outcomes of the pilot group with the new evaluation system were compared to the reference groups taught in Spanish (in which the evaluation was not modified), amounting to a total of 166 students. Additionally, the results were compared to the information from the previous year for all groups (Spanish and English).

### 2.2.2. Qualitative Data

The results of the student surveys were used as indicators for self-perceived learning experiences. Specifically, the students were surveyed concerning their satisfaction with the contents of the course, its organization, and teacher performance. For reference, the teacher in charge of teaching the subject to the English group was the same in the two years
of this study. The sections contained in the survey that allow students to evaluate teacher performance are as follows: the teacher stimulates learning appropriately; the teacher plans and coordinates the class well and clearly exposes the explanations; the teacher adequately resolves doubts and guides students in the development of tasks; and in general, I am satisfied with the teaching of the teacher of the subject. In a similar fashion, to assess the subject in more general terms, the students also indicate their global satisfaction with the subject, estimate the number of hours per week devoted to it (classes aside), and value the increase in knowledge, competencies, and skills acquired.

To gather an estimation of the degree of benefits in the implementation of the new evaluation system, the scores obtained in these questions in the student survey were compared to the scores from previous years. Thus, an indication of whether the path was correct was extracted from these initial results. These data were complemented with the number of passing students and the average grade obtained in the pilot group.

In summary, the methodology followed presents a mixed approach (quantitative and qualitative). To discuss the first, second, and third hypotheses, a survey was used, which the students completed at the end of the course, answering a series of questions about their perception of the subject (the questions are included in Table 1), as well as their opinion of the teachers who taught the subject. On the other hand, to address the fourth hypothesis, an approach based on the grades obtained by the students in different assessments throughout the course was proposed. All these results serve as guides, extracted from the outcomes of this initial implementation of the new evaluation system, to gauge the effectiveness of the assessment method. While tentative in nature, due to the very early stages of the application of the method, they are notably useful in understanding whether its repercussions could point to beneficial progress in the teaching of the course. A comparison was made between the results from the pilot group, taught in English, and two control groups. On the one hand, the pilot group was compared to the rest of the groups of the subject taught in Spanish in the same academic year. On the other hand, the pilot group was compared to the results from the previous year in the English group. To minimize language bias and temporal bias, mainly relative values with respect to time and the pilot group will be discussed.

Table 1. Relative variation in subject perception provided by the students. Highlights in green and red represent positive and negative outcomes, respectively.
$\left.\left.\begin{array}{cccc}\hline & 2021 & 2022 & \begin{array}{c}\text { Percentage of Variation, Values Obtained } \\ \text { in 2022 with Respect to 2021 }\end{array} \\ \text { Pilot with Respect } \\ \text { to All Groups }\end{array}\right) \begin{array}{c}\text { Pilot with Respect } \\ \text { to All Groups }\end{array}\right)$

## 3. Results

Throughout this section, the results of this initial implementation of the flippedassessment system will be discussed. It is imperative to note that, while the results are fairly interesting and indicative of whether the system is advantageous or counterproductive, they are not fully conclusive. Their tentative nature should be accounted for when understanding the effectiveness of the assessment system. Future work and research lines will be proposed and discussed later in this article, which should assist in complementing the initial findings herein reported.

### 3.1. Student Self-Perceived Measurements of The Learning Experience

This subsection presents the results collected from the surveys carried out by the students about their self-perception of the teaching performance and their degree of satisfaction with the subject.

Table 1 summarizes the relative variation in the students' perception between the years 2021 and 2022. It is worth noting that the "pilot group" refers to the English-taught group irrespective of the year. The survey is divided into student perception of teacher performance and the global organization of the subject. The survey results for the pilot group are compared to those of all groups within the same academic year. Additionally, the variation within the pilot group for the two years under study is reported.

The results seem to indicate a positive trend after the implementation of the new assessment system. Taking as a reference the scenario in the year 2021, the pilot group presents significantly lower student satisfaction than the rest of the groups. Coincidentally with the new assessment system in 2022, the indicators for satisfaction appear to mutate quite positively from the previous year. As can be seen in the 2022 column, the indicators corresponding to the teacher performance in the pilot group are very close to those for the other groups. This apparent change in the perception of teacher performance would seem to be matched with the global perception of the subject, which also appears to improve from 2021 to 2022. In addition to the conceived reversion of the worse results for the pilot group, the questions about global satisfaction with the subject would seem to be more positively answered with the implementation of the flipped-assessment system. Variations in the table are highlighted in red and green color to aid visualization. Red represents a negative outcome, such as low satisfaction with the course or a very high number of hours devoted to the subject, as opposed to positive outcomes, which are highlighted in green. Differences below $5 \%$ were not considered relevant enough, given the initial nature of these results, and are therefore not highlighted.

A few more inferences from Table 1 are noteworthy. In the matter of teacher performance, apparently, the reference groups (i.e., the Spanish-taught groups) do not present a relevant shift between the years 2021 and 2022. This permanence motivates the assumption that the changes in the results in the pilot group are related to a beneficial connection between the teacher and the students, although these results still need to be further sustained. This insight is supported as well by the negligible difference between the pilot and the reference groups in the year 2022. There seems to be an important jump for the pilot group between the two years of the study, with a reported improvement in the way the teacher plans and coordinates the class by $37 \%$. Teacher performance is also positively assessed in the way a teacher resolves doubts and guides students in the development of tasks, with a change of $33 \%$. These results could arise from a parallel effect of the assessment system. As the students had to formulate their own questions for the tests, they needed to strive to follow the subject and understand the contents taught to a much higher degree. This situation might have been the origin of engaging in a more active interaction in class, which in turn gave the teacher an opportunity to resolve more doubts. This increased engagement was reported by the professor teaching the subject, although there was no quantifiable indicator for it. Implementing such an indicator would have disturbed the natural flow of the class and might have been counterproductive. However, a parallel result supports this observation: the number of emails sent to the teacher with doubts about the contents
was reduced by $5 \%$ in the pilot group. Additionally, visits to the professor's office hours decreased by $7 \%$. The students would then have perceived a responsive attitude in the teacher concerning the way in which they solved doubts during class, as well as retaining the students' attention, possibly making the students perceive better planning of the lesson insofar as their comprehension of the topic was better. This initial finding is also supported by the general satisfaction with the teacher in the pilot group, which increased by $8 \%$ with respect to the previous year.

One more comment should be made about the results of the student surveys in this first year of the new assessment system. While the students seemingly perceived their increase in knowledge and competencies was very positive ( $22 \%$ increase with respect to the previous year), the overall estimation of the time devoted to the subject was lower than previously reported (decrease of $-32 \%$ ). However not conclusive in nature, these early observations could hint at the lessons being more advantageous when the students need to extract viable questions to evaluate their own knowledge.

### 3.2. Student's Grades

A more quantitative analysis of the repercussions of the new flipped-assessment method is presented in this subsection. As previously stated, a student's involvement in the evaluation of the subject was made possible by formulating questions for the two mid-term exams that are part of the continuous evaluation. It is worth noting that a temporary contingency method for the use of unethical practices was implemented, which consisted of disregarding the validity of questions posed by the students if they were answered correctly by a suspiciously large portion of the class.

Focusing on the ten questions posed by the students in each partial test, some interesting observations arise. In the first test, on average the students performed $125 \%$ better on the students-posed questions compared to those formulated by the teachers. However, the opposite result was noted on the second multiple-choice test, where students obtained on average $10 \%$ lower grades on their questions than on the teachers' questions. Furthermore, comparing the results of the second to the first test, student performance was on average $31 \%$ poorer on questions they formulated, while student performance on questions posed by the teachers increased by $71 \%$.

Given the tentative nature of the conclusions after one year of system implementation, two comparisons were established in order to acquire a somewhat more complete view of the impact of the new evaluation format on the academic results obtained by the students. On the one hand, the results obtained by the pilot group were compared with the results obtained in the year before the implementation of the new evaluation system. On the other hand, to have a reference point, an analysis of the variation in the results obtained by the rest of the groups compared to the previous year before the implementation of the new system was also included.

In Figure 1, the relative changes in the resulting average grade for each of the evaluative assessments can be appreciated for both groups analyzed with respect to the year before the implementation of the self-assessment (pilot group and rest of the groups). For the first test exam, an increase in the average grade obtained by the pilot group of $73 \%$ can be observed, while in the rest of the groups, only an increase of $18 \%$ is observed. As discussed in the following section, this could have been due to dishonest behavior in the pilot group student. However, on the second test, the variation found in the average grade obtained in the pilot group and in the rest of the groups was identical and equal to $-21 \%$, which points to reasons unrelated to the impact of the evaluation system but rather more related to the difficulty of the proposed exam. The next analyzed assessment is the evaluation of the subject's project, which is part of the continuous evaluation. In this case, it can be perceived that the resulting average grade in the pilot group is slightly lower than that obtained by the declared group in the previous year, with no variation in the average grade for the rest of the groups. In the case of the pilot group, it should be noted that the project is an open project where students indirectly set its difficulty based on some boundary conditions that
they define themselves. An apparent explanation of the results is related to the effect of the new evaluation system. Deriving from higher engagement, the students might have been more curious and eager to face new challenges, which led them to set an increased difficulty in their projects.


Figure 1. Percentage of grade variation obtained in 2022 compared with those obtained in 2021.
Among all these initial findings, the most surprising and relevant result is the average grade obtained by the pilot group on the final exam of the subject, which increased by $29 \%$, far from the decrease in the other groups of the subject ( $-1 \%$ ). A reminder should be made, pointing to the relevance of this observation. The new evaluation system involves the students in the creation of the mid-term tests but excludes them from the creation of the final exam of the course. The final exam is developed by the teachers and is common and carried out by all the groups at the same time. Regarding the overall average grade for the subject, it should be noted that the pilot group surpassed the average grade obtained by the equivalent group in the previous year by $14 \%$, while in the rest of the groups, only an increase of $3 \%$ occurred.

One of the original objectives of this work was to reduce the gap in training results between the English-taught pilot group and the Spanish-taught groups. To analyze more clearly the impact of the introduction of the self-assessment system on the pilot group, Figure 2 shows the relative differences in the mean value of the results obtained in the pilot group and in the rest of the groups in the year of implementation of the pilot project (2022) and the previous year.

Regarding the first test exam, it can be observed that in the pilot group, grades were obtained on average $31 \%$ lower than those obtained by the rest of the groups prior to the implementation of the self-assessment evaluation. However, with the implementation of the new evaluation system, this distance appears to be shortened, reaching values of practically identical mean grades ( $1 \%$ higher in the pilot group). On the second multiplechoice test, the distance in terms of academic results remained unchanged compared to the year before the implementation of the new evaluation system. Concerning the evaluation of the subject's project, there is no perceived relevant difference in the average grade obtained by the groups (only a $1 \%$ of difference). Analyzing the mean grades obtained in the pilot group and the rest of the groups, it can be seen that not only has the pilot group improved its results compared to the rest of the groups but it surpassed them, obtaining an average grade of $8 \%$ higher than the average grade achieved by the rest of the groups. The final result of the impact of the pilot project translates into the fact that the average grade in the subject obtained by the pilot group went from being $7 \%$ lower in the pilot group to $3 \%$ higher compared with the rest of the groups.


Figure 2. Percentage of grade variation obtained in the pilot group compared with those obtained in the rest of the group for the years 2021 and 2022.

## 4. Discussion

This section will expand on the relevance of these initial findings and their implications, setting the ground for complementary work that should sustain these inferences and test these tentative observations with more conclusive analyses. This section is subdivided into the four hypotheses described at the beginning of this work, with an additional subsection concerning future work.

### 4.1. Effect of Student Self-Assessment on Intrinsic Learning Motivation

As mentioned in previous sections, the first situations of self-assessment in the pilot group were disruptive with respect to the expectation. The first mid-term exam that incorporated student participation resulted in a large difference in the success rate of the students from their own questions to the questions posed by the teacher. Specifically, the average grade on the student questions was $125 \%$ higher than on the teacher questions. Throughout this work, a concern has been raised a few times about possible unethical behavior in the formulation of the students' pool of questions, triggered by the commonly accepted consequence of extrinsic motivation for seeking a high grade above knowledge acquisition. A similar observation was previously made by other researchers, who claimed that when students had the opportunity to self-assess their own knowledge and that grade affected their final grade, they showed a tendency to assign themselves higher grades than those assigned by teachers [44]. In our study, it was observed that students chose to divide into groups to study and formulate their questions. The methodology used by the students in the first test exam favored those who were not initially as keen to learn and were rather motivated by grades, as they were able to obtain better results in the half of the exam that was student-formulated (as they could know most of the answers beforehand). However, this behavior was reverted on the second mid-term exam, in which the success rate in the questions posed by the teachers increased by $71 \%$. This tendency reversion was more strongly noted on the final exam, as stated previously, where students were not involved whatsoever in its design, and still, the performance was increased with respect to the references. The global outcome of the results, while still preliminary, seems to point toward the increase in student motivation to learn and properly acquire knowledge, closing the perceived gap between teacher and student expectations of the knowledge assessment.

Therefore, considering these initial results, the first hypothesis would seem to be supported by the outcomes of this pilot group, showing a positive effect of the self-assessment system on the intrinsic motivation of the students.

### 4.2. The Contribution of Student Self-Assessment to Improving Their Perception of the Teaching Work and the Subject

A relevant change induced by the implementation of the self-assessment system was a subjective perception noted by the teacher related to student engagement. During the lessons, there was a noted increase in the attention paid by the students, and in their participation engaging in interactions with the teacher and among themselves. Such an outcome is consistent with other studies previously published [45]. Additionally, the teacher's perception of this change was met with the same feeling by the students, as shown in the surveys submitted. The student's evaluation of the teacher assigned to the pilot group increased by more than $30 \%$ compared to the previous year. On a similar note, students reported a generally increased feeling of acquiring knowledge and competencies. The answers to this question in the student survey were $22 \%$ more positive in the year of implementation of the new system than in the previous year. All of this comes together with the estimated time commitment in the year 2022, when students reported having devoted $32 \%$ less time than in the previous year, which would point to enhanced classroom achievement relating to the higher student engagement reported by the teacher.

The results herein mentioned would seem to be in line with the hypothesis under discussion. By being involved in deciding what questions would be used to test their grasp of the material, students appear to be more engaged during the teaching of the lesson. This increased engagement is realized by asking for further clarifications, answering follow-up questions posed by the teacher, and interacting with each other during classroom time. Judging from the initial results of this study, this enhanced classroom performance could be related not only to the overall achievement of the students in their examinations but also to an improved perception of the teaching strategies and of the subject. In other words, this might be generating not only a better understanding of the material but also a kinder feeling toward the subject. These results are supported by other researchers, who point out that self-assessment also increases student efficiency [11,46].

Therefore, the preliminary study here reported would seem to confirm the second hypothesis, stating that student self-assessment helps improve the student's perception of the subject and the teaching.

### 4.3. Influence of Self-Assessment on Student Engagement and Classroom Dynamics

As stated above, on the first multiple-choice test, a large difference was noted in student performance between the questions formulated by students and the questions formulated by teachers. After completing this exam, the students were informed of their grades. The large deviation between their achievements in each part made the students reflect and take full responsibility. As hinted at earlier, this points to a psychological aspect of the evaluation. As extrinsic motivation is likely to promote behaviors focused on achieving higher grades instead of grasping better knowledge, there is a possibility that it might trigger dishonest behaviors. However, the maturity of the students is a key factor in the efficiency of self-assessment methods. On these lines, a more mature and responsible reaction is triggered by confronting students with their apparent malpractices. This side effect of the new methodology was also reported by other researchers in similar studies [47,48]. Additionally, it caused students to have more critical and profound thinking, modifying their approach to the lessons and the preparation for the second multiple-choice test. The change in approach and dynamics was seen in the results for this second test, where students performed better on the questions posed by the teachers than on the student-formulated ones. The possible explanation for this observation is twofold. Firstly, the students interacted more proactively in the classroom and exhausted the clarifications needed to better grasp the material. Secondly, they explored more profound and complex aspects of the concepts, in order to formulate more appropriate questions. This would have yielded an overall extraordinary performance in answering the teacher's evaluation questions.

A preliminary conclusion can be drawn from this result. After an initial setback, probably related to an extrinsic-motivation mindset, the dynamics of the class changed
radically, greatly improving student efficiency. This might hint at asserting the third hypothesis: self-assessment appears to have a positive influence on student engagement and classroom dynamics. This observation is in accordance with similar works by other researchers [49,50].

### 4.4. Relationship Between Self-Assessment and Better Academic Results

The discussion of this last hypothesis is hindered by a comment already made throughout this paper, but which is, nonetheless, worth revisiting. While intrinsic motivation (i.e., a student's will to learn) is commonly conceived as a more relevant impulse toward the beneficial learning of the material, it is difficult to gauge. On the other hand, extrinsic motivation (i.e., rewards or punishments) is likely to trigger a non-desired focus on achieving higher grades, which in turn might sometimes lead to unethical behavior. Concerns can be raised about coupling an extrinsic-motivation mindset with self-assessment methods. As discussed previously, a provisional contingency method was devised for this study. However, the maturity of the group in which self-assessment is implemented contributes highly to the good practice of the method. For this reason, in this subsection, the focus is shifted toward the second and the last examination in the course, disregarding the first test, in which student performance was drastically higher on their own questions than on the teachers' questions.

On the second mid-term test, the grades obtained by the students on the part of the exam formulated by themselves were similar to the grades corresponding to the part of the exam formulated by teachers. Similar results have been reported by other researchers in similar scenarios [2,3,51]. Students obtained on average $10 \%$ worse grades on the part of the exam formulated by themselves. It is worth indicating that student participation in formulating their own questions triggers them to better understand the material, solve doubts that arise during the preparation of the questions, and better comprehend the evaluation system. This fact was also highlighted by other researchers [52]. It is true that the effect of self-assessment is more notable once students have faced this format for the first time and have received feedback, and that could also explain the grades obtained on the second exam. This has the potential to increase student involvement and motivation, resulting in improved academic achievement [53] as demonstrated by the good average subject grades obtained by students compared to the year before the implementation of the new evaluation system ( $14 \%$ higher). Additionally, it is worth noting that the performance of the pilot group also increased on the final exam, in which they did not partake in creating the questions. This fact follows the results of the second mid-term test, in which students performed notably well on the teacher-posed questions.

The initial outcomes of this early-stage implementation of the self-assessment system seem to hint toward the validity of the fourth hypothesis. Evaluation systems that involve and incorporate students seem to result positively in the understanding of the material, which in turn yields better academic results.

### 4.5. Limitations and Future Works

The results presented in this study are not concluding in nature and should be reviewed while considering their limited scope. They are the outcomes of a study in its early stages. Nevertheless, although tentative, the results are positive and provide interesting insights into how self-assessment could affect the classroom and improve student learning. This, along with the fact that this study was performed during a finished academic year, are the elements that point to a partially concluded work that is suitable for reporting in an article.

In accordance with these positive results obtained, some further work needs to be performed in order to confer these insights with a more conclusive nature. A largerscale study should be designed, including different types of subjects, a greater number of pilot groups, and over a longer period of time. Additionally, studying the effect of self-assessment practices on repeating students would be of the utmost interest. Students
who have previously failed the subject are evidence of a failed engagement attempt, that can be attributed to a number of causes. Self-assessment practices might improve student involvement in understanding the subject and could be a key element to their success when retaking the course. Therefore, in future work, a differentiation between first-time and repeating students would be very interesting.

## 5. Conclusions

This study presented both quantitative and qualitative results that allow us to relate students' self-assessment to their motivation to learn, and in turn, to the academic results obtained and the perception they have of the quality of the education received. While the study is in the early stages and the results are still of a tentative nature, some concluding remarks can be drawn that help shape the future in this line.

The main concluding remarks of these initial findings are as follows:

1. The first hypothesis appears to be valid: self-assessment seems to help motivate students, even in the presented case where students do not necessarily have to answer their own questions but those of their classmates.
2. The second hypothesis is also apparently well-posed. On average, students appear to be more satisfied with the subject when the self-assessment system is implemented. Students positively value the teacher's work in all areas, highlighting how they organize classes, explain content, and resolve doubts. Additionally, they perceive that they have acquired more knowledge in a more time-efficient manner.
3. In the matter of classroom dynamics and student engagement, self-assessment might promote a better learning environment and greater camaraderie, which would support the third hypothesis.
4. Finally, the hypothesis regarding academic results also seems to be valid. In quantitative terms, the implemented self-assessment system appears to have allowed for an increase in student knowledge, resulting in higher academic results on average than before its implementation.
These remarks would benefit from further work confirming their initial validity, exploring the possible extrapolation to subjects of different fields, and including a larger number of groups and a focus in different kinds of students.

Author Contributions: Conceptualization, J.D.-Á.; Methodology, J.D.-Á. and A.D.-Á.; Formal analysis, J.D.-Á., A.D.-Á. and R.M.; Data curation, J.D.-Á.; Writing-review \& editing, J.D.-Á., A.D.-Á., R.M. and M.H.M.; Project administration, J.D.-Á.; Funding acquisition, M.H.M. All authors have read and agreed to the published version of the manuscript.
Funding: Authors gratefully acknowledge the financial support through the grant PID2020-118480RBC22 funded by MCIN/AEI/10.13039/501100011033.

Institutional Review Board Statement: Not applicable.
Informed Consent Statement: Not applicable.
Data Availability Statement: The raw/processed data required to reproduce the above findings cannot be shared at this time due to legal/ethical reasons.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of this study; in the collection, analyses, or interpretation of data; in the writing of this manuscript; or in the decision to publish the results.

## References

1. Rodríguez Esteban, M.A.; Frechilla-Alonso, M.A.; Saez-Pérez, M.P. Implementación de la evaluación por pares como herramienta de aprendizaje en grupos numerosos. Experiencia docente entre universidades = Implementation of the evaluation by pairs as a learning tool in large groups. Teaching experience between universities. Adv. Build. Educ. 2018, 2, 66. [CrossRef]
2. Leach, L. Optional self-assessment: Some tensions and dilemmas. Assess. Eval. High. Educ. 2012, 37, 137-147. [CrossRef]
3. Brown, G.T.L.; Andrade, H.L.; Chen, F. Accuracy in student self-assessment: Directions and cautions for research. Assess. Educ. Princ. Policy Pract. 2015, 22, 444-457. [CrossRef]
4. Panadero, E.; Brown, G.T.L.; Strijbos, J.W. The Future of Student Self-Assessment: A Review of Known Unknowns and Potential Directions. Educ. Psychol. Rev. 2016, 28, 803-830. [CrossRef]
5. Andrade, H.L. A Critical Review of Research on Student Self-Assessment. Front. Educ. 2019, 4, 87. [CrossRef]
6. Panadero, E.; Jonsson, A.; Strijbos, J.W. Scaffolding Self-Regulated Learning Through Self-Assessment and Peer Assessment: Guidelines for Classroom Implementation. Enabling Power Assess. 2016, 4, 311-326.
7. Yan, Z.; Brown, G.T.L. A cyclical self-assessment process: Towards a model of how students engage in self-assessment. Assess. Eval. High. Educ. 2017, 42, 1247-1262. [CrossRef]
8. Mendoza, N.B.; Yan, Z. Studies in Educational Evaluation Exploring the moderating role of well-being on the adaptive link between self-assessment practices and learning achievement. Stud. Educ. Eval. 2023, 77, 101249. [CrossRef]
9. Yan, Z.; Lao, H.; Panadero, E.; Fernández-Castilla, B.; Yang, L.; Yang, M. Effects of self-assessment and peer-assessment interventions on academic performance: A meta-analysis. Educ. Res. Rev. 2022, 37, 100484. [CrossRef]
10. Panadero, E.; Alqassab, M. Cambridge Handbook of Instructional Feedback; Cambridge University Press: Cambridge, UK, 2018; ISBN 9781316832134.
11. Bartimote-Aufflick, K.; Bridgeman, A.; Walker, R.; Sharma, M.; Smith, L. The study, evaluation, and improvement of university student self-efficacy. Stud. High. Educ. 2016, 41, 1918-1942. [CrossRef]
12. You Searched for How to Successfully Introduce Self-Assessment in Your Classroom-THE EDUCATION HUB. Available online: https:/ /theeducationhub.org.nz/?s=How+to+successfully+introduce+self-assessment+in+your+classroom (accessed on 12 August 2023).
13. Hearn, J.; McMillan, J.H. Student Self-Assessment: The Key to Stronger Student Motivation and Higher Achievement. Educ. Horiz. 2008, 87, 40-49.
14. Mphahlele, L. Students' Perception of the Use of a Rubric and Peer Reviews in an Online Learning Environment. J. Risk Financ. Manag. 2022, 15, 503. [CrossRef]
15. Yan, Z.; Wang, X.; Boud, D.; Lao, H. The effect of self-assessment on academic performance and the role of explicitness: A meta-analysis. Assess. Eval. High. Educ. 2023, 48, 1-15. [CrossRef]
16. Larsari, V.N.; Dhuli, R.; Koolai, Z.D. An Investigation into the Effect of Self-assessment on Improving EFL Learners' Motivation in EFL Grammar Achievements. J. Engl. A Foreign Lang. Teach. Res. 2023, 3, 44-56.
17. Chambers, A.W. Increasing Student Engagement with Self-Assessment Using Student-Created Rubrics. J. Scholarsh. Teach. Learn. 2023, 23, 96-99.
18. Self-Assessment as a Tool for Teacher's Professional Development—LessonApp. Available online: https:/ /lessonapp.fi/self-assessment-as-a-tool-for-teachers-professional-development/ (accessed on 9 August 2023).
19. Zimmerman, B.J. Motivational Sources and Outcomes of Self-Regulated Learning and Performance. In Handbook of Self-Regulation of Learning and Performance; Routledge: New York, NY, USA; Abingon, UK, 2015.
20. Mendoza, N.B.; Yan, Z.; King, R.B. Computers \& Education Supporting students' intrinsic motivation for online learning tasks: The effect of need-supportive task instructions on motivation, self-assessment, and task performance. Comput. Educ. 2023, 193, 104663.
21. Zimmerman, B.J. Chapter 2—Attaining Self-Regulation: A Social Cognitive Perspective. In Handbook of Self-Regulation; Boekaerts, M., Pintrich, P.R., Zeidner, M., Eds.; Academic Press: San Diego, CA, USA, 2000; pp. 13-39; ISBN 978-0-12-109890-2.
22. Sharma, R.; Jain, A.; Gupta, N.; Garg, S.; Batta, M.; Dhir, S. Impact of self-assessment by students on their learning. Int. J. Appl. Basic Med. Res. 2016, 6, 226. [CrossRef]
23. Andrade, H.L.; Andrade, H.; Du, Y. Student Perspectives on Rubric-Referenced Assessment. Educ. Couns. Psychol. Fac. Scholarsh. 2005, 10, 3.
24. Black, P. and Wiliam, D. Assessment and Classroom Learning. Assess. Educ. Principles. Policy Pract. 1998, 5, 7-74. [CrossRef]
25. Larsari, V.N. A comparative study of the effect of using self-assessment and traditional method on improving students' academic mo-tivation in reading competency: The case of primary school. In Proceedings of the 1st International Congress (ICESSER), Ankara, Turkey, 29-30 May 2021.
26. Margevica-grinberga, I.; Vitola, K. Teacher Education Students' Self -assessment in COVID-19 Crisis. Psychol. Educ. 2021, 58, 2874-2883.
27. Leenknecht, M.; Wijnia, L.; Köhlen, M.; Fryer, L.; Rikers, R.; Loyens, S. Formative assessment as practice: The role of students' motivation. Assess. Eval. High. Educ. 2021, 46, 236-255. [CrossRef]
28. Schunk, D.H.; Pintrich, P.R.; Meece, J.L. Motivation in Education: Theory Research and Applications, 3rd ed.; Merrill Prentice Hall: Hoboken, NJ, USA, 2008.
29. Yarborough, C.B.; Fedesco, H.N. Motivating Students. Vanderbilt University Center for Teaching. 2020. Available online: https: / / cft.vanderbilt.edu/ /cft/guides-sub-pages/motivating-students/ (accessed on 12 August 2023).
30. Ryan, R.M.; Deci, E.L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am. Psychol. 2000, 55, 68-78. [CrossRef]
31. Gómez-Carrasco, C.J.; Monteagudo-Fernández, J.; Sainz-Gómez, M.; Moreno-Vera, J.R. Effects of a gamification and flippedclassroom program for teachers in training on motivation and learning perception. Educ. Sci. 2019, 9, 299. [CrossRef]
32. Vansteenkiste, M.; Sierens, E.; Goossens, L.; Soenens, B.; Dochy, F.; Mouratidis, A.; Aelterman, N.; Haerens, L.; Beyers, W. Identifying configurations of perceived teacher autonomy support and structure: Associations with self-regulated learning, motivation and problem behavior. Learn. Instr. 2012, 22, 431-439. [CrossRef]
33. Deci, E.L.; Koestner, R.; Ryan, R.M. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychol. Bull. 1999, 125, 627-700. [CrossRef]
34. Hattie, J.; Timperley, H. The power of feedback. Rev. Educ. Res. 2007, 77, 81-112. [CrossRef]
35. Gan, Z.; He, J.; Zhang, L.J.; Schumacker, R. Examining the Relationships between Feedback Practices and Learning Motivation. Measurement 2023, 21,38-50. [CrossRef]
36. Russo, P.; Papa, I.; Lopresto, V. Impact damage behavior of basalt fibers composite laminates: Comparison between vinyl ester and nylon 6 based systems. In Proceedings of the 33rd Technical Conference of the American Society for Composites 2018, Seattle, WA, USA, 24-27 September 2018; DEStech Publications Inc.: Lancaster, PA, USA, 2018; Volume 5, pp. 3347-3360.
37. Muenks, K.; Yang, J.S.; Wigfield, A. Associations between grit, motivation, and achievement in high school students. Motiv. Sci. 2018, 4, 158-176. [CrossRef]
38. Habók, A.; Magyar, A.; Németh, M.B.; Csapó, B. Motivation and self-related beliefs as predictors of academic achievement in reading and mathematics: Structural equation models of longitudinal data. Int. J. Educ. Res. 2020, 103, 101634.
39. Möller, J.; Pohlmann, B.; Köller, O.; Marsh, H.W. A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. Rev. Educ. Res. 2009, 79, 1129-1167. [CrossRef]
40. Huang, C. Self-concept and academic achievement: A meta-analysis of longitudinal relations. J. Sch. Psychol. 2011, 49, 505-528. [PubMed]
41. Lazowski, R.A.; Hulleman, C.S. Motivation Interventions in Education: A Meta-Analytic Review. Rev. Educ. Res. 2016, 86, 602-640. [CrossRef]
42. Lee, Y.; Rofe, J.S. Paragogy and flipped assessment: Experience of designing and running a MOOC on research methods. Open Learn. 2016, 31, 116-129. [CrossRef]
43. Toivola, M. Flipped Assessment—A Leap towards Assessment for Learning; Edita: Helsinki, Finland, 2020; ISBN 9513775151.
44. Tejeiro, R.A.; Gómez-Vallecillo, J.L.; Romero, A.F.; Pelegrina, M.; Wallace, A.; Emberley, E. La autoevaluación sumativa en la enseñanza superior: Implicaciones de su inclusión en la nota final. Electron. J. Res. Educ. Psychol. 2017, 10, 789-812. [CrossRef]
45. Sloan, J.A.; Scharff, L.F. Student Self-Assessment: Relationships between Accuracy, Engagement, Perceived Value, and Performance. J. Civ. Eng. Educ. 2022, 148, 04022004. [CrossRef]
46. Liu, J. Correlating self-efficacy with self-assessment in an undergraduate interpreting classroom: How accurate can students be? Porta Linguarum 2021, 2021, 9-25.
47. Bourke, R. Self-assessment in professional programmes within tertiary institutions. Teach. High. Educ. 2014, 19, 908-918. [CrossRef]
48. Ndoye, A. Peer/Self Assessment and Student Learning. Int. J. Teach. 2017, 29, 255-269.
49. Van Helvoort, A.A.J. How Adult Students in Information Studies Use a Scoring Rubric for the Development of Their Information Literacy Skills. J. Acad. Librariansh. 2012, 38, 165-171. [CrossRef]
50. Siow, L.-F. Students' perceptions on self- and peer-assessment in enhancing learning experience. Malays. Online J. Educ. Sci. 2015, 3, 21-35.
51. Barana, A.; Boetti, G.; Marchisio, M. Self-Assessment in the Development of Mathematical Problem-Solving Skills. Educ. Sci. 2022, 12, 81. [CrossRef]
52. Panadero, E.; Romero, M. To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. Assess. Educ. Princ. Policy Pract. 2014, 21, 133-148. [CrossRef]
53. Barney, S.; Khurum, M.; Petersen, K.; Unterkalmsteiner, M.; Jabangwe, R. Improving students with rubric-based self-assessment and oral feedback. IEEE Trans. Educ. 2012, 55, 319-325. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

