

Developing Students' "Soft Skills" through the Flipped Classroom: Evidence from an International Studies Class

ANDREA BETTI, PABLO BIDERBOST, AND AURORA GARCÍA DOMONTE
Pontifical University of Comillas, ICAI-ICADE, Madrid

Abstract: The goal of this study is to evaluate the application of the flipped classroom (FC) format in the international studies curriculum. Previous research has examined the impact of the FC on students' performance, operationalized by test scores, and demonstrated the utility of this technique in the learning process. Our research goes beyond student performance to evaluate the impact of the FC on student class attendance and the development of a set of soft skills, such as teamwork, critical thinking, self-efficacy, academic self-concept, and perception of learning. In our experiment, we compare an FC with a non-flipped class that combines traditional lecturing with other active learning techniques, such as presentations, teamwork activities, and problem-based debates. The study finds that the FC did not have a substantial impact on either students' performance, attendance, or soft skills. The differences between the two groups were too small to corroborate any tendency in favor of one or the other format. For the most part, students performed in a similar way. This can suggest that the benefits of an FC format might be less when compared with other active learning techniques.

Resumen: El objetivo de este estudio es evaluar la aplicación del formato de salón de clases invertido (Flipped Classroom, FC) en el currículo de los estudios internacionales. Las investigaciones anteriores han analizado el impacto del FC en el desempeño de los estudiantes, aplicadas a través de los resultados de exámenes, y demostraron la utilidad de esta técnica en el proceso de aprendizaje. Nuestra investigación va más allá del rendimiento académico para evaluar el impacto del FC en la asistencia a clases de los estudiantes y el desarrollo de un conjunto de habilidades sociales, como el trabajo en equipo, el pensamiento crítico, la autoeficacia, el autoconcepto académico y la percepción del aprendizaje. En nuestro experimento, comparamos un FC con una clase no invertida que combina las lecciones tradicionales con otras técnicas de aprendizaje activo, como las presentaciones, las actividades de trabajo en equipo y los debates basados en problemas. Tras completar el estudio, se descubre que el FC no tuvo un impacto significativo en el desempeño, en la asistencia ni en las habilidades sociales de los estudiantes. Las diferencias entre los dos grupos fueron demasiado pequeñas para corroborar alguna tendencia a favor de un formato u otro. En la mayoría, el desempeño de los estudiantes fue similar. Esto puede sugerir que es posible que los beneficios del formato de FC sean menores en comparación con otras técnicas de aprendizaje activo.

Résumé: L'objectif de cette étude est d'évaluer l'application du format Classe inversée dans le cursus d'études internationales. Des recherches précédentes ont examiné l'impact de la classe inversée sur les performances des étudiants, exprimées par des résultats de tests, et ont démontré l'utilité de cette technique dans le processus d'apprentissage.

Notre recherche va au-delà des performances des étudiants pour évaluer l'impact de la classe inversée sur l'assiduité des étudiants et le développement d'un ensemble de compétences comportementales, telles que le travail d'équipe, la pensée critique, l'auto-efficacité, le concept de soi académique et la perception de l'apprentissage. Dans notre expérience, nous avons comparé une classe inversée à une classe non inversée alliant des cours magistraux traditionnels à d'autres techniques d'apprentissage actif, telles que des présentations, des activités de travail d'équipe et des débats basés sur des problèmes. L'étude a permis de constater que la classe inversée n'avait pas d'impact substantiel, que ce soit sur les performances, l'assiduité ou les compétences comportementales des étudiants. Les différences entre les deux groupes ont été trop minces pour corroborer toute tendance en faveur d'un format ou de l'autre. Les performances des étudiants ont en majeure partie été similaires. Cela peut suggérer que les avantages d'un format Classe inversée sont peut-être moins importants si nous les comparons à ceux d'autres techniques d'apprentissage actif.

Keywords: international relations, teaching innovation, flipped classroom, active learning, soft skills

Palabras clave: relaciones internacionales, innovación en la enseñanza, salón de clases invertido, aprendizaje activo, habilidades sociales

Mots clés: relations internationales, innovation d'enseignement, classe inversée, apprentissage actif, compétences comportementales

Introduction

Universities have often regarded a student's brain as an "empty container into which the professor pours knowledge" (King 1993, 30). However, in the last twenty years, scholars have begun debating the best ways to teach and the best ways for students to learn. Many scholars criticize the traditional teaching format, based on frontal lectures, in which the teacher explains the contents, with students listening, often passively and taking notes. Now, thanks to the widespread communication technology, particularly during the COVID-19 crisis, scholars and teachers have an opportunity to research the effectiveness of different learning styles.

The term "active learning" is increasingly common in the education field. It is usually associated with "student activity and engagement in the learning process" (Prince 2004, 223). The flipped classroom (FC) is one of the most recent examples of active learning techniques. The basic idea is "that which is traditionally done in class is now done at home, and that which is traditionally done as homework is now completed in class" (Bergmann and Sams 2012, 13). The learning process is inverted, with students watching video lectures outside class and later doing homework, exercises, and other activities in class. With this format "students do the lower levels of cognitive work (remember, understand) outside of class and, thus, can focus on the higher levels (apply, analyse, evaluate) inside the class under the direction of the instructor and the support of fellow students" (Yamarick 2019, 1). While the FC has been applied relatively extensively in education, engineering, and languages, a limited number of studies has explored its effects in the field of Politics and International Studies (IS). The goal of this article is to examine the application of the FC in an IS class on "Regional Studies: Latin-America."

Most investigations have explored the impact of the FC on students' performance in terms of exam results. This has offered valuable insights about the usefulness of the technique. However, to better involve students in the learning process, we must look beyond their performance. In addition to exploring the FC's impact on students' academic performance, our study analyzes whether the FC can be a useful resource to stimulate their attendance and the development of a set of soft skills, like teamwork, critical thinking, self-efficacy, academic self-concept, and perception of learning. After comparing an FC with a non-flipped class that combines traditional lecturing with other active learning techniques, the study finds that the FC did not have a substantial impact on either students' performance, attendance, or soft skills. The differences between the two groups were too small to corroborate any tendency in favor of one or the other format. For the most part, students performed in a similar way.

In the first section, we review the literature on the use of the FC. In the second section, we explain our research design and the type of statistical analysis. In the third section, we discuss the results of the analysis, and in the final section, we identify the limitations of the study together with some venues for future research.

Flipped Classroom: The Literature

The FC originated in the context of "student-centred learning," which is usually defined "in opposition to teacher-centred models." In the latter, "the acquisition of knowledge" takes place "outside of the context in which it will be used" and tends to be based on the idea of "instructional delivery," which includes "lecture, homework, and exams" (Hamdan et al. 2013, 7). In the former, students can "actively engage in activities such as . . . collaborative discussion [and] problem-solving" (Mazur, Brown, and Jacobsen 2015, 5). Using video lectures and other contents that are uploaded to the web, teachers can "reduce the amount of time spent in class on lecturing, opening up class time for the use of active learning strategies," more focused on "understanding and application (critical and creative thinking) than on recall of facts (content/basic thinking)" (Baker 2000, 11).

Several scholars observed a positive effect on performance, in the sense that students exposed to the FC would obtain better academic results (Flores, Del-Arco, and Silva 2016; Hsieh, Vivian Wu, and Marek 2017). Other advantages would be an improvement in their engagement (Elmaadaway 2017; Karabulut-Ilgü et al. 2018) and motivation (Hsieh, Vivian Wu, and Marek 2017). Nevertheless, there are studies that did not find any significant difference between a flipped and non-flipped format in terms of either students' grades or their course evaluation responses (McLaughlin et al. 2013). Some even found partially negative effects of the FC on perceptions, with students preferring traditional teaching (Roehling et al. 2017).

The application of the FC in the field of Politics and IS has been more limited. Some scholars observed that this technique would improve performance, for example, by increasing the possibility that students "actively engage the material during class and also remain for question and answer sessions" (Touhchon 2015, 38). Others did not find clear effects. In terms of grades, for Lambach, Karger, and Goerres (2017, 561), there would be "no evidence of an overall association between exam results and the course format." Moreover, students would split in terms of perceptions, with half of them preferring conventional teaching and another half preferring the FC (Lambach, Karger, and Goerres 2017, 562). In a similar way, Jenkins (2015, 610) reported that students are quite prudent about the implementation of the FC, meaning that they tend to prefer "mixed class sessions" as compared to "having all flipped-class sessions."

Each of these studies compared active learning formats with fully traditional teaching formats. This can lead to partially biased results, especially due to the increasing criticism against traditional teaching. For this reason, Jensen et al.

proposed a comparison of two different teaching formats that are both structured around the same active learning technique, such as the 5-E learning cycle.¹ One teaching format would be inverted, the other one would not. This way, the authors tried to limit the risk of comparing two different formats, making it difficult to identify recognizable effects (Jensen, Kummer, and Godoy 2015).

A few studies proposed comparisons either between different types of FCs or between the FC and other active learning techniques. In the former group, researchers found that the FC can improve grades when integrated with “cooperative learning activities,” such as students working in groups (Foldnes 2016) or when administered in a self-regulated way, in which students monitor their own learning performance (Lai and Hwang 2016). In the latter group, researchers observed a positive effect of the FC on students’ performance and satisfaction when compared to other active learning techniques, such as simulations (Davies, Dean, and Ball 2013). For others, the FC would enhance students’ satisfaction, especially when they perceived it as an instrument to improve online teaching (Tang et al. 2020), or when the FC is integrated with other learning management systems (LMSs), such as Moodle (Nouri 2016). Nevertheless, other scholars noticed that, even when it improves grades, the FC does not necessarily improve perceptions. Students would express more satisfaction with other types of video classes, such as lecture capture backup, that is videos that are not intended to replace the frontal lecture but to review the contents after each class (Missildine et al. 2013). In one of the few studies in the field of Politics that compared the FC with other non-traditional formats, such as online teaching, Whitman Cobb (2016) reported an improvement of both students’ performance and satisfaction.

In our study, we made a similar methodological decision. Instead of studying the differences between a fully traditional and an inverted teaching format, we compare an FC with a non-flipped class that combines traditional lecturing with other active learning techniques, such as presentations, teamwork activities, and problem-based debates.

Flipped Classroom, Attendance, and “Soft Skills”

On the one hand, hard skills are usually understood as the “technical expertise and knowledge needed for a job.” They are “achievements that are included on a resumé, such as education, work experience, knowledge, and level of expertise.” Examples of hard skills are “typing, writing, math, reading, and the ability to use software programs” (Robles 2012, 457). On the other hand, soft skills are “interpersonal qualities, also known as people skills, and personal attributes that one possesses,” such as responsibility, communication, and teamwork (Robles 2012, 453). There has been a tendency in education to underestimate that “success in life depends on personality traits that are not well captured by measures of cognition” (Heckman and Krautz 2012, 3). Lacking these soft skills can “sink the promising career of someone who has technical ability and professional expertise but no interpersonal quality” (Robles 2012, 459). Few authors explored the impact of active learning on the development of students’ soft skills. This makes it important to generate new data about the effects of the FC in terms of improvement (or non-improvement) of such skills.

First, many scholars recognize the importance of managing teamwork skills. Salas, Sims, and Burke (2005, 558–59) identified five main components of teamwork: “team leadership, mutual performance monitoring, backup behaviour, adaptability,

¹Technique that consists of “five instructional phases.” The first three phases, “engage,” “explore,” and “explain,” “are used to facilitate the content attainment.” The “elaborate” phase “is the one in which students apply the concepts they have constructed through the content attainment stage, that is, the concept application phase.” Finally, “evaluate” is the phase that tests “students’ understanding of the concepts they have just learned” (Jensen, Kummer, and Godoy 2015, 3).

and team orientation.” Several studies revealed that the FC can improve students’ teamwork (Foldnes 2016; McNally et al. 2017; Karabulut-Ilgü et al. 2018). Nevertheless, in Politics and IS, there is more debate. Some political scientists generally argued that the FC “creates a learning space that is more favourably received by students who prefer collaborative learning” (Lambach, Karger, and Goerres 2017, 563), while others observed that students are “less positively inclined toward the group work that came in the flipped-class sessions” (Jenkins 2015, 610). Except for Foldnes (2016), who studied the FC in relation with cooperative learning techniques, these studies are all based on comparisons between the FC and fully traditional teaching formats.

Second, one of the most oft-mentioned soft skills is critical thinking, defined as the capacity “to think reflectively and judge skilfully, so as to decide what information is reliable and what actions should be taken” (Kong 2014, 3). It involves learning to “evaluate a proposed solution” or “the reasonableness of a hypothesis,” or “which of two alternative methods is a more effective and efficient way of solving given problems” (Anderson et al. 2001, 84). As such, it is not a hard skill, such as a quantifiable competence or a specific level of expertise. Rather, it is an interpersonal quality developed through experience. In Politics and IS, scholars found a positive relation between critical thinking and active learning techniques, such as team-based learning, debates, simulations (Truby, Weiss, and Rousseau 2014) or problem-based learning (Burch 2000). The FC too was associated with the development of critical thinking. In Politics, Lambach, Karger, and Goerres (2017, 556) observed that through the FC, students can acquire competences in which they “synthesise, develop an independent opinion and formulate critical statements.” However, this study did not empirically test the effect of the FC on this skill. In other disciplines, more systematic studies detected a positive impact of the FC on critical thinking (McLaughlin et al. 2013; Kong 2014; Hsieh, Vivian Wu, and Marek 2017). Nevertheless, none of these studies are based on a comparison between different types of active learning techniques.

Another skill we focused on was self-efficacy, which psychologist Albert Bandura (1994, 72) defined as “people’s beliefs about their capabilities.” This can translate into approaching “difficult tasks as challenges to be mastered rather than as threats to be avoided.” Self-efficacy is an important indicator of the students’ capacity to pursue “challenging goals” and “recover ... after failures or setbacks.” No political scientist that we know of has explored the relation between active learning and self-efficacy. Scholars from other disciplines, such as McNally et al. (2017), compared an FC with a traditional learning environment and did not observe any significant impact of the FC on this skill. Lai and Hwang (2016) were among the few to compare different types of FCs and to find that its self-regulated version can improve self-efficacy.

We then looked at academic self-concept. This includes “self-confidence,” which is the “level of confidence a person has in his or her ability to perform certain behaviours,” “importance,” which is “a person’s perceptions about the importance of performing certain behaviours,” and “outcome confidence,” which is “the degree of confidence . . . that positive outcomes will result from performing certain behaviours” (Gresham 1995, 19–20). Self-concept allows students to “rate how confident they feel in performing tasks . . . their probable outcomes . . . and how important these tasks are for them” (Gresham 1995, 19). An improvement in academic self-concept can positively influence achievement and emotions, such as happiness and self-esteem (Marsh and Martin 2011). We could not find any study that tested the impact of active learning on this soft skill.

The last soft skill that we analyzed is the perception of students’ learning. The debate on this issue has generated contrasting results. In studies comparing the FC with traditional teaching formats, the results showed that students exposed to the FC would achieve better grades and also express more satisfaction of their

perception of increased learning. This has been the case both in the field of Politics (Touchton 2015) and in other disciplines (Chen et al. 2014; Hsieh, Vivian Wu, and Marek 2017; Awidi and Paynter 2019). However, others obtained less coherent results. In some cases, although it increased students' engagement, the FC did not significantly improve their satisfaction (McLaughlin et al. 2013). In other cases, students who were exposed to active learning had a tendency to learn more than students who were not, as indicated by grades. However, the former perceived that they learned less, because of the "increased cognitive effort required during active learning" (Deslauriers et al. 2019, 19 251). In the studies that compared the FC with other active learning formats, one can find a similar variety of results. Some found that the FC would improve both grades and the perception of learning (Nouri 2016; Tang et al. 2020), while others discovered that, despite an improvement in their grades, students were less satisfied with the FC, as it "required more work" (Missildine et al. 2013, 599). These contradictory results remind us of the difficulty to measure the impact of a teaching technique by relying on students' perceptions. Sometimes they can be poor judges of their own learning progress. We will return to this problem in the discussion.

Finally, our study explored the impact of the FC on attendance, which is a good indicator of the students' motivation and engagement with their learning process. Most studies have so far explored whether recorded video and audio lectures discourage students' attendance. Some did not detect any specific relation between these types of teaching formats and attendance (Copley 2007; Von Konski, Ivins, and Gribble 2009; Larkin 2010). On the contrary, others observed that recorded lectures would make students perceive attendance as less indispensable (Traphagan, Kucsera, and Kishi 2010; Jensen 2011). Among the scholars that explored the flipped format, some detected an improvement in attendance (Chen et al. 2014; Elmaadaway 2017), while others observed a negative effect (Yamarick 2019). All these studies are based on comparisons between active learning techniques, including the FC, and fully traditional teaching formats.

Data and Methods

By relying on a quantitative approach, this exploratory study aims to understand whether the implementation of an FC can improve students' performance in terms of academic achievement, soft skills, and class attendance.

Our sample consisted of sixty-three students enrolled in a dual degree in international relations and business administration. In addition, they were enrolled in a diploma in "Development of Personal, Communication, and Professional Abilities." At the university where we collected this sample, this diploma is mandatory for all the degrees offered. The purpose of this diploma is to develop students' social, interpersonal, and professional skills.

The quasi experiment took place in one section of the IS class "Regional Studies: Latin America." This section—"Politics in Latin America"—dealt with the national and international dimensions of Latin American political systems and represented around 33 percent of the class. This meant the first five weeks of the semester. We administered the experiment only in this section to avoid too disruptive and overwhelming changes that could make the students' perceptions biased.

In this section, the class was divided into two groups, taught by the same professor. In the first group—the control group composed of thirty-two students—the class was organized in a semi-traditional way, combining traditional frontal lectures with a variety of active learning activities, such as problem-based debates, teamwork activities, and presentations. The professor taught the control group four hours per week, with two hours dedicated to frontal lecturing and the other two dedicated to active learning activities. The traditional lecture was combined with videos, such as news or short documentaries, taken from YouTube and not produced by the

professor. The professor also gave several individual and group exercises, focusing on questions and debates relating to the lecture's contents.

The second group—the experimental group composed of thirty-one students—was entirely taught through the FC format.² The professor removed frontal lecturing from the classroom and substituted it with some video lectures, self-produced through the application *Kaltura*, and assigned the videos in advance. All the four hours per week were dedicated to active learning activities in class, such as presentations, debates, and exercises, guided by the professor and related to the content explained in the videos. The application *Poll Everywhere* conducted some of the activities and helped the professor to give different surveys. Moreover, the professor used other techniques, like logical problems, case studies, trivia quizzes, word search puzzles, photo find puzzles, crosswords, alphabet soups, and specific research tasks on the web for the key concepts comprehension. Students had to perform these exercises in the classroom, either individually or in groups. The professor used similar active learning techniques and exercises for both groups.³ The main difference between the two groups was that in the FC all the classroom time was spent on active learning.

We divided students into two groups based on the alphabetical order of their last name. This is the criterion that this specific university usually adopts to divide the largest lecture classes into subgroups.

Data: Student Characteristics, Academic Performance, Attendance, and Soft Skills

To collect students' data, we selected these variables: gender, membership of the control or experimental group, and the academic performance in terms of grades in the degree in international relations, in the degree in business administration, and in the diploma. These are the five independent variables of our study.⁴

Table 1 presents the descriptive statistics for these variables, divided in terms of control and experimental group. There are no significant differences between these factors in the two samples, except for gender. The control group is composed mostly of women, while in the experimental group women are a minority. Moreover, another difference between the two samples is the type of teaching methodology that students were exposed to: FC for the experimental group and traditional class integrated with other active learning techniques for the control group. The other factors, more clearly academic, allow us to assume homogeneity between both groups.

To measure the academic performance of students, we compared the students' results in three different exams. First, students filled out a survey at the beginning of the section "Politics in Latin America," before studying the contents. This measured their prior knowledge of the subject. Second, students filled out a survey at the end of the section, after the professor taught them the contents. And third, students took the exam at the end of the section after studying the contents. Attendance data were collected through a sheet that students signed before each class. Since we wanted to test the effects of the FC on attendance, in our experiment, attendance did not have any impact on the final grade. We took attendance only as a control for data collection. This guaranteed that attendance was not influenced by other factors, like the requirement to attend a certain number of lessons to pass the class.

² Students of the experimental group did not have any previous experience with the FC format.

³ After attending several FC courses at the university which held the experiment, the professor implemented the FC for the first time. However, it is important to also note that the professor had extensive experience using the other active learning techniques mentioned, such as problem-based debates, teamwork activities, and presentations.

⁴ We could consider other variables, such as age, number of years spent at university, or whether they are part-time or full-time students. Nevertheless, these variables would not have been helpful to discriminate our sample. The students of our sample are all almost the same age, with small differences of maximum one year younger or older, they are all enrolled in the third year of their dual degree, and they are only allowed to be full-time students.

Table 1. Descriptive statistics of student characteristics

	Control group		Experimental group		Difference (Experimental – Control)	
	Mean	Standard deviation	Mean	Standard deviation	Mean	p-value
Gender						
Male = 0; Female = 1	0.780	0.420	0.450	0.506	-0.330	(0.007)***
Grade average international relations studies						
0–10	7.736	0.639	7.707	0.522	-0.029	(0.847)
Grade average business administration studies						
0–10	7.296	0.811	7.363	0.636	0.067	(0.716)
Grade average diploma						
0–10	8.491	0.384	8.521	0.380	0.029	(0.761)
Observations	32		31		63	

Notes: Means, standard deviation, or p-value of control group, experimental group, and difference between them.

***Significance difference in means across the two groups at 1 percent level of confidence.

We performed a description of the data by considering the average results, the typical deviation, and a *dif in dif* (DiD) analysis. We aimed to measure students' evolution throughout the section, relying on the use or nonuse of the FC as the discriminant variable. After that, we applied a linear regression to our five independent variables.

Finally, we collected data on the soft skills, by giving a twenty-eight-question survey to both groups.⁵ We applied a linear regression to each group of variables. We collected data on the soft skills twice, at the beginning and at the end of the section.

Table 2 presents the data obtained through the twenty-eight-question survey that we gave to both groups during the first and the last day of class of the section.

The twenty-eight questions were about the soft skills. For the questions about self-efficacy (1–10), academic self-concept (17–19), and learning perception (20), we used a Likert scale from 0 to 5, whereas 0 meant “completely disagree” and 5 meant “completely agree.” For the questions about teamwork (11–16), we used a scale to measure the negative or positive perceptions of students, whereas 0 meant “very negative perception” and 5 meant “very positive perception.” Finally, to measure critical thinking (21–28), we used a set of questions whereas only one was correct, so that the set of answers was a dichotomous variable in which 0 meant “wrong” and 1 meant “right.”⁶

Model

We used mean comparisons and a multivariate linear regression, as the following:

$$X_{i,m} = A_1 + A_2 \text{Group}_m + A_3 \text{Gender}_m + A_4 \text{IRGA}_m + A_5 \text{BAGA}_m + A_6 \text{DGA}_m + \varepsilon_i \quad (1)$$

⁵The Institute of Science Education at the University where the experiment took place provided the questionnaire. It was based on established questions previously used by this Institute. A pretest was done with five students to validate its wording and comprehension.

⁶To measure critical thinking (21–28), we calculated the average data of the eight questions (21–28). We assigned 0 to the wrong answers and 1 to the correct answers. A result below 0.5 meant that students mostly answered in a wrong way, above 0.5 meant that students mostly answered in a correct way, and equal to 0.5 meant that students answered the same number of correct and wrong answers.

Table 2. Twenty-eight-question survey items

Self-efficacy		Completely disagree	Completely agree
1	I think I am going to get some excellent grades this year.	0	5
2	If I make an effort, I think I have enough capacity to achieve a good academic record.	0	5
3	I believe that I am able to understand even the most difficult topics in this course.	0	5
4	I think I have enough capacity to understand a subject, quickly and well.	0	5
5	I think I can pass the courses quite easily and even get good grades.	0	5
6	Although teachers are demanding and strict, I have great confidence in my own academic ability.	0	5
7	I think that I am prepared and well qualified to achieve academic success.	0	5
8	When they ask me to do projects or homework, I am sure that I will do them well.	0	5
9	I work effectively in any team, no matter who the teammates are.	0	5
10	Considering the difficulty of the degree, what I am learning, and my own abilities, I think I will be fine when I finish (the degree).	0	5
Teamwork ^a		Very negative perception	Very positive perception
11	Participation in teamwork sharing information, knowledge, and experiences.	0	5
12	Acceptance and compliance with the rules agreed upon in the group (deadlines, parts of the work, format, etc.).	0	5
13	Action to face the conflicts of the team in this subject.	0	5
14	Commitment to the management and operation of the team	0	5
15	Management of meetings effectively.	0	5
16	Communication and cohesion within the group.	0	5
Academic self-concept		Completely disagree	Completely agree
17	Regarding my grades, I think I am among the best ten students in class	0	5
18	My grades so far have been mostly as	0	5
19	In general, I consider myself a very good student	0	5
Learning perception		Few	A lot
20	Regardless of your results in the exams, think how much you will learn in this section about Latin American politics	0	5
Critical thinking		Wrong	Right
21–28	Application of a set of questions about a fantasy story read by students in advance in order to test their capacities to use logical and critical thinking.	0	1

^aBefore our experiment, students had already attended a class of the diploma, called “Teamwork,” which means that they already had a perception of their teamwork skills. At the beginning of our experiment, we measured this initial perception. We assumed that this was greater than zero since students had already attended the “Teamwork” class of the diploma. At the end of the experiment, we measured the students’ perception of this skill, by considering whether there were any negative or positive changes in their initial perception.

In which:

- X_i is each one of the dependent variables collected to measure academic performance, class attendance, and soft skills.
- m is the observation of each student, from 0 to 63.
- A is the estimated coefficient for each one of the independent variables, which are:
 - *Group*: dichotomous variable, which indicates whether the student has received or not received the FC, that is to say, whether he/she belongs to the control (0) or the experimental (1) group.
 - *Gender*, dichotomous variable that reflects students' gender, in which *male* equals to 0 and *female* equals to 1.
 - *IRGA*, continuous variable that reflects the average grade in the degree in international relations achieved at the time of the experiment. It can have a value between 0 and 10.
 - *BAGA*, continuous variable that reflects the average grade in the degree in business administration achieved at the time of the experiment. It can have a value between 0 and 10.
 - *DGA*, continuous variable that reflects the average grade in the diploma. It can have a value between 0 and 10.

Based on the types of data, we made ordinary least squared (OLS) estimations of the equation (1) for the variables of academic achievement and attendance. For the soft skills, we made ordinal logistic regressions (OLR) for the first twenty questions of the survey and binary logistic regressions (BLR) for the questions 21 to 28, in both cases of the equation (1).⁷

Analysis

Academic Performance and Attendance

We made a description of the data by considering the average results, the typical deviation, and the difference analysis to measure students' evolution throughout the class. We relied on the use or nonuse of the FC as the discriminant variable. Table 3 presents the results. There are no substantial differences between the two groups, experimental and control, in terms of the summary measures of performance and attendance.

Table 4 presents the results of the multivariate regression for achievement and attendance. In terms of achievement, the average grade of the diploma influences only the final exam grade significantly. In terms of attendance, the average grade of the diploma and gender positively influence the results; women and the best average grade of the diploma tend to have a higher index of attendance to class. This means that the teaching methodology used in class—FC or traditional class integrated with other active learning techniques—does not significantly affect any of the three measurements of academic achievement. Neither does it affect the attendance to class.

Soft Skills

Table 5 presents the sample means of the soft skills perceived by students across ten items of self-efficacy, six items of teamwork, three items of academic self-concept, one item of learning perception, and eight items of critical thinking. The left panel shows the values of the first-day survey, the middle panel the values of the last-day

⁷ OLS was applied for the summary measure of critical thinking.

Table 3. Academic performance and attendance

	Control		Experimental		Difference (Experimental – Control)	
	Mean	Standard deviation	Mean	Standard deviation	Mean	<i>p</i> -value
Initial test						
0–10	4.391	1.293	4.306	0.928	–0.084	(0.768)
Second test						
0–10	4.969	1.350	5.113	1.606	0.144	(0.701)
Difference						
Second – Initial test						
0–10	0.578	1.454	0.806	1.667	0.228	(0.564)
Final exam						
0–10	8.204	1.955	8.007	2.036	0.198	(0.695)
Attendance						
0–5	3.840	0.954	3.90	1.106	0.060	(0.820)
Observations	32		31		63	

Notes: Means, standard deviation or *p*-value of control group, experimental group, and difference between the two groups.

Table 4. Impact of the FC, gender, and average grades on the academic performance and attendance

Independent variables	Dependent variables X _i : academic performance and attendance			
	Initial test	Second test	Final exam	Attendance
Flipped classroom	–0.242	0.144	–0.210	0.253
No = 0; Yes = 1	(0.428)	(0.709)	(0.698)	(0.300)
Gender	–0.504	–0.076	–0.336	0.683
Male = 0; Female = 1	(0.108)	(0.848)	(0.541)	(0.007)***
IRGA	0.399	0.802	1.147	0.211
0–10	(0.533)	(0.325)	(0.314)	(0.679)
BAGA	0.060	0.242	–0.305	0.123
0–10	(0.905)	(0.704)	(0.732)	(0.759)
DGA	–0.028	–0.623	–1.557	0.994
0–10	(0.946)	(0.230)	(0.034)**	(0.003)***
Observations	63	63	63	63
R ²	10.10 percent	15.90 percent	9.60 percent	31.30 percent
Sig.	(0.287)	(0.071)*	(0.316)	(0.001)***

Notes: Each entry for the independent variables is the estimated coefficient and its *p*-value.

*, **, ***: Significance difference at 10 percent, 5 percent, and 1 percent level of confidence.

survey, and the right panel the difference (last minus first). In each panel, the data are collected from the control group (first column), the experimental group (second column) and for the difference (experimental minus control, third column). For the study of differences, we used ANOVAs, with the *p*-values appearing in parenthesis.

As table 5 indicates, at the beginning of the experiment, students' perceptions of the soft skills were similar in both groups, except for one item of self-efficacy (1), one of teamwork (12), and one of academic self-concept (18), which, on average, show significantly higher results for the experimental group. Moreover, in terms of critical thinking, three items (26, 27, 28), and the aggregate of the soft skill, that is students' development, were significantly lower in the experimental group than in the control group.

Table 5. Statistics of soft skills

	First day				Last day				Difference (last – first)			
	Control	Experimental	p-value	(experimental-control)	Control	Experimental	p-value	(experimental-control)	Control	Experimental	p-value	Difference (experimental – control)
Self-efficacy												
1.	3.72	3.84	0.12 (0.544)	0.52 (0.046)**	3.81	4.02	0.21 (0.303)	0.09 (0.644)	0.09	0.17	0.09 (0.644)	0.09 (0.644)
2.	3.25	3.77			3.47	3.68	0.21 (0.358)	–0.10 (0.241)	0.22	–0.10	–0.32 (0.241)	–0.32 (0.241)
3.	4.22	4.16	–0.06 (0.834)		4.28	4.13	–0.15 (0.538)	–0.03 (0.721)	0.06	–0.03	–0.10 (0.721)	–0.10 (0.721)
4.	3.72	4.06	0.35 (0.193)		3.94	4.06	0.13 (0.639)	–0.22 (0.405)	0.22	0.00	–0.22 (0.405)	–0.22 (0.405)
5.	3.78	3.90	0.12 (0.682)		3.88	4.13	0.25 (0.333)	0.13 (0.627)	0.09	0.23	0.13 (0.627)	0.13 (0.627)
6.	3.44	3.55	0.11 (0.634)		3.63	3.68	0.05 (0.847)	–0.06 (0.832)	0.19	0.13	–0.06 (0.832)	–0.06 (0.832)
7.	3.56	3.81	0.24 (0.283)		3.69	4.00	0.31 (0.234)	0.07 (0.789)	0.13	0.19	0.07 (0.789)	0.07 (0.789)
8.	4.00	3.97	–0.03 (0.889)		3.97	4.19	0.23 (0.387)	0.26 (0.359)	–0.03	0.23	0.26 (0.359)	0.26 (0.359)
9.	3.47	3.52	0.05 (0.855)		3.47	4.06	0.60 (0.016)**	0.54 (0.043)**	0.00	0.55	0.54 (0.043)**	0.54 (0.043)**
10.	3.75	3.77	0.02 (0.930)		3.75	3.97	0.22 (0.383)	0.19 (0.429)	0.00	0.19	0.19 (0.429)	0.19 (0.429)
Teamwork												
11.	4.03	3.90	0.13 (0.636)		4.03	4.26	0.23 (0.363)	0.35 (0.197)	0.00	0.35	0.35 (0.197)	0.35 (0.197)
12.	3.77	3.94	0.175 (0.192)		3.88	3.89	0.007 (0.958)	–0.168 (0.148)	0.11	–0.05	–0.168 (0.148)	–0.168 (0.148)
13.	3.72	3.71	–0.01 (0.961)		3.59	3.74	0.15 (0.386)	0.16 (0.468)	–0.13	0.03	0.16 (0.468)	0.16 (0.468)
14.	3.69	4.13	0.44 (0.054)*		4.09	4.00	–0.09 (0.602)	–0.54 (0.019)**	0.41	–0.13	–0.54 (0.019)**	–0.54 (0.019)**
15.	3.66	3.97	0.31 (0.203)		3.84	3.97	0.12 (0.607)	–0.19 (0.499)	0.19	0.00	–0.19 (0.499)	–0.19 (0.499)
16.	3.91	4.03	0.13 (0.516)		4.06	4.13	0.07 (0.684)	–0.06 (0.741)	0.16	0.10	–0.06 (0.741)	–0.06 (0.741)
17.	3.72	3.65	–0.07 (0.772)		3.72	3.35	–0.36 (0.169)	–0.29 (0.244)	0.00	–0.29	–0.29 (0.244)	–0.29 (0.244)
18.	3.91	4.16	0.26 (0.263)		3.97	4.13	0.16 (0.459)	–0.09 (0.712)	0.06	–0.03	–0.09 (0.712)	–0.09 (0.712)
Academic self-concept												
19.	3.28	3.52	0.23 (0.109)		3.28	3.42	0.14 (0.392)	–0.96 (0.339)	0.00	–0.10	–0.96 (0.339)	–0.96 (0.339)
20.	3.19	3.48	0.30 (0.184)		3.13	3.39	0.26 (0.345)	–0.03 (0.888)	–0.06	–0.10	–0.03 (0.888)	–0.03 (0.888)
Learning perception												
21.	2.91	3.13	0.22 (0.091)*		2.94	3.10	0.16 (0.281)	–0.06 (0.535)	0.03	–0.03	–0.06 (0.535)	–0.06 (0.535)
22.	3.75	3.94	0.19 (0.224)		3.78	3.77	–0.01 (0.968)	–0.19 (0.079)*	0.03	–0.16	–0.19 (0.079)*	–0.19 (0.079)*
23.	4.38	4.19	–0.18 (0.355)		4.41	4.23	–0.18 (0.335)	0.00 (0.995)	0.03	0.03	0.00 (0.995)	0.00 (0.995)

Table 5. Continue

	First day			Last day			Difference (last – first)		
	Control	Experimental	Difference (experimental – control) <i>p</i> -value	Control	Experimental	Difference (experimental – control) <i>p</i> -value	Control	Experimental	Difference (experimental – control) <i>p</i> -value
Critical thinking	0.59	0.46	–0.12 (0.007)***	0.59	0.55	–0.04 (0.424)	0.01	0.09	0.09 (0.115)
21.	0.31	0.23	–0.09 (0.446)	0.31	0.39	0.08 (0.542)	0.00	0.16	0.16 (0.190)
22.	0.25	0.29	0.04 (0.724)	0.34	0.16	–0.18 (0.099)*	0.09	–0.13	–0.22 (0.130)
23.	0.81	0.77	–0.04 (0.713)	0.75	0.84	0.09 (0.393)	–0.06	0.06	0.13 (0.293)
24.	0.84	0.74	–0.10 (0.326)	0.72	0.87	0.15 (0.140)	–0.13	0.13	0.25 (0.033)**
25.	0.28	0.35	0.07 (0.538)	0.41	0.42	0.01 (0.918)	0.13	0.06	–0.06 (0.672)
26.	0.66	0.39	–0.27 (0.033)**	0.72	0.42	–0.30 (0.016)***	0.06	0.03	–0.03 (0.838)
27.	0.84	0.58	–0.26 (0.021)**	0.72	0.65	–0.07 (0.538)	–0.13	0.06	0.19 (0.164)
28.	0.66	0.32	–0.33 (0.008)***	0.75	0.68	–0.07 (0.531)	0.09	0.35	0.26 (0.074)*
Observations	32	31	63	32	31	63	32	31	63

Notes: Each entry is the sample means. The values for the difference column (last – first) are the changes in values from the first to the last day. The values for the difference columns (experimental – control) are the changes in values between the two groups.

*, **, ***; Significant difference in means across the two groups at 10 percent, 5 percent, and 1 percent level of confidence.

Something similar occurred at the end of the experiment (last day), where we found few differences between the two groups in terms of soft skills development. Only three items presented a significant difference, one in terms of self-efficacy (experimental higher than control, item 8) and two in terms of critical thinking (experimental lower than control, items 22 and 26). Finally, as the data in the right column show, after comparing the results of the experimental group at the end of the experiment (last day) with the differences of its initial results (first day) and with the results of the control group, students exposed to the FC improved in one item of self-efficacy (8), and in two items of critical thinking (24 and 28) while they worsened in one of teamwork (12) and in one of academic self-concept (19).

Table 6 presents the results of the regressions used to measure whether the methodology implemented in class had any effect on the soft skills as perceived by students. We present the odd ratios on the different soft skills of the independent variable, which reflect the use or nonuse of the FC for the items 1–20, and the estimated coefficient for the items 21–28. The results of the survey from the first day of class are presented in the left panel. The middle panel presents the results of the survey administered on the last day of class, while the right panel presents the estimations of the DiD analysis.

For this DiD analysis, we used the following regression model:

$$X_i = B_1 + B_2 \text{Group} + B_3 \text{Moment} + B_4 (\text{Group} * \text{Moment}) + \varepsilon_i. \quad (2)$$

In which:

- X_i is each one of the dependent variables collected in this part of the study, in order to measure the soft skills.
- B are the estimated coefficients for each one of the independent variables, which are:
 - *Group*: dichotomous variable, which indicates whether the student has received or not received the FC, no (0) or yes (1); and
 - *Moment*, dichotomous variable that reflects the moment of the survey, first day being 0 and last day being 1.

Results in table 6 show a significant effect among the students who were exposed to the FC (last day) in terms of the development of one item of self-efficacy (item 8, where the experimental group improves) and two items of critical thinking (item 24, where the experimental improves, and item 26, where the experimental worsens). The measurement of the impact of the use of the FC on the students of the experimental group (DiD analysis) shows that these students (experimental) were less inclined toward "acceptance and compliance with the rules agreed upon in the group (deadlines, parts of the work, format, etc.)," that is item 12 of teamwork. However, they performed better in terms of one item of critical thinking (24).

Discussion

After comparing an FC with a semi-traditional class integrated with other active learning techniques, we observed no substantial difference between the control and the experimental group in terms of academic performance. On one hand, the data visualization (figure 1) shows some differences that were undetected by the chosen statistical techniques. Students of the experimental group outperformed the students of the control group in the second test and in terms of the progression (difference between the initial and second test in a before–after logic). This means that the experimental group started with lower results but finished with better

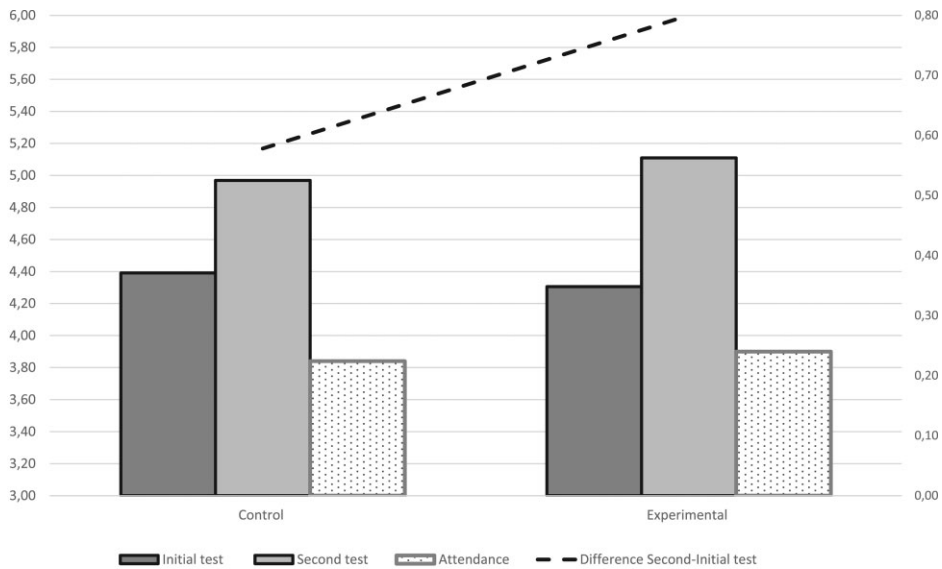
Table 6. Impact of the FC on the soft skills

Dependent Variable	First day		Last day		Difference	
	Odd ratio/coefficient	p-value	Odd ratio/coefficient	p-value	Odd ratio/coefficient	p-value
Self-efficacy						
1.	2.971	(0.048)**	1.628	(0.354)	0.530	(0.336)
2.	0.946	(0.918)	0.409	(0.107)	0.513	(0.326)
3.	1.806	(0.268)	0.699	(0.513)	0.465	(0.260)
4.	1.244	(0.668)	1.215	(0.716)	1.300	(0.693)
5.	1.011	(0.984)	0.674	(0.444)	0.913	(0.890)
6.	1.248	(0.674)	1.291	(0.635)	1.313	(0.680)
7.	0.794	(0.661)	1.155	(0.782)	1.586	(0.487)
8.	1.617	(0.343)	3.949	(0.014)**	2.749	(0.124)
9.	0.935	(0.897)	1.605	(0.352)	1.425	(0.593)
10.	0.587	(0.317)	1.479	(0.467)	2.088	(0.277)
Teamwork						
11.	1.386	(0.545)	2.055	(0.202)	1.605	(0.498)
12.	3.481	(0.019)**	1.257	(0.675)	0.289	(0.067)*
13.	1.654	(0.341)	1.547	(0.383)	0.639	(0.499)
14.	1.731	(0.349)	1.851	(0.304)	0.847	(0.824)
15.	1.047	(0.929)	0.626	(0.377)	0.623	(0.472)
16.	1.222	(0.695)	1.343	(0.571)	0.779	(0.704)
Academic self-concept						
17.	3.570	(0.047)**	1.688	(0.325)	1.026	(0.970)
18.	6.605	(0.065)*	1.912	(0.351)	0.803	(0.774)
19.	2.454	(0.184)	1.286	(0.676)	0.519	(0.358)
Learning perception						
20.	0.888	(0.818)	0.770	(0.624)	1.012	(0.986)

Table 6. Continue

Dependent Variable	First day		Last day		Difference	
	Odd ratio/coefficient	p-value	Odd ratio/coefficient	p-value	Odd ratio/coefficient	p-value
Critical thinking						
21.	0.896	(0.863)	1.732	(0.361)	2.165	(0.323)
22.	1.296	(0.690)	0.415	(0.183)	0.299	(0.149)
23.	1.060	(0.934)	1.735	(0.428)	2.191	(0.379)
24.	0.770	(0.725)	3.683	(0.090)*	4.961	(0.082)*
25.	1.726	(0.361)	0.949	(0.926)	0.751	(0.701)
26.	0.356	(0.073)*	0.318	(0.057)*	0.854	(0.833)
27.	0.325	(0.093)*	0.918	(0.886)	2.775	(0.211)
28.	0.147	(0.003)***	0.631	(0.452)	2.806	(0.183)
Critical thinking (summary measures)	-0.093	(0.054)*	-0.019	(0.723)	0.085	0.191
Observations	63		63		126	

Notes: Each entry for dependent variables is the odd-ratio (critical thinking summary measures: estimated coefficient) and p-value of the independent variable FC or not.
* **, ***; Significance difference at 10 percent, 5 percent, and 1 percent level of confidence.



Initial test, second test and Attendance: Principal axis (left).
 Difference Second-Initial test: Secondary axis (right).

Figure 1. Performance in attendance and grades (average data).
 Initial test, second test, and attendance: principal axis (left).
 Difference second-initial test: secondary axis (right).

results. On the other hand, the data show that the use of the FC was helpful to align the slightly different levels of knowledge detected through the initial test among students in the two groups. Nevertheless, a similar progress took place also in the group that was not exposed to the FC.

For this reason, our results corroborate those studies that did not identify any specific effect of the FC on students' performance (McLaughlin et al. 2013). Especially when compared with other active learning techniques, the FC alone might not be sufficient to determine students' performance. Students exposed to the FC would perform equally well compared to students exposed to other active learning formats (Jensen, Kummer, and Godoy 2015). We, thus, consider it safer to conclude that the teaching technique is not necessarily the factor that produces a "discernible differentiated effect on student learning" (Powner and Allendoerfer 2008, 85). A wide array of techniques, resources, or formats could similarly improve students' performance. In line with Jensen et al. (2015, 9), the effects of the FC are less evident when both classes are based on an active learning approach. This makes it important to focus on the "active-learning style of instruction rather than the order in which the instructor participated in the learning process." In some studies, the FC is an effective technique for improving students' performance, when implemented as a complement to other "cooperative learning activities," (Foldnes 2016) or when students use it as an instrument to monitor their own learning performance (Lai and Hwang 2016).

In terms of attendance, it was not possible to contrast our results with studies that compared different active learning techniques because all the studies that we found were based on comparisons between fully traditional and active learning formats. Nevertheless, studying attendance in an FC is important because it relates with one of the main concerns of the literature, which is whether recorded video classes reduce the students' will to go to class. Our tentative results corroborate the studies that did not find any specific correlation, negative or positive, between these

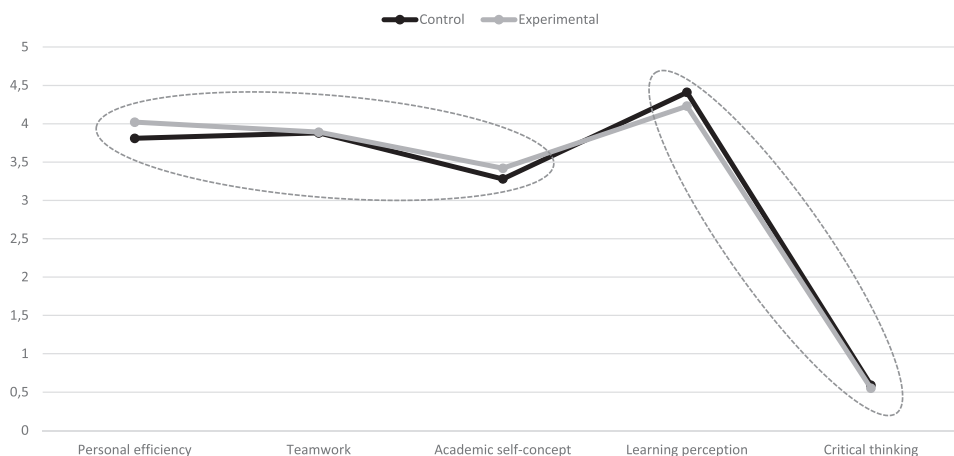


Figure 2. Performance in soft skills (average data last day).

types of classes and attendance (Copley 2007; Von Konsi, Ivins, and Gribble 2009). Even though our data visualization shows that students exposed to the FC slightly outperformed the students of the control group (figure 1), the chosen statistical techniques did not detect a substantial impact of the FC on attendance. The significant correlation we identified in our study was between the average grade achieved in the “Diploma in Development of Personal, Communication, and Professional Abilities” and attendance. The fact that students recording a high academic performance in this diploma were also those who attended class the most could mean that the interpersonal abilities provided by this type of education generate a stronger motivation to go to class. Nevertheless, due to the specific nature of this diploma, we cannot relate this conclusion to any previous study.

These results do not allow us to conclude that the FC *per se* improves students' attendance (Chen et al. 2014; Elmaadaway 2017). However, they do not corroborate either argument that “by viewing the content screencasts alone,” students will regard attendance as “optional” (Traphagan, Kucsera, and Kishi 2010; Yeung and O'Malley 2014, 63; Yamarick 2019). While new data should be collected on this point, our experiment suggests that “fears that the increasing availability of technology-enhanced educational materials have a negative impact on lecture attendance seem unfounded.” (Billings-Gagliardi and Mazor 2007, 573).

Our study did not detect a substantial impact of the FC on the soft skills either. Even though one can observe some small differences in some of the items in table 5, such differences were not significant for the most part.

With regard to self-efficacy, as table 5 suggested, the only significant improvement with the FC was in item 8, which refers to the sense of efficacy that students felt when doing projects or homework. The data visualization confirmed this point, in terms of the average data of this soft skill (figure 2). However, there are no other significant differences. While our results cannot be related to any previous study in the field of politics or IS, we can tentatively conclude that flipping the classroom is not necessarily beneficial to improve students' self-efficacy. This conclusion is in line with those studies that compared different types of FCs and discovered that the technique in itself is neutral in terms of the effects on students' self-efficacy. Other aspects, for example, the way in which the FC is administered, seem to be more relevant (Lai and Hwang 2016). One of the promises of the FC is to empower students to become more aware of their learning. The FC has the potential to transform the classroom into a more individualized experience where students learn at their own pace and apply their knowledge to solve problems under the supervision

of the teacher and in collaboration with their classmates (Baker 2000; Bergmann and Sams 2012). Nevertheless, for the FC to improve self-efficacy, it needs to be designed with that specific purpose, for example, by providing students with instruments to assess their own learning process (Lai and Hwang 2016). Future studies should assess the best strategies to achieve this goal.

Academic self-concept has not been explored in many studies either. For example, King (1993) expected collaborative learning to have positive effects on students' self-concept. Yet, this conclusion was not corroborated by any systematic study. Our empirical analysis did not provide evidence in favor of this expectation. On one hand, item 19 in table 5, which refers to the feeling of being a very good student, suggested that the control group improved while the FC group worsened. On the other hand, as figure 2 shows, the FC group performed slightly better in terms of average data. However, the differences were quite small and did not allow the argument that the FC helped improve the academic self-concept in any significant way. We do not deny that active learning can be beneficial for increasing students' self-confidence to perform complex tasks and to trust in their capacities to obtain the expected outcomes (Gresham 1995). However, it is not clear that the FC is the best candidate to improve this aspect.

As to the perception of learning, in our study, figure 2 shows a slightly better performance of the control group. This initial finding was not surprising. Other scholars observed that students can associate active learning with an "increased cognitive effort" and with the sensation of learning less (Missildine et al. 2013; Deslauriers et al. 2019, 19251). Nevertheless, none of the items in table 5 shows significant differences between the two groups. This can mean that active learning techniques, such as the FC, are not the main determinants of students' perception of learning. A positive or a negative perception of learning can depend, for example, on the professor's capacity to give students adequate feedback after the explanation of contents and the performance of active learning activities, or on his/her competence in managing the necessary technology (Murillo, Sánchez, and Godoy-Caballero 2019).

The difficulty reaching a conclusion on this point reminds us of the complexity of measuring skills learned from a teaching technique, particularly as we rely on students to perceive their own skill advancement. Taking students' feedback at face value is problematic because it can be difficult for students to evaluate how much or how well they are learning. Along these lines, McNally et al. (2017, 292) argued that "preferences alone may not be the most informative aspect on which to evaluate a flipped classroom environment." Preferences can depend on "a number of different factors," such as "novelty," which are not related to any pedagogical element. Nevertheless, we do not think that this is enough a reason not to study students' perceptions. Students are the main actors of any learning process. It is, thus, important to explore their perceptions about what techniques can hinder or favor their learning. Studying students' perceptions of progress remains a valid instrument to measure their personal and academic progress.

In terms of teamwork, our results do not corroborate those studies that found a positive impact of the FC, in comparison with traditional teaching (Karabulut-Ilgu et al. 2018) or other active learning techniques (Foldnes 2016). However, our results cannot corroborate either of those studies that found that the FC can reduce the students' disposition to work in group (Jenkins 2015). In our study, the control group performed slightly better than the FC group in one item (12), which refers to the acceptance and compliance with the rules agreed upon in the group, such as deadlines. This is somehow similar to one of Shannon Jenkins' (2015, 610) conclusions. After implementing an FC in an "Introduction to American Politics" course, she observed that students were "less positively inclined" toward working in groups. However, single questionnaire items can hardly show differences for an entire category. Moreover, as our data visualization shows in terms of average data of the last day survey, the two groups performed virtually in the same way. Along these lines,

the main tendency was that the FC did not have a significant impact on teamwork. Further exploration based on larger samples will need to validate this initial finding, because it would go against one of the main promises of active learning, the ability to empower students to become responsible for their own work.

As for critical thinking, the consensus on the positive effects of active learning and the FC is so far unanimous. Several studies concluded that the FC tends to have a positive impact on students' critical thinking (Kong 2014; Hsieh, Vivian Wu, and Marek 2017). Nevertheless, our study could only in part confirm this. In table 5, students exposed to the FC performed significantly better in comparison with their results of the first day, as indicated by items 24 and 28, related to their capacity to use logical and critical thinking. However, figure 2 shows a slightly better performance of the control group in terms of the average data of the last day. In conclusion, the differences between the two formats were not significant. This substantially null effect could lead to a partial revision of one of the main conclusions of the literature on the positive impact of the FC on critical thinking. Far from arguing that our study can refute this conclusion, it would be beneficial to further test these existing expectations. Critical thinking is not only associated with the capacity to evaluate information in an independent way, but also with the capacity to deepen the knowledge of basic concepts. The capacity of active learning techniques to improve this aspect is yet to be demonstrated.

Without forgetting the exploratory nature of our study, which needs to be expanded by using larger samples over longer time spans, the main pattern that we observe in our data is that the FC did not have a substantial impact on either students' performance, attendance, or soft skills. The differences between the two groups were too small to corroborate any tendency in favor of one or the other format.

These results can indicate that it is not enough to flip a class or use a certain active learning technique to improve learning. Sometimes while studying the existing literature on active learning, one has the impression that teaching formats are analyzed in an impersonal way. The teachers' skills are not always taken into account in terms of management of the techniques or his/her capacity to motivate students to improve. These skills are crucial to "think carefully about how to deliver content outside the class, how to ensure that students understand that content, and how to integrate it with in-class activities" (Jenkins 2015, 610). With this, we do not argue against the possible utility of the FC or other active learning formats. There is no question that the possibility for students to acquire some knowledge of the contents before class can save precious time in the classroom. This is highly beneficial for both students and teachers, as time can be dedicated to the resolution of problems or doubts related to the contents of the class (Enfield 2013). Moreover, the possibility for students to watch video lectures when and where they find it more convenient can be an efficient way to use time outside the classroom (Bergmann and Sams 2012). However, our study is an invitation to implement these techniques with caution, without disregarding the centrality of the teachers and their competences for the learning process (Jenkins 2015; Yen et al. 2018). For these techniques to be effective, they need to be carefully planned and implemented in a gradual rather than abrupt way. In addition, not all subjects or teaching styles might be equally amenable to these techniques (Lambach, Karger, and Goerres 2017, 564).

One of the reasons some studies do not seem to assign sufficient importance to these aspects can reside in the type of comparisons conducted. Comparing active learning with traditional teaching formats can motivate instructors to make their teaching more effective. Nevertheless, the differences between traditional and active learning can make it difficult to understand the effects of the two formats. Moreover, it can lead to the obvious conclusion that active is better than passive. This does not significantly increase our knowledge of what aspects of active learning are more useful or viable to improve teaching. More comparisons between the FC and

other active learning techniques could give us a clearer picture of their viability and usefulness.

In our study, the implementation of a different active learning format in each group made the benefits of the FC unclear (Jensen, Kummer, and Godoy 2015). For the most part, students performed in a substantially similar way. This can suggest that the benefits of an FC format might be less, when compared with other active learning techniques.

Conclusions, Limitations and Future Research

The influence of the FC on students' achievement, attendance, and soft skills has, so far, been studied most extensively in the fields of education, natural science, and languages. For this reason, our study can offer a relevant contribution to understanding the implications of this active learning technique in the field of Politics, and, more specifically, IS. The FC has been acquiring importance due to its capacity to connect students who cannot be physically present in a classroom. In times of the COVID-19 pandemic, the use of video lectures and other interactive technology can greatly help share academic knowledge.

We are also aware of the limitations of this study. The experiment explored the impact of the FC only in one section of a single class over a single semester. The need to avoid the sensation of a too disruptive change among students mostly motivated this gradual and partial implementation. However, we are aware that such a limited implementation can reduce the possibility of identifying clear effects. For this reason, it is necessary to repeat the experiment, using larger samples of participants and longer periods of the class and relying on different analytical techniques, possibly based on a mixed method approach (Betti, Biderbost and García Domonte 2020). This approach is useful to obtain more reliable knowledge because it allows us to integrate the statistical analysis with other techniques, such as student focus groups. Although it was not possible to do so in this study, due to time and financial limitations, it is our intention to use this mixed approach in future research.

Nevertheless, we believe that our experiment, as an exploratory study, provides a useful starting point. Future studies should aim to validate several hypotheses as to how students' soft or hard skills can be improved by active learning techniques. These studies should also assess whether the FC can be beneficial for lower levels of cognitive work, such as content retention and understanding, or for higher levels, such as analysis and evaluation. Moreover, it would be important to identify the most effective components of the FC or what teaching strategies can make it more useful. Since it is still a partially unexplored field, the available studies on whether the FC can favor or hinder students' critical skills lead to conflicting conclusions, requiring further research and analysis.

Finally, future studies should examine the possible drawbacks of the FC. Some studies have started to discuss the risk that recorded video lectures reduce students' capacity to maintain attention, for example, due to the impossibility to interrupt the teacher with questions (Toto and Nguyen 2009). Students used to a traditional model of teaching, based on frontal lectures, might feel abandoned or unable to adequately follow the explanations of the contents (Talbert 2017). This is a particularly sensitive issue in times of rolling lockdowns and online teaching, which can increase the digital divide among students with unequal technology access. In terms of teachers, future studies should assess the economic impact of investing resources to train professors in the management of the technology necessary to produce video and other interactive contents (Collins and Halverson 2010). Acquiring these skills can also mean a considerable workload for the teacher (Tunks 2012). Our article could not deal with these issues, but future research should assess these possible drawbacks.

Acknowledgments

The Pontifical University of Comillas, ICAI-ICADE (Madrid) financed the research for this article (Name of the research project: "Teaching Innovation in International Relations: A Comparative Study of the Flipped Classroom and the Traditional Classroom").

Conflict of interest

The authors have no conflicts of interest to declare.

References

- ANDERSON, LORIN W., DAVID KRATHWOHL, PETER AIRASIAN, KATHLEEN A. CRUIKSHANK, RICHARD E. MAYER, PAUL PINTRICH, JAMES RATHS, AND MERLIN C. WITTRICK, eds. 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- AWIDI, ISALAH T., AND MARK PAYNTER. 2019. "The Impact of a Flipped Classroom Approach on Student Learning Experience." *Computers & Education* 128: 269–83.
- BAKER, WESLEY J. 2000. "The Classroom Flip: Using Web Course Management Tools to Become the Guide by the Side." In *Selected Papers from the 11th International Conference on College Teaching and Learning*, edited by Jack A. Chambers, 9–17. Jacksonville, FL: Florida Community College at Jacksonville.
- BANDURA, ALBERT. 1994. "Self-efficacy." In *Encyclopedia of Human Behavior*, vol. 4, edited by Ramachandran Vylayanur, 71–81. New York: Academic Press.
- BERGMANN, JONATHAN, AND AARON SAMS. 2012. *Flip your Classroom: Reach Every Student in Every Class Every Day*. Eugene: ISTE.
- BETTI, ANDREA, PABLO BIDERBOST, AND AURORA GARCÍA DOMONTE. 2020. "Flipping the Classroom in Political Science: Student Achievement and Satisfaction." *Revista de Ciencia Política* 40 (3): 589–615.
- BILLINGS-GAGLIARDI, SUSAN, AND KATHLEEN M. MAZOR, 2007. "Student Decisions about Lecture Attendance: Do Electronic Course Materials Matter?" *Academic Medicine* 82 (10): 573–76.
- BURCH, KURT. 2000. "A Primer on Problem-Based Learning for International Relations Courses." *International Studies Perspectives* 1 (1): 31–44.
- CHEN, YUNGLUNG, YUPING WANG, KINSHUK, AND NIAN-SHING CHEN. 2014. "Is FLIP Enough? Or Should We Use the FLIPPED Model Instead?" *Computers & Education* 79: 16–27.
- COLLINS, ALLAN, AND RICHARD HALVERSON. 2010. "The Second Educational Revolution: Rethinking Education in the Age of Technology." *Journal of Computer Assisted Learning* 26: 18–27.
- COPLEY, JONATHAN. 2007. "Audio and Video Podcasts of Lectures for Campus-Based Students: Production and Evaluation of Student Use." *Innovations in Education and Teaching International* 44 (4): 387–99.
- DAVIES, RANDALL S., DOUGLAS L. DEAN, AND NICK BALL. 2013. "Flipping the Classroom and Instructional Technology Integration in a College-Level Information Systems Spreadsheet Course." *Educational Technology Research and Development* 61 (4): 563–80.
- DESLAURIERS, LOUIS, LOGAN S. MCCARTHY, KELLY MILLER, KRISTINA CALLAGHAN, AND GREG KESTIN. 2019. "Measuring Actual Learning Versus Feeling of Learning in Response to Being Actively Engaged in the Classroom." *Proceedings of the National Academy of Sciences* 116 (39): 19251–57.
- ELMAADAWAY, MOHAMMED A. 2017. "The Effects of a Flipped Classroom Approach on Class Engagement and Skill Performance in a Blackboard Course." *British Journal of Educational Technology* 49 (3): 479–91.
- ENFIELD, JACOB. 2013. "Looking at the Impact of the Flipped Classroom Model of Instruction on Undergraduate Multimedia Students at CSUN." *TechTrends* 57 (6): 14–27.
- FLORES, OSCAR, ISABEL DEL-ARCO, AND PATRICIA SILVA. 2016. "The Flipped Classroom Model at the University: Analysis Based on Professors' and Students' Assessment in the Educational Field." *International Journal of Educational Technology in Higher Education* 13 (1): 1–12.
- FOLDNES, NJAL. 2016. "The Flipped Classroom and Cooperative Learning: Evidence from a Randomised Experiment." *Active Learning in Higher Education* 17 (1): 39–49.
- GRESHAM, FRANK M. 1995. "Student Self-Concept Scale: Description and Relevance to Students with Emotional and Behavior Disorders." *Journal of Emotional and Behavioral Disorders* 3 (1): 19–26.
- HAMDAN, NOORA, PATRICK MCKNIGHT, KATHERINE MCKNIGHT, AND KARI AREFSTROM. 2013. "A Review of Flipped Learning." *Flipped Learning Network* 1–20.
- HECKMAN, JAMES J., AND TIM D. KAUTZ. 2012. "Hard Evidence on Soft Skills." Working Paper 18121. National Bureau of Economic Research. <http://www.nber.org/papers/w18121>.

- HSIEH, JUN S. C., WEN-CHI VIVIAN WU, AND MICHAEL W. MAREK. 2017. "Using the Flipped Classroom to Enhance EFL Learning." *Computer Assisted Language Learning* 30 (1–2): 1–21.
- JENKINS, SHANNON. 2015. "Flipping the Introductory American Politics Class: Student Perceptions of the Flipped Classroom." *PS: Political Science and Politics* 48 (4): 607–11.
- JENSEN, JAMIE L., TYLER A. KUMMER, AND PATRICIA GODOY. 2015. "Improvements from a Flipped Classroom May Simply Be the Fruits of Active Learning." *Life Sciences Education* 14 (1): 1–12.
- JENSEN, SCOTT A. 2011. "In-Class versus Online Video Lectures: Similar Learning Outcomes, but a Preference for In-Class." *Teaching of Psychology* 38 (4): 298–302.
- KARABULUT-ILGU, ALIYE, SUHAN YAO, PETER SAVOLAINEN, AND CHARLES JAHREN. 2018. "Student Perspectives on the Flipped-Classroom Approach and Collaborative Problem-Solving Process." *Journal of Educational Computing Research* 56 (4): 513–37.
- KING, ALISON. 1993. "From Sage on the Stage to Guide on the Side." *College Teaching* 41 (1): 30–35.
- KONG, SIU C. 2014. "Developing Information Literacy and Critical Thinking Skills through Domain Knowledge Learning in Digital Classrooms: An Experience of Practicing Flipped Classroom Strategies." *Computers & Education* 78: 160–73.
- LAI, CHIU-LIN, AND GWO-JEN HWANG. 2016. "A Self-Regulated Flipped Classroom Approach to Improving Students' Learning Performance in a Mathematics Course." *Computers & Education* 100: 126–40.
- LAMBACH, DANIEL, CAROLINE KARGER, AND ACHIM GOERRES. 2017. "Inverting the Large Lecture Class: Active Learning in an Introductory International Relations Course." *European Political Science* 16 (4): 553–69.
- LARKIN, HELEN E. 2010. "'But They Won't Come to Lectures ...' The Impact of Audio Recorded Lectures on Student Experience and Attendance." *Australasian Journal of Educational Technology* 26 (2): 238–49.
- MARSH, HERBERT W., AND ANDREW J. MARTIN. 2011. "Academic Self-Concept and Academic Achievement: Relations and Causal Ordering." *British Journal of Educational Psychology* 81 (1): 59–77.
- MAZUR, AMBER D., BARBARA BROWN, AND MICHELE JACOBSEN. 2015. "Learning Designs Using Flipped Instruction." *Canadian Journal of Learning and Technology* 41 (2): 1–26.
- MCLAUGHLIN, JACQUELINE E., LA TOYA M. GRIFFIN, DENISE A. ESSERMAN, CHRISTOPHER A. DAVIDSON, DYLAN M. GLATT, MARY T. ROTH, NASTARAN GHARKHOLONAREHE, AND RUSSELL J. MUMPER. 2013. "Pharmacy Student Engagement, Performance, and Perception in a Flipped Satellite Classroom." *American Journal of Pharmaceutical Education* 77 (9): 196.
- MCNALLY, BRENTON, JANINE CHIPPERFIELD, PAT DORSETT, LETITIA DEL FABBRO, VALDA FROMMOLT, SANDRA GOETZ, AND JOANNE LEWOHL et al. 2017. "Flipped Classroom Experiences: Student Preferences and Flip Strategy in a Higher Education Context." *Higher Education* 73: 281–98.
- MISSILDINE, KATHY, FOUNTAIN REBECCA, LYNN SUMMERS, AND KEVIN GOSSELIN. 2013. "Flipping the Classroom to Improve Student Performance and Satisfaction." *Journal of Nursing Education* 52 (10): 597–99.
- MURILLO, LUIS R., JOSÉ ÁNGEL LÓPEZ SÁNCHEZ, AND ANA LUISA GODOY-CABALLERO. 2019. "How the Flipped Classroom Affects Knowledge, Skills, and Engagement in Higher Education: Effects on Students' Satisfaction." *Computers and Education* 141.
- NOURI, JALAL. 2016. "The Flipped Classroom: For Active, Effective, and Increased Learning—Especially for Low Achievers." *International Journal of Educational Technology in Higher Education* 13: 1–10.
- POWNER, LEANNE C., AND MICHELLE G. ALLENDOERFER. 2008. "Evaluating Hypotheses about Active Learning." *International Studies Perspectives* 9: 75–89.
- PRINCE, MICHAEL. 2004. "Does Active Learning Work? A Review of the Research." *Journal of Engineering Education* 93 (3): 223–31.
- ROBLES, MARCEL M. 2012. "Executive Perceptions of the Top 10 Soft Skills Needed in Today's Workplace." *Business Communication Quarterly* 75 (4): 453–65.
- ROEHLING, PATRICIA, LINDSEY M. ROOT LUNA, FALLON J. RICHIE, AND JOHN J. SHAUGHNESSY. 2017. "The Benefits, Drawbacks, and Challenges of the Flipped Classroom in an Introduction to Psychology Course." *Teaching of Psychology* 44 (3): 183–92.
- SALAS, EDUARDO, DANA E. SIMS, AND SHAWN C. BURKE. 2005. "Is There a Big Five in Teamwork?" *Small Group Research* 36 (5): 555–99.
- TALBERT, ROBERT. 2017. *Flipped Learning: A Guide for Higher Education Faculty*. Sterling: Stylus.
- TANG, TAO, ATEF M. ABUHMALD, MELAD OLAMAIT, DANA M. OUDAT, MAGED ALDHAEABI, AND EBRAHIM BAMANGER. 2020. "Efficiency of Flipped Classroom with Online-Based Teaching under COVID-19." *Interactive Learning Environments*. doi: 10.1080/10494820.2020.1817761.
- TOTO, R., AND DUC H. NGUYEN. 2009. "Flipping the Work Design in an Industrial Engineering Course." In *ASEE/IEEE Frontiers in Education Conference*, San Antonio, TX.
- TOUCHTON, MICHAEL. 2015. "Flipping the Classroom and Student Performance in Advanced Statistics: Evidence from a Quasi-Experiment." *Journal of Political Science Education* 11 (1): 28–44.
- TRAPHAGAN, TOMOKO, JOHN W. KUCSERA, AND KYOKO KISHI. 2010. "Impact of Class Lecture Webcasting on Attendance and Learning." *Educational Technology Research and Development* 58: 19–37.

- TRUBY, KATHERINE, MEREDITH L. WEISS, AND DAVID L. ROUSSEAU. 2014. "Teaching the Unfamiliar to a Crowd." *PS: Political Science and Politics* 47 (1): 189–94.
- TUNKS, KARIN W. 2012. "An Introduction and Guide to Enhancing Online Instruction with Web 2.0 Tools." *The Journal of Educators Online* 9 (2): 1–16.
- VON KONSKI, BRIAN, JIM IVINS, AND SUSAN J. GRIBBLE. 2009. "Lecture Attendance and Web-Based Lecture Technologies: A Comparison of Student Perceptions and Usage Patterns." *Australasian Journal of Educational Technology* 25 (4): 581–95.
- WHITMAN COBB, WENDY N. 2016. "Turning the Classroom Upside Down: Experimenting with the Flipped Classroom in American Government." *Journal of Political Science Education* 12 (1): 1–14.
- YAMARICK, STEVEN. 2019. "Flipping the Classroom and Student Learning Outcomes: Evidence from an International Economics Course." *International Review of Economics Education* 31: 1–12.
- YEN, SHU-CHEN, YAFEN LO, ANGELA LEE, AND JUDELMAY ENRIQUEZ. 2018. "Learning Online, Offline, and In-Between: Comparing Student Academic Outcomes and Course Satisfaction in Face-To-Face, Online, and Blended Teaching Modalities." *Education and Information Technologies* 23 (5): 2141–53.
- YEUNG, KAY, AND PATRICK J. O'MALLEY. 2014. "Making the Flip Work: Barriers to and Implementation Strategies for Introducing Flipped Teaching Methods into Traditional Higher Education Course." *New Directions* 10 (1): 59–63.

© 2022 International Studies Association. Copyright of International Studies Perspectives is the property of Oxford University Press / USA and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.