

An Action Research Study from Implementing the Flipped Classroom Model in Primary School History Teaching and Learning

Vasiliki Aidinopoulou^{1,2} and Demetrios G. Sampson^{3,4*}

¹Greek-French School Jeanne d' Arc, Greece // ²Department of Digital Systems, University of Piraeus, Greece //

³School of Education, Curtin University, Australia // ⁴Information Technologies Institute, CERTH, Greece // aidinopoulou@gmail.com // demetrios.sampson@curtin.edu.au

*Corresponding author

ABSTRACT

The benefits of the flipped classroom (FC) model in students' learning are claimed in many recent studies. These benefits are typically accounted to the pedagogically efficient use of classroom time for engaging students in active learning. Although there are several relevant studies for the deployment of the FC model in Science, Technology, Engineering and Maths (STEM) subjects, and at Higher Education and/or High School, there are very few works studying FC in social studies and at primary school level. This paper presents an action research focused on the implementation of the FC model in teaching social studies in primary school. The main scope of this action research, conducted over an entire school year with two different History classes (one representing the experimental group that followed the FC model and the other representing the control group following the traditional lecture based approach) was to compare the use of classroom time for student-centered learning activities and the resulted learning outcomes related to both traditional learning goals of a history course (that is, memorization of historical content) and more ambitious ones such as the cultivation of historical thinking skills (HTS). The study revealed that indeed, the classroom based sessions of the experimental group were used for engaging student-centered activities and that this resulted into better learning outcomes in terms of demonstrating critical HTS. Thus, this initial action research provides encouraging evidences for the potential benefits of the FC model in primary school social studies courses.

Keywords

Flipped classroom model, Primary school, History teaching, Historical thinking skills

Introduction

Traditional teaching of social studies at school level is challenged by both the use of technology and the adoption of inquiry-based teaching strategies in other subjects, such as Science, Technology, Engineering, and Math (STEM) (Bishop & Verleger, 2013; Hwang et al., 2015; Keengwe & Onchwari, 2015). Typically, in traditional history/social studies school curricula, emphasis is given in memorizing large amounts of historical content namely, names, dates and facts, making these courses less attractive to students (Fielding, 2005). Furthermore, still, many history/social studies school teachers adopt traditional teaching strategies, using most of the classroom time for lecturing and assessing students' ability to memorize content. As a result, students do not actively engage in learning and assessment activities that promote their historical/critical thinking development (Gaughan, 2014). It is also argued that students have common misconceptions about historical knowledge, primary sources, human motivation and historical change which are not easily overcome with traditional teaching strategies (Epstein, 2012). As a result, there are systematic efforts to enhance teaching, learning and assessment of history and social studies at K-12, by exploiting innovative pedagogical designs supported by digital technologies (Lyons, 2008).

On the other hand, the FC model has gained prominence over the past years, as a technology-supported pedagogical innovation which uses classroom time for students to actively engage in interactive learning activities, including personalized feedback and scaffolding from the teacher, while teachers' traditional lecturing is delivered out of the classroom time with asynchronous video lectures (Chen et al., 2014). Despite the wide take-up of the FC model in Higher Education and STEM subjects (Bishop & Verleger, 2013; Keengwe & Onchwari, 2015; Sergis et al., 2017), there are limited efforts in studying its application in primary school and in history/social studies courses, in particular.

In this paper, it is argued that adopting the FC model in a primary school history course has the potential to use classroom time in a more efficient way, leading to enhanced students' learning experiences and outcomes. Therefore, the paper reports on the design and results of an action research implemented to investigate this hypothesis and provide evidence on the added value of the FC model.

The remainder of the paper is structured as follows: the Background section presents an overview of the FC model and its implementation in teaching history/social studies school courses. The Action Research Method

section presents the methodology and research questions of the action research. The Educational Design section describes the design of the educational intervention used in the action research. The Results section presents the findings obtained in relation to the two research questions and, finally, the Discussion and Future Work section discusses the lessons learnt and outlines potential future research.

Background

The FC model is an emerging blended learning model widely used both in school and university formal educational settings. Based on sound pedagogical theoretical principles, FC targets to exploit classroom time and space for appropriately designed interactive learning activities differentiated according to individual and group students' needs (DeLozier & Rhodes, 2016). This section presents an overview of the use of digital technologies in school history teaching and learning, and the implementation of the FC model in K-12 history/social studies.

Technology supported history teaching and learning in K-12

As in all disciplines, the integration of educational technologies in social studies is constantly increasing (Lee & Friedman, 2009). Over the past decade, many history teachers are reported to be hesitant to adopt digital technologies into their classrooms and still rely on the textbook and their own lectures for teaching history (Scheuerell, 2015). Surveys showed that some teachers were skeptical about the effectiveness of technology, while others expressed concerns about the amount of time they had to devote in order to design new activities and material and also the technological competence they should acquire (Townsend, 2010).

On the other hand, research studies have demonstrated the potential of digital technologies in social studies at school level in relation to the development of historical critical thinking skills (Green et al., 2013). To this end, a wide range of digital technologies have been used in history classrooms. Digital libraries, for example, allow teachers and students to freely view and download primary and secondary sources for use in the classroom, providing students with the possibility to examine different perspectives of historical context (Scheuerell, 2015). Simulations, forums, wikis, blogs, documentary film making, social media, whiteboards, web quests and voting technologies are tools that history teachers have been using in an effort to transform the history classroom into a student-centered environment (Haydn et al., 2014). Thus, it is established that K-12 history teaching and learning is now widely exploiting digital technologies.

Using the flipped classroom model in history teaching and learning in K-12

The FC model has been only sporadically implemented in history and social studies at middle and/or high school courses. Typically, in such implementations, students watch video lectures before coming to class. The purpose of these videos is to present the historical content, in addition to the material presented in the course textbook. These video lectures are highly connected with classroom time and are meant to trigger student's interest for the classroom activities that will follow (Gaughan, 2014). The video lectures are usually made by the teachers with narration, text and enriched with maps, annotations and images. For the creation and distribution of the videos, widely available tools are used, such as iTunes and YouTube (Kotlik, 2014).

On the other hand, the classroom-based face-to-face sessions are used for discussions on the previously watched video lectures, and other student-engaging tasks, such as primary and secondary sources analysis, debating, peer reviewing or simulations (Bergmann & Sams, 2012). From case studies, it is reported that more history teachers tend to adopt the popular flipped-mastery system, according to which, students have to prove mastery in content using their own pace and means of proving academic performance (Bergmann & Sams, 2014).

In history courses that use the FC model, emphasis is given in the cultivation of historical thinking skills. As students interact with material, they learn how to use historical context and perspectives different to their own, towards developing critical historical thinking skills, they analyze and hypothesize, and in general act as historians (Gaughan, 2014; Bergmann & Sams, 2015). Consequently, the memorization of historical content is only a minor learning goal as most goals concern skill development (Fulton, 2014).

Regarding the evaluation of the FC teaching approach in history courses, studies report improvements in students' learning outcomes both in historical content memorization as well as in the cultivation of historical critical thinking skills. Furthermore, students and teachers demonstrate a positive attitude towards the FC model

as it increases interest and engagement and makes students more responsible in their studies (Fielding, 2005; Gaughan, 2014; Bergmann & Sams, 2015).

In this context, the present study reports on the design and implementation of an action research in order to provide additional evidence of the impact of the FC model on the exploitation of face-to-face classroom-based teaching time as well as the students' learning outcomes. The following section outlines the action research method adopted in this work.

The action research method

The action research was designed and conducted using McKernan's model (McKernan, 1991). This model, which is a form of scientific inquiry (McKernan, 1988), meets Lewin's four original action research phases: plan, act, observe and reflect (Lewin, 1948), but also adds a spiral flow that allows the researcher to reflect on and redesign his action research. It gives the opportunity to the action researcher to go through all the phases again, via a new action cycle and bring additional elements into the study. This section presents the general purpose and research questions of the study as well as the study's context and characteristics of participants. In addition, the phases of the two action research cycles are described. The section is concluded with the description of the study's instruments and the data collection and analysis methods.

Research questions

The general purpose of the action research was to investigate the extent to which the implementation of the FC model can improve students' learning outcomes and also lead to better use of classroom time. More specifically, our scope was to compare the use of classroom time for student-centered learning activities and the resulted learning outcomes related to both traditional learning goals of a history course (that is, memorization of historical content) and more ambitious ones such as the cultivation of historical thinking skills (HTS), between two different History classes (one following the FC model and the other following the traditional lecture-based approach). As a result, the following research questions were defined:

- **RQ1.** Does the implementation of the FC model in a primary school history course lead to the use of classroom time for more student-centered activities?
- **RQ2.** Does this contribute to better students' learning outcomes compared to traditional teaching strategies?

The first research question concerns the exploitation of classroom based teaching time. We argue that in history teaching better use of classroom time is achieved when it accommodates student-centered activities that promote cognitive goals of the higher levels of Bloom's revised taxonomy such as Analyze, Evaluate and Create (Krathwohl, 2002), as opposed to activities that promote cognitive goals of the lower levels of Bloom's taxonomy, such as Remembering. The FC model has been attributed with the capacity to allow the teacher to facilitate such learning and assessment activities since it releases time by moving lecture-based content delivery out of the classroom time (Bergman & Sams, 2012). For the second research question to be investigated, the selection of learning goals is required. The specific learning goals of the action research are explicitly described in the "Educational Design" section.

Study context

The study was conducted in a two-term-long (24 weeks) history course at primary school. More specifically, the "Roman and Byzantine History" course in the 5th Grade of the Greek primary school national curriculum was the context of this study. The course's syllabus, in brief, covers a long period of sixteen centuries (146 BC - 1453 AD) from the subjugation of Greece by the Romans to the conquest of Constantinople by the Turks. This includes extensive names, dates and facts that have to be memorized by the students and, moreover, the textbook involves challenging vocabulary in relation to the so far cognitive experiences of students. Thus, a teacher who needs to follow such a syllabus faces all the difficulties outlined in the previous sections. The action research aimed to investigate whether the implementation of the FC model could provide a solution to this educational problem.

Participants

The participants of the study were 49 eleven-year old students who attended the “Roman and Byzantine History” course at the Greek – French “Jeanne d’ Arc” Primary School in Greece in two different classes of Grade 5. From these, 26 students (11 boys, 15 girls) formed the experimental group and 23 students (7 boys, 16 girls) formed the control group. These two groups were selected because (a) all students had access to a computer and an internet connection and (b) they demonstrated similar learning achievements in the history courses that they had attended during the previous two school years. Thus, overall, we consider that the selected groups did not present any significant differences in number of students in each group, gender distribution, or prior knowledge that could bias the results. All students were familiar with the use of computers but were unfamiliar with the Moodle (<https://moodle.org/>) learning management system (LMS) that was used to facilitate the implementation of the FC model. Therefore, classroom time had to be devoted for students to get acquainted with the LMS and their new role in the new teaching strategy.

Procedure

The action research was conducted in two action cycles following the McKernan spiral model as Table 1 depicts. The first cycle took place in four phases. First phase was the Plan. It was conducted from July - August 2014 before the beginning of the first term of the 2014 school year. During this phase, the educational design took place. Second and third phases in the research were Action and Observation. They were completed during the first term from September 11, 2014 until December 10, 2014 when formally the first term ended. The term’s ending was followed by the Christmas holiday season during which the Reflection phase took place. From the Reflection phase, emerged the revised design of the educational intervention. The second cycle of the action research began with the formation of the revised educational design. The revised plan was implemented (Action and Observation) during the second term, from January 8, 2015 until March 10, 2015, when the second term formally ended. Then, followed the reflection of the second cycle and during the period March - June 2015 the data of both terms were gathered and results presented.

Table 1. Action research procedure (Time frame and phases)

1 st term - 1 st action research cycle	
Plan	July - August 2014 Selection of resources, creation of new material, design of activity flow and assessment methods.
Act and Observe	September – December 2014 Implementation of the FC model. Data collection from teacher logs and first term’s assessments.
Reflect and Re-plan	December 2014 (holiday season) Reflection on the first term’s observations and revision of educational design.
2 nd term - 2 nd action research cycle	
Act and Observe	March – June 2015 Implementation of the revised plan. Collection of data via teacher logs and second term’s assessments.
Reflect	June 2015 Analysis of both terms’ research data. Presentation of action research results.

Instruments

In order to collect data concerning the use of classroom time, teacher logs were created that documented for every teaching hour, the flow of learning activities and the time spent in each activity. The learning goal of the memorization of the historical content was assessed with the standardized tests of the National curriculum. In order to assess historical thinking skills cultivation, assessment activities were created by the teacher that included open and closed ended questions on primary and secondary sources. For the evaluation of the students’ answers the “Assessing Historical Thinking and Understanding (ARCH): Historical Thinking Skills Rubric” was used (UMBC Center for History Education, 2013). Each assessment question was linked to one learning goal and was evaluated using the rubric’s criteria. The tests were common for both groups and included the same grading scale (from 0 to 10).

Data analysis

Following data collection, the data analysis was divided into the following tasks: (a) analyze the distribution of classroom time in learning activities for the experimental and the control group during the first and second term; (b) analyze the students' learning outcomes in the historical content memorization assessments for the experimental and the control group in the first and second term; (c) calculate the Mann-Whitney U test and effect size in order to statistically examine the significance of differences in student historical content memorization achievement scores between the experimental and the control group; (d) analyze the student's learning outcomes in the HTS cultivation assessments for the experimental and the control group in the first and second term; (e) calculate the Mann-Whitney U test and the effect size in order to statistically examine the significance of differences in student HTS cultivation achievement scores between the experimental and the control group; (f) an overall evaluation to identify the benefits and possible challenges of the FC model. Mann-Whitney U tests were employed to investigate for potential significant differences in the assessment scores between the two groups, since the data did not follow a normal distribution and the sample size was not large. The calculation of the effect size was based on the standardized mean difference and evaluated based on Cohen's *d* coefficient (Cohen, 1994). Since the effect size provides a quantitative means to measure the "strength" of an observed case, the adopted "strength" intervals are defined as: *small* (effect size ≤ 0.3), *medium* ($0.3 < \text{effect size} \leq 0.5$), *large* ($0.5 < \text{effect size} \leq 0.7$) and *very large* ($0.7 < \text{effect size}$) (Cohen, 1994). These intervals will be exploited during the analysis of the results regarding students' learning outcomes (RQ2), so as to provide a measure of the impact of the FC model.

Educational design

The purpose of the educational design was to create the educational intervention to be used in the action research. This section describes the five phases (analysis, design, development, implementation, evaluation) of the ADDIE model (Dick, Carey & Carey, 2001) as applied to the study.

Analysis Phase: During this phase, the educational problem that primary school history courses face was identified. Matters such as the emphasis in historical content memorization as the prime learning outcome, insufficient teaching time for engaging students in learning activities that support historical thinking skills cultivation, lack of student interest and common student historical misconceptions, raised the need for a new approach that could improve history teaching and learning. The analysis phase also included the identification of the experimental and control groups' characteristics as described in the previous section, as well as the students' learning environment characteristics.

Design Phase: During this phase, learning goals were defined and the flow of activities and assessment methods were selected. The learning goals of the intervention were formed after studying the following international standards for teaching history: (a) the National Standards for Social Studies Teachers (Myers, 2000), (b) the Common Core State Standards for History/Social Studies - Grade 6-8 and (c) the UCLA Historical Thinking Standards (Common Core State Standards Initiative, 2010). These learning goals were common for both research groups and they were grouped into categories and subcategories:

- **Category A: Historical content memorization**, namely, to remember and recall names, dates and facts from the long-term memory.
- **Category B: Historical Thinking skills cultivation:**
 - **Understanding the concept of time**, namely, to distinguish between past, present and future time, to identify and use the temporal structure in a historical narrative, to measure and calculate calendar time, to interpret data presented in time lines and create time lines.
 - **Understanding historical sources**, namely, to identify the author of the historical document, to identify the historical facts of the source, to differentiate between historical facts and historical interpretations, to draw upon data in historical maps, visual, literary and musical sources.
 - **Historical analysis and interpretation**, namely, to analyze cause-and-effect relationships, to argue using historical evidence, to hypothesize on how different historical actions could have led to different results, to formulate historical questions, to evaluate the actions and decisions of historical persons based on their results.

Apart from the learning goals, the flow of activities is elaborated in the implementation phase. The selected assessment methods have been described in the previous section.

Development Phase: During this phase, new material was created to successfully support the implementation of the FC approach. First of all, a series of video lectures was created that presented the content material. The creation of videos followed standard design principles (Koumi, 2006) with the typical duration being 9-10 minutes and the lectures usually being divided into three parts (3 minutes each). The video lectures and all new content were provided to students via the Moodle LMS. In addition, activities were developed by exploiting Moodle, such as quizzes, wikis and forums for students to study and interact. All the above contributed to the application of the FC model.

Implementation Phase: During this phase, the educational intervention was implemented. The procedure of the intervention is presented in Table 2 and then further elaborated.

Table 2. Educational intervention procedure

Control group	Experimental group
<i>Before Class</i>	<i>Before Class</i>
	Studied material using the Moodle LMS
	Interacted at the forum
	Took quizzes
	Studied extra material related to the lesson (optional)
<i>During Class</i>	<i>During Class</i>
Lectures on the new historical content	Q&A session with the teacher
Q&A session with the teacher	Historical content memorization assessment
Teacher's feedback	Teacher's feedback
HTS cultivation activities (if time was available)	HTS cultivation activities
<i>After Class</i>	<i>After Class</i>
Studied material using the textbook and all digital resources except from the video lectures	Reward (digital badges)
Interacted at the forum	
Took quizzes	

The students of the experimental group prior to attending classroom activities studied the video lectures in their online digital classroom that was implemented using the Moodle LMS. They interacted at the forum by publishing questions, comments or answering to their classmates' questions. They also answered quizzes based on what they had previously studied. For those who had the time and interest, extra material was provided, relevant to the video lectures. During the face-to-face classroom sessions, first, the students were involved in Q&A activities with the teacher in order to solve any questions that remained unsolved. The teacher then assessed their historical content memorization by verbally asking questions. The assessment was followed by the teacher's feedback on students' answers. The remaining teaching time was spent in engaging with historical thinking skills (HTS) cultivation activities, such as collaborative activities and debates. The teacher observed the whole process and gave directions and feedback. At the end of each unit the assessments of the two main categories of learning goals took place. Students completed standardized tests with open and closed ended questions referring to historical content memorization and HTS cultivation. The assessments were followed by the teacher's feedback on students' learning outcomes and the students were rewarded with digital badges.

The students of the control group during the face-to-face classroom sessions were exposed to traditional lectures on the new historical content. After the lectures, Q&A sessions with the teacher followed in order to respond to questions that were created during the delivery of the new content. The remaining classroom time was invested in engaging with HTS cultivation activities. In most cases though, the remaining time was not sufficient. When the students of the control group returned home, they had to study and practice using their textbook and the historical content that was delivered to them. They could also use digital resources (articles, images, presentations) that were published in the LMS (the control group had the same access to resources as the experimental group except for the video lectures). If they had any further questions they could publish them in the forum. Following their home-based individual study, they attended classroom-based follow-up sessions which started with Q&A activities and then the teacher assessed their historical content memorization and HTS cultivation (same as in the experimental group). The assessment was followed by the teacher's feedback on students' answers and the delivery of the new historical content. At the end of each unit the assessments of the two main categories of learning goals took place. Students completed standardized tests which were common to the experimental group.

Evaluation Phase: During this phase, the completed educational intervention was evaluated. All data collected from the implementation phase were processed and the results were presented.

Results

Results regarding the use of classroom time (RQ1)

The purpose of the study was to examine the benefits of the FC model relating to the use of classroom time and the students' learning outcomes. This section summarizes the results that corresponded to the two research questions. Table 3 presents the distribution of classroom time in learning activities for the experimental and the control group in the first and second term.

Table 3. Distribution of classroom time in learning activities

	Lecture	Assessment of historical content memorization (verbal)	Q & A session/ Feedback	HTS cultivation activities	Standardized tests
1 st term					
Control group	220 min.	130 min.	145 min.	115 min.	90 min.
Experimental group	0 min.	95 min.	100 min.	440 min.	90 min.
2 nd term					
Control group	210 min.	110 min.	110 min.	115 min.	75 min.
Experimental group	0 min.	150 min.	120 min.	195 min.	75 min.

As Table 3 depicts, lecture was the main activity that the control group spent most classroom time in, whereas the experimental group spent no time in this activity. This is due to the implementation of the FC approach in the experimental group that transferred lectures entirely to the individual learning space. The control group also spent more time in (verbal) assessment of historical content memorization than the experimental group. The increase of classroom time spent by the experimental group in (verbal) assessment of historical content memorization in the second term is due to the revision of educational design, as a result of teacher's reflection to first term teaching and learning analysis. It was observed that the students of the experimental group who were less often verbally assessed on historical content memorization were not as well prepared as the students of the control group that were more often verbally assessed on historical content memorization. For that reason, in the second action research cycle, the teacher spent more time assessing that learning goal in the experimental group.

Moreover, in the first term, the control group spent more classroom time in Q&A sessions and feedback than the experimental group. This is due to the fact that the experimental group students participated more actively in the Moodle forum, asking most of their questions before coming to class. The students of the control group preferred to ask questions and, hence, received feedback at the time when the new content was delivered to them, during classroom sessions. In the second term though, the classroom time spent in Q&A sessions and feedback by the experimental group was increased, because the time spent in assessment of historical content memorization had also increased. Therefore, new questions came up especially from students that came poorly prepared to the class.

A significant difference of classroom time invested by the experimental group in HTS cultivation activities is observed in the first term. This is due to the implementation of the FC model that released classroom time and allowed students to engage in HTS cultivation activities. This finding provides evidence to positively answer our first Research Question, namely that the FC model allows classroom time to be more effectively used, because it is dedicated in student-centered engaging activities that facilitate the development of higher cognitive goals such as HTS cultivation. The decrease in classroom time spent by the experimental group in HTS cultivation activities in the second term is due to (a) the revision of the educational design as described above and (b) the fact that teaching hours became fewer in the experimental group compared to the first term because of school schedule (field trips etc.). Finally, the two study groups spent equal classroom time in standardized tests, which were common for both groups.

Results regarding to students' learning outcomes (RQ2)

As shown in Figure 1 and Table 4, during the first term, the experimental group had an increased improvement in students' score means. More specifically, the group reported (compared each time to the previous assessment) (a) in the 1st formative assessment a 4.38% increase and (b) in the 2nd formative assessment, a 1.7% increase. Correspondingly, the control group had an increase followed by a decrease in students' score means. The group

reported (a) in the 1st formative assessment a 3% increase, and (b) in the 2nd formative assessment, a 1.5% decrease.

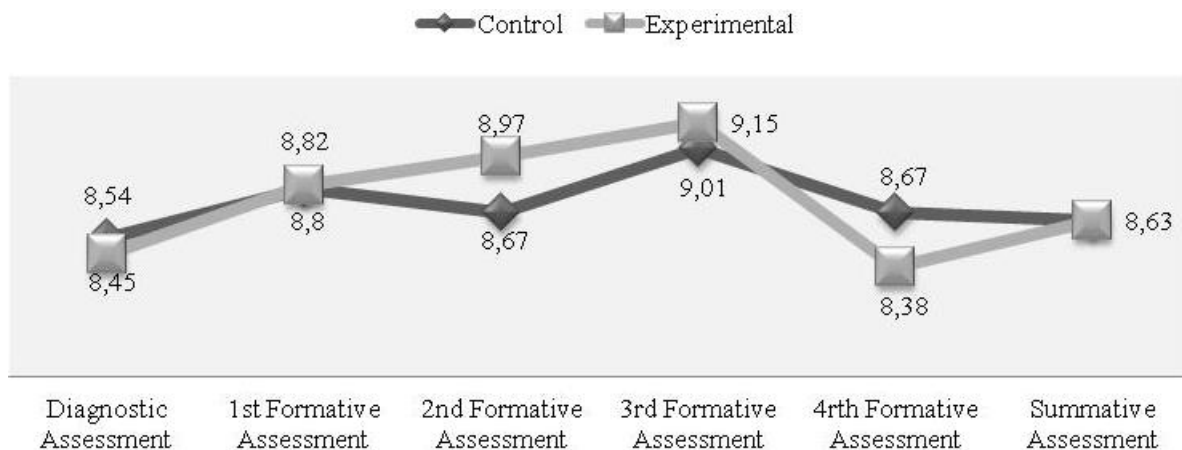


Figure 1. Assessment of historical content memorization (Study groups' achievement score means in standardized tests)

In the second term, the experimental group and the control group reported similar performances. The experimental group reported in the 4th formative assessment a decrease of 8.4%, compared to the previous assessment. This decline may be due to increased difficulty of the certain standardized test. In the summative assessment there was a 2.98% increase. The control group also had in the 4th formative assessment a 3.8% decrease and a 0.5% increase in the summative assessment.

Table 4. Historical content memorization: Mann-Whitney U test and Effect Size results

Assessment of historical content memorization	Mean (SD)		U	p	Effect Size (ES)
	Control group [N = 23]	Experimental group [N = 26]			
Diagnostic Assessment	8.54 (1.49)	8.45 (1.41)	277.5	.672	0.06
1 st Formative Assessment	8.80 (1.39)	8.82 (1.36)	293.5	.912	0.02
2 nd Formative Assessment	8.67 (1.54)	8.97 (0.93)	295.0	.936	0.25
3 rd Formative Assessment	9.01 (0.95)	9.15 (0.93)	272.0	.587	0.15
4 th Formative Assessment	8.67 (1.48)	8.38 (1.34)	251.0	.335	0.21
Summative Assessment	8.63 (1.59)	8.63 (1.53)	280.0	.702	0.003

As the Table 4 depicts, no statistical significance was evident between groups. However, it is evident from the data that the experimental group had in general the same or higher achievement scores in historical content memorization than the control group. There is only one different observation, the fourth formative assessment, which is considered an isolated event. The calculation of the effect size further attests to this conclusion, since it depicts a low effect in the differences in student historical content memorization achievement scores between the groups. Therefore, the data suggest that the FC model had a limited positive impact on enhancing students' content memorization capacity.

Table 5 depicts the results of the Mann-Whitney U test based on the assessment scores for each student group in terms of the three Historical Thinking skills categories. As the Table 5 depicts, regarding "Understanding the concept of time," no statistically significant difference exists in the diagnostic test, therefore the two groups had similar prior skill. However, the data for the remaining three tests signify that the experimental group showed a statistically significant improvement compared to the control group. The effect size supports this finding, attributing medium and large effects to the exploitation of the FC model. In a similar vein, the data for the "Understanding historical sources" category attest to (a) the initial similar skill level of the two groups (no statistical significance exists), and (b) to the statistical significant improvement of the experimental group in the subsequent three assessment tests. The large effect size results again corroborate these findings. Finally, in the "Historical analysis and interpretation" category, a similar pattern of results is observed, namely the two groups showed uniform skill levels during the diagnostic test, but subsequent statistically significant differences in favor of the experimental group (with corresponding medium to large effect sizes).

Overall, the analysis of data signifies that the implementation of the FC model had a significant positive impact on developing students' HTS, but not so regarding their historical content memorization capacity. This is an

interesting finding which provides evidence on the *added value* of the FC model, namely its potential for facilitating the development of students' historical thinking skills.

Table 5. HTS cultivation: Mann-Whitney U test and effect size results

Assessment of historical thinking skills	Mean (<i>SD</i>)		<i>U</i>	<i>p</i>	Effect Size (ES)
	Control group [<i>N</i> = 23]	Experimental group [<i>N</i> = 26]			
Understanding the concept of time					
Diagnostic Assessment	5.33 (1.74)	5.58 (1.91)	273.5	.580	0.13
1 st Formative Assessment	5.87 (2.08)	7.02 (1.73)	217.5	< .05	0.60
2 nd Formative Assessment	5.98 (1.54)	6.92 (0.93)	215.0	< .05	0.53
Summative Assessment	6.20 (1.59)	7.12 (1.53)	225.0	< .05	0.45
Understanding historical sources					
Diagnostic Assessment	5.22 (1.83)	5.58 (1.91)	266.0	.477	0.19
1 st Formative Assessment	5.54 (1.89)	6.73 (1.54)	203.0	< .05	0.70
2 nd Formative Assessment	5.65 (1.72)	6.54 (1.24)	217.0	< .05	0.60
Summative Assessment	5.76 (1.76)	6.83 (1.51)	210.0	< .05	0.65
Historical analysis and interpretation					
Diagnostic Assessment	5.00 (1.69)	5.38 (1.83)	263.0	.430	0.22
1 st Formative Assessment	5.98 (2.10)	6.92 (1.29)	208.0	< .05	0.55
2 nd Formative Assessment	6.09 (1.29)	7.02 (1.53)	204.5	< .05	0.55
Summative Assessment	6.30 (2.24)	7.12 (1.53)	205.5	< .05	0.43

Discussion and future work

Benefits of the flipped classroom model

Overall, we can conclude that the FC model contributed to a more effective use of classroom time since it released time from lecturing to more engaging student-centered learning activities. Consequently, in the experimental group, more classroom time was available for implementing learning activities that promote learning goals relevant to HTS.

Regarding students' learning outcomes, the data showed that the FC model had a low positive impact on developing students' traditional historical content memorization achievement scores between the two groups. This means that both study groups had similar learning outcomes in historical content memorization although the students of the control group spent significantly more classroom time with teacher's lectures compared to the students of the experimental group who spent no time for live lectures during classroom-based face-to-face sessions. This observation can provide promising research hypotheses for future investigations:

- The teacher's lectures in the control group may not have been engaging enough in order to increase students' learning outcomes in historical content memorization.
- The video lectures developed for the experimental group may have been as engaging as needed in order to promote learning outcomes in historical content memorization similar to the teacher's live lectures.

Regarding students' HTS, however, there was a clear positive impact of the FC model. More specifically, the experimental group showcased better learning outcomes than the students of the control group (in a statistically significant manner and with medium to large effect sizes). If we take under consideration that the HTS cultivation is a process that takes place mainly inside the classroom, it is reasonable to assume that there is a cause and effect relationship between the classroom time spent on HTS cultivation and the learning outcomes of HTS cultivation. In other words, the reason that the students of the experimental group had better learning outcomes than the students of the control group, may strongly be the fact that the FC model facilitated more classroom time to be invested on HTS cultivation activities.

Overall, potential future work could include additional longitudinal action research studies to provide additional evidence on the impact of the FC model on students' performance and learning experiences. Furthermore, these works could also adopt a more granulated level of analysis for studying and revealing how the FC impacts the learning patterns (e.g., learning material study patterns or engagement patterns) of individual students.

Challenges of the flipped classroom model

Aside from the benefits of the FC model that are described above, there are certain challenges in its use that need to be taken into consideration. First of all, it is most likely that students are not acquainted with the FC model and the different learning roles and activities it introduces. This can be even more relevant in school environments where the traditional teaching model is predominantly used. As a result, before implementing the FC model, the teacher should devote teaching time in order to “train” the students. Training students on the way they should study history or any other subject matter may prevent future obstacles in the learning process. Furthermore, there is a need for parents to be familiarized with the FC model especially in primary school where the family has an important role in student’s learning and can contribute to or obstruct the model’s efficiency.

The process of creating a course based on the FC model requires significant time invested on behalf of the teacher and also some level of technological competence. That is, the teacher needs to develop learning activities and resources for the classroom. In addition, it was observed that some students of the experimental group were not self-motivated enough and they did not thoroughly prepare themselves by watching and interacting with the video lectures nor did they study the material in general. The result was that they came to class unprepared. This led to less participation and engagement from those students who consequently fell behind. Therefore, it is important that the teacher is empowered with tools that facilitate him in monitoring individual students’ home-based preparation activities and collect data that can help him re-design activities based on differentiated instruction principles. Overall, if the FC model is to be widely adopted, tools for Teaching and Learning Analytics (Sergis & Sampson, In Press) are required to support teachers as reflective practitioners.

Acknowledgements

The work presented in this paper was conducted as part of the first author’s MSc Diploma Thesis at the MSc Programme “Digital Systems and Services (Track on e-Learning)” of the Department of Digital Systems, University of Piraeus, Greece and it has been partly supported by the State Scholarship Foundation (IKY) under the grant No 30275/2013B (IKY Fellowships of excellence for postgraduate studies in Greece – Siemens Programme).

References

- Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. Washington, DC: Internal Society for Technology in Education.
- Bergmann, J., & Sams, A. (2014). Flipping for mastery. *Educational Leadership*, 71(4), 24-29.
- Bergmann, J., & Sams, A. (2015). *Flipped learning for social studies instruction*. Washington, DC: Internal Society for Technology in Education.
- Bishop, J., & Verleger, M. A. (2013, June). *The Flipped classroom: A Survey of the research*. Paper presented at 2013 ASEE Annual Conference, Atlanta, Georgia.
- Chen, Y., Wang, Y., Kinshuk, & Chen, N. S. (2014). Is FLIP enough? Or should we use the FLIPPED model instead? *Computers & Education*, 79, 16-27.
- Cohen, J. (1994). The Earth is round ($p < .05$). *American Psychologist*, 49(12), 997-1003.
- Common Core State Standards Initiative. (2010). *Common core state standards for English language arts & literacy in history/social studies, science and technical subjects*. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers.
- DeLozier, S. J., & Rhodes, M. G. (2016). Flipped classrooms: A Review of key ideas and recommendations for practice. *Educational Psychology Review*, 1-11.
- Dick, W., Carey, L., & Carey, J. O. (2001). *The Systematic design of instruction*. New York, NY: Longman.
- Epstein, T. (2012). Preparing history teachers to develop young people’s historical thinking. *Perspectives on History*, 50(5), 36.
- Fielding, J. (2005). Engaging students in learning history. *Canadian Social Studies*, 39(2). Retrieved from <http://files.eric.ed.gov/fulltext/EJ1073994.pdf>

- Fulton, K. P. (2014). *Time for learning: Top 10 reasons why flipping the classroom can change education*. Singapore: Corwin Press.
- Gaughan, J. E. (2014). The Flipped classroom in world history. *The History Teacher*, 47(2), 221-244.
- Green, T., Ponder, J., & Donovan, L. (2013). Educational technology in social studies education. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 573-582). New York, NY: Springer New York.
- Haydn, T., Stephen, A., Arthur, J., & Hunt, M. (2014). *Learning to teach history in the secondary school: A Companion to school experience*. London, UK: Routledge.
- Hwang, G. J., Lai, C. L., & Wang, S. Y. (2015). Seamless flipped learning- a mobile technology enhanced flipped classroom with effective learning strategies. *Journal of Computers in Education*, 2(4), 449-473.
- Keengwe J., & Onchwari, G. (2015). *Handbook of research on active learning and the flipped classroom model in the digital age*. Hershey, PA: IGI Global.
- Kotlik, R. H. (2014). The Flipped model in an advanced placement United States history course. In J. Keengwe, G. Onchwari, & J. Oigara, (Eds.), *Promoting Active Learning through the Flipped Classroom Model* (pp. 208-225). Hershey, PA: Information Science Reference.
- Koumi, J. (2006). *Designing video and multimedia in open and flexible learning*. London, UK: Routledge.
- Krathwohl, D. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory into Practice*, 41(4), 212-218.
- Lee, J. K., & Friedman, A. M. (2009). *Research on technology in social studies education*. London, UK: Routledge.
- Lewin, K. (1948). *Resolving social conflicts: Selected papers on group dynamics*. New York, NY: Harper & Brothers.
- Lyons, J. F. (2008). *Teaching history online*. London, UK: Routledge.
- McKernan, J. (1988). The Countenance of curriculum action research: Traditional, collaborative and critical-emancipatory conceptions. *Journal of Curriculum and Supervision*, 34(3), 173-200.
- McKernan, J. (1991). *Curriculum action research. A Handbook of methods and resources for the reflective practitioner*. London, UK: Kogan Page.
- Myers, C. B., Adler, S., Brandhorst, A., Dougan, A. M., Dumas, W., Huffman, L., ... & Helmkamp, C. J. (2006). *National Standards for Social Studies Teachers. National Council for the Social Studies*. Retrieved from <http://www.htsb.org/wp-content/uploads/2014/01/2004-NCSS-Soc-Standards.pdf>
- Scheuerell, S. K. (2015). *Technology in the middle and secondary social studies classroom*. London UK: Routledge.
- Sergis, S., & Sampson, D. (In Press). Teaching and Learning Analytics to support Teacher Inquiry: A Systematic Literature Review. In P. Peña-Ayala (Ed.), *Learning analytics: Fundamentals, applications, and trends: A view of the current state of the art*, Heidelberg: Springer
- Sergis, S., Vlachopoulos, P., Sampson, D. G., & Pelliccione, L. (2017). Implementing teaching model templates for supporting flipped classroom-enhanced STEM education in Moodle. In A. Marcus-Quinn & T. Hourigan (Eds.), *Handbook for Digital Learning in K-12 Schools* (pp. 191-215). New York, NY: Springer New York.
- Townsend, R. B. (2010). Assimilation of new media into history teaching: Some snapshots from the Edge. *Perspectives on History*, 48(9), 24-26.
- UMBC Center for History Education. (2013). *ARCH historical thinking skills rubric – Elementary*. Retrieved from <http://goo.gl/JIGmgJ>