Article

Evidence-Informed Teaching: Investigating Whether Evidence from 'Flipping the Classroom' Research Improves Students' Motivation for Mathematics

Adri Dierdorp

Freudenthal Institute for Mathematics and Science Education, Utrecht University,
3508 AD Utrecht, The Netherlands; a.dierdorp@uu.nl; Tel.: +31-648819775

Abstract: This study from 2019 investigates whether the impact on a STEM teacher's evidence-informed teaching approach using the evidence of flipping the classroom research improves students' (13–14 years old) motivation in a Dutch setting and if this approach allows students to perform better. We report this approach in line with the cycle of expansive learning of Engeström. We asked: "To what extent can evidence based on the flipping the classroom approach improve the motivation and results of grade 8 preuniversity track students doing mathematics?", followed by the subquestions: "To what extent does education by the FtCA increase student motivation?" and "To what extent does education by the FtCA ensure better test results for students?". A questionnaire is used to investigate to what extent the motivation of students increased, and a teacher is interviewed about his experiences with the "flipping the classroom" model. To test whether the results have improved, a pre- and post-test is taken and analyzed. A significant increase in both intrinsic and extrinsic motivation has been found, and students gained a stronger sense of autonomy, competence, and belonging. The test results improved, but the difference is not statistically significant. However, despite the disappointing test results, the teacher was very positive about the new way of working.

Keywords: evidence informed teaching; flipping the classroom; motivation; expansive learning

1. Introduction

Teaching at a high school based on "evidence-informed teaching" (EIT) is increasingly popular with a desired impact on the teaching [1–4]. Evidence-informed teaching is about applying robust external evidence, comparing past experience with present experience challenged by external evidence, and assessing the evidence in order to improve your own teaching practice. EIT can be implemented on the macro or micro level as a real educational change. As a reason for educational change, Burner [5] mentions, among others, "developments in research into teaching and learning approaches". However, for a single teacher, the reason can be the necessity to improve the students' motivation.

Student motivation to learn mathematics can often be low, resulting in poor performance. Stakeholders in mathematics education have begun to experiment teaching differently from traditional education in order to address this problem [6]. The infusion of technology can especially be very useful in classrooms teaching secondary level mathematics [7]. If teaching deviates from classroom-based, front-based education, this is quickly referred to as an alternative form of education, which has a positive or negative connotation depending on the messenger. The traditional form of education is traditional because it seemed to work well for a long time. One disadvantage of this frontal way, however, could be that in conceptual development, the student often does not have to think for himself and must only reproduce well. Often, the concepts do not take root, turning out not to have been learned in the long term [8]. Alternative approaches are sought.

The "flipped classroom approach" (FtCA) is an approach in which students use the Internet, videos, or author audio-visual recordings to be introduced to pre-recorded

concepts outside of the traditional classroom [9–11] as part of their homework. After they watch the digital information (often at home), they discuss the content in the next class with their peers and teacher and possibly clear up any misconceptions. In class they are expected to discuss, explain, and extend the concepts they learned [12]. Thus, homework in a traditional setting becomes what the students have to do in class, and vice versa. In fact, the classroom is not "flipped", but the way of learning in and outside the classroom is.

The FtCA is becoming increasingly popular when teaching STEM [13–15]. In this study, we explicitly mention mathematics as being part of STEM education. Educational strategies following the FtCA seem to entice greater levels of motivation in students [16]. However, few studies have been published with reference to the FtCA [17], and worse is the amount of literature available on the FtCA which is exclusively devoted to mathematics in high schools [12].

The aim of this presented study is to investigate the impact on a young teacher's EIT approach whether using the FtCA improves the motivation of students in a Dutch setting (eighth grade preuniversity level) during mathematics education, and if this approach allows the students to perform better. We define motivation as the desire and willingness to actively participate in the learning process offered by the teacher. In this study we follow a STEM teachers' EIT practice in line with the cycle of expansive learning of Engeström [18].

1.1. Theoretical Background

In this section we give a short description of the expansive learning cycle and some theoretical background about the FtCA and motivation.

1.1.1. Expansive Learning

Implementing new didactics means introducing new knowledge and skills, instead of improving existing ones. Engeström [18] speaks of "expansive learning" and has developed a cycle to support the research, teaching, and implementation of expansive learning (see Figure 1).
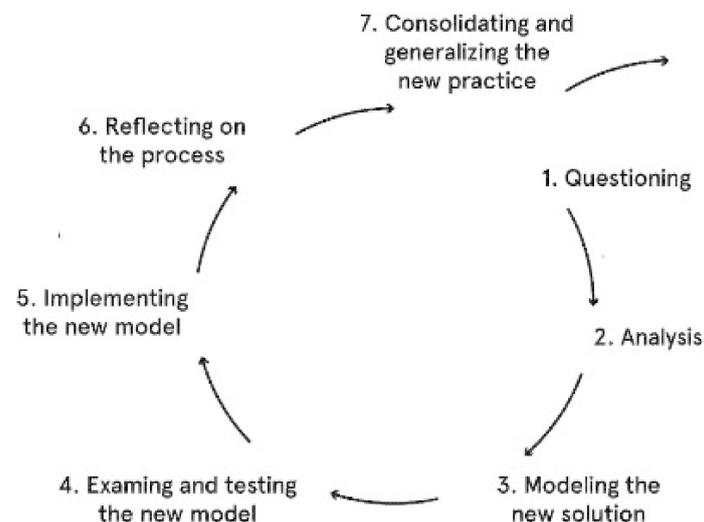

Figure 1. Expansive learning cycle. After "Perspectives on activity theory" by Engeström, Y., 1999, p. 399, Cambridge: Cambridge University Press).

"Expansive learning" consists of seven basic actions and fits very well with the key elements (ke1, ke2, ke3, ke4) of a FtCA. The first action, asking questions, occurs when the student perceives contradictions and questions, criticizes, or rejects certain aspects of accepted practice and wisdom (ke1, ke2, ke3). Analyzing occurs when he or she asks questions and investigates to find causes or explanatory principles (ke2, ke3). Modelling refers to the action by which the researcher or student constructs an explicit model of the idea that explains the problematic situation and provides a solution (ke3). Examining

the model means trying the model to understand its dynamics, potential, and limitations (ke3, ke4). Implementation of the model refers to practical applications, enrichments, and conceptual extensions (ke3, ke4). The actions that follow are to reflect on the process and consolidate its results (ke2, ke3, ke4).

1.1.2. Flipping the Classroom

In this study, we see lecturing as the traditional model of classroom instruction; the teacher provides information during class and responds to questions, while when students need help or feedback, they must defer directly to the teacher. Lecturing is often focused on an explanation of content using a lecture style [19]. Even in class discussions, the teacher controls the flow of the conversation and is the centre of the lecturing. Mazur [8] writes the following (p. 51) in his farewell article: "I once heard somebody describe the lecture method as a process whereby the lecture notes of the instructor get transferred to the notebooks of the students without passing through the brains of either. That is essentially what is happening in classrooms around the globe." The FtCA is one response to such observations. More and more reports [13–15] about the FtCA are appearing on different levels. In Spain, a training program for future Primary Education teachers showed a very positive impact on motivation, and the FtCA was applied to the curriculum [20]. As mentioned earlier, a student goes through the learning process via the FtCA exactly the other way around from the traditional setting of lecturing. That which normally happens during a lesson and that which is normally seen as homework changes. Referring to Bloom's taxonomy [21], this means that students are gaining knowledge and comprehension (the lower levels of cognitive work) outside of class, and focusing on application, analysis, synthesis, and/or evaluation (the higher forms of cognitive work) in class, together with their peers and teacher [22]. Having mentioned this, we can conclude that the FtCA is not just about doing homework. Brame gives us a list of key elements for a successful FtCA:

ke1. Provide an opportunity for students to gain first exposure prior to class.
ke2. Provide an incentive for students to prepare for class.
ke3. Provide a mechanism to assess student understanding.
ke4. Provide in-class activities that focus on higher level cognitive activities.

Provide an Opportunity for Students to Gain First Exposure Prior to Class

The way to provide an opportunity can differ in each lesson. For example, students can study theory from books or articles themselves, but it is also possible to view videos or podcasts selected or made by the teacher. It is important that it is possible for all students to prepare for the lesson.

Provide an Incentive for Students to Prepare for Class

As the student prepares himself out of sight of the teacher, and (social) control is thus lost, it is important that the student receives something in return when he has prepared himself. Frequently used resources are points for doing homework. However, participation in (online) "discussion boards", online quizzes, worksheets, or short writing assignments could also provide an incentive for students.

Provide a Mechanism to Assess Student Understanding

Bransford, Brown, and Cocking [23] wrote: "To develop competence in an area of inquiry, students must: (a) have a deep foundation of factual knowledge, (b) understand facts and ideas in the context of a conceptual framework, and (c) organize knowledge in ways that facilitate retrieval and application" (p. 16). This citation understates the need to provide a mechanism to assess student understanding. During a lesson, often at the beginning, it will be necessary to check whether the pupils indeed have the skills and/or knowledge that they should have due to the preparation for the lesson. Starting with a classroom question that students must answer individually or in groups is the most common way to do this. Digital means facilitate this but are not necessary.

Provide in-Class Activities That Focus on Higher Level Cognitive Activities

The preparation for the lessons is mainly focused on the lower cognitive levels, such as memorization and reproduction. In this way, during the lessons, there can be room for activities aimed at the higher cognitive levels, depending on the need for this at that point in the curriculum. Bransford, Brown, and Cocking [23] stress that "A 'metacognitive' approach to instruction can help students learn to take control of their own learning by defining learning goals and monitoring their progress in achieving them" (p. 18). The immediate feedback that occurs following the FtCA could also help students recognize and think about their own growing understanding.

In summary, teaching by the FtCA means that students first absorb new theory, or what there is to learn, outside of class, often by reading pieces or with the help of tutorial videos, after which they actually absorb the material in class, through problems to resolve or to discuss [22]. This means that according to Bloom's taxonomy [20,24], students acquire lower levels of cognition outside of class and acquire the higher levels in class [25]. In class, students could learn to take control of their own learning by monitoring their progress or by defining learning goals [23]. By formulating learning goals and monitoring the progress of students in achieving them, this metacognitive approach can be further enhanced. Although the latter is not the primary goal of FtCA, the initial assignments and any collaboration assignments can contribute to this.

Saunders [12] used the theories of Vygotsky and Bandura to explain the relationship between the FtCA, mathematics achievement, and students' critical thinking skills in secondary mathematics classrooms. These theories suggest that when students learn through social interactions, in groups, or in collaboration with the teacher (facilitation), they retain the self-discovered knowledge and information apprehended with teacher assistance and actually enjoy learning mathematics [26].

The FtCA seems to be successful. However, there are few reports on research to improve the mathematical learning performance of middle school students [27]. Fortunately, more and more reports are appearing. Wei [28] claimed in a study with 88 sixth-grade students that the FtCA significantly improves the students' mathematical learning performance. Saunders [12], however, found no significant difference between the focus group and control group.

1.2. Motivation

Saunders [12] used the theories of Vygotsky and Bandura to explain the relationship between the FtCA, mathematics achievement, and students' critical thinking skills in secondary mathematics classrooms. These theories suggest that when students learn in groups or through social interactions, the self-discovered knowledge and information apprehended with teacher assistance last longer and the students are more motivated to do mathematics [26].

A change in student motivation could be a positive effect of a new form of didactics. Ryan and Deci [29] distinguish intrinsic and extrinsic motivation within the self-determination theory, which is based on the relationship between the will to do something (determination) and our experience of three basic psychological needs: 1: autonomy [30,31], 2: competence [32,33], and 3: connectedness [34,35]. These needs influence the locus of control: the extent to which someone feels that he can make his own choices and achieve a goal with the resulting behaviour. A more internal locus of control leads to a higher degree of self-determination and thus intrinsic motivation for a particular task [29]. When less autonomy, sense of competence, and connectedness are felt, a more extrinsic form of motivation will arise to perform a task. Intrinsic motivation is the strongest form of motivation that influences behaviour [36,37]. When external rewards and punishments are involved and the behaviour can be described as more compliant, it is referred to as extrinsic motivation, with an external locus of control. Even when it involves internal rewards, such as ego, self-control, or a sense of personal importance, there may be extrinsic motivation. We speak of extrinsic motivation (the reward is still outside the task itself) with an internal

locus of control; it comes from the individual and is not imposed. When there is an internal locus of control, behaviour is more strongly influenced than when there is a more external locus of control [38]. If a task does not provide a feeling of autonomy, competence, and belonging, then there will be no question of any of the abovementioned behaviours, and it is referred to as amotivation.

Subjective desirable results in relation to the FtCA, such as satisfaction and positivity, appear in more studies. Saunders [12], however, explained that students who really took advantage of the flipped classroom format were the students who were already intrinsically motivated before the study. Additionally, Missildine, Fountain, Summers, and Gosselin [39] found that students were less satisfied with the setting in which the FtCA was applied than with the traditional setup. One reason for this also confirms that it is necessary to actually be able to go through the teaching material before participating in the classroom lesson. For many students there was no reliable internet connection during the study. This connection was essential to be able to study the teaching materials during the lesson.

1.3. Research Questions

In the theoretical background we see that the FtCA seems hopeful, but this is not always the case. To investigate the effects of the FtCA in a Dutch grade 8 preuniversity track setting we ask the next research question:

To what extent can evidence based on the "flipping the classroom" approach improve the motivation and results of grade 8 preuniversity track students doing mathematics?

As seen in the theoretical background, student outcomes are influenced by motivation. To improve performance as much as possible, there must be a higher intrinsic motivation or extrinsic motivation with an internal locus of control. It is important to apply didactics in such a way that students feel a high degree of autonomy, competence, and connectedness. Therefore, in order to answer the research question, an answer is sought to the following subquestions:

1. To what extent does education by the FtCA increase student motivation?
2. To what extent does education by the FtCA ensure better test results for students?

2. Materials and Methods

2.1. The Participants

In this pilot study from 2019, we follow Teacher A, an experienced teacher who completed teacher training to obtain his bachelor's degree. For the experiment, two of his 8th grade (13–14 years old) Dutch preuniversity track classes were chosen. Due to persistent disappointing results in these groups, it was decided to try a different way of organizing education. Both groups have the same teacher. The two groups are randomly assigned as the "Focus group" on which the experiment will take place and the "Control group." The Focus group is taught according to the FtCA and the Control group is taught in the traditional way. The number of students and boys and girls per group can be read in Table 1.

Table 1. Number of students (and the distribution boys, girls) of Test group and Control group.

Group	Number of Students	Number of Boys	Number of Girls
Focus group	29	14	15
Control group	30	15	15
Total	59	29	30

Furthermore, Teacher B, a colleague of Teacher A, was involved for triangulation reasons. During the questioning phase (ke1) of the Engeström cycle, Teacher A struggled with low motivated students and criticized the traditional teaching methods at his school. During ke2 (analysis), he analysed the report of the education inspection of the government,

scientific literature, and comments of colleagues. In the modelling phase (ke3), he designed a lesson series based on the FtCA.

2.2. The Design

To design a series of lessons based on the FtCA, we have drawn up a number of design requirements based on the literature:

1.   The lessons must be suitable for the discipline of grade 8 mathematics of the Dutch preuniversity track.
2.   The students all have access to the material to prepare for a lesson.
3.   Students have access to a virtual learning environment (VLE).
4.   Each lesson should have a short baseline to determine the extent to which students master the concepts.
5.   The teaching materials should give the students the feeling that they have a better command of the concepts and skills.
6.   Lessons should be suitable for students to work together.
7.   The didactics must provide an incentive for students to actively participate in the lessons.

To provide students an opportunity to gain first exposure prior to class, the students of the Focus group have access to the materials to prepare for a lesson. This material consists of a textbook and additional material in the form of explanatory videos and additional exercises that are available through the VLE. The VLE is where all supporting material is housed and collected on one chapter page (CP) per chapter. Each CP includes the theory, instructional videos, exercises, and an assessment question. In addition, students have their standard notebook, textbook, and an iPad with a digital book with answers at their disposal. In addition to the CPs, a forum has been created where students can ask questions. The CPs are always accessible to the students, at home and at school. The layout of the pages consists largely of videos in which the material is explained. The videos on a CP mostly accompany the textbook. However, the theory and explanations are slightly more extensive. Sometimes the teacher must make his own videos when none are available.

The Focus group is given homework consisting of three parts: assessing assignments from the previous lesson, reading theory, and watching videos. Furthermore, they receive an assessment of one or two assignments that are not too difficult in order to check whether they understood the theoretical basis. This check forms the baseline measurement at the beginning of the lesson. For questions outside of the lessons, a forum has been set up within the VLE to which all students of the grade have access. A lesson then starts with the next assessment: a task to check which students understood the concepts and which did not. Students individually have a maximum of 5 min to come to an answer. Then, the assignment is assessed in class. Depending on the length of the assignment, several students are appointed to explain the steps they have taken and why. Control of the written work takes place while the students are busy with the assessment. An example of an assessment question is, for example, a question in which students must link the correct formula to the correct graph (see Figure 2).
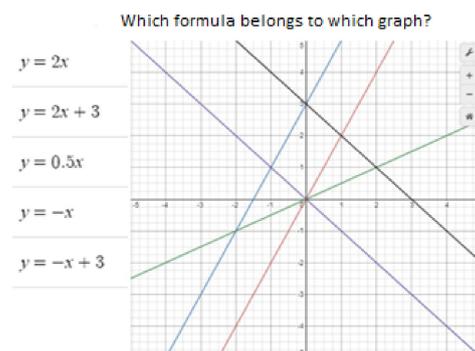

Figure 2. Example of an assessment question (translated).

Then, the class is split into two parts. The students who have done their homework, have no questions about it, and have completed the assessment correctly may choose whether they prefer to continue working in the classroom or on the so-called study balcony. The study balcony is a place above the classroom for working in silence. Students who were unable to complete the assessment and/or had any questions about the materials or the work they did, stay in the classroom and receive extra explanation in a smaller group.

A teaching assistant is present on the study balcony during 2 of the 3 h of class to keep order, but practice shows that this is hardly necessary. It turns out that students enjoy sitting and working together in small groups. It seldom happens that a student does not finish his work during the designated lesson time. A tutor, a senior student paid by the parents association, is available for the remainder of the class when no teaching assistant is present. Students on the balcony who do have questions about exercises or theory are free to walk downstairs and ask the teacher questions when he is available. Five minutes before the end of the lesson, all students return to the classroom, where they end together by completing a group assignment to reflect upon the lesson material.

In order to provide students the opportunity to view more examples (and practice themselves), the teacher creates teacher videos that cover more examples than in the already available videos on the CP. In order to respond as clearly, but also as quickly as possible, to what students need, a Wacom Bamboo Slate (digital notebook) is used in the so-called live mode, linked to the iPads. Everything the teacher writes on the Slate the students see on the screens of their iPads. When the teacher discusses the assignment, he makes screen recordings which are later cut or deleted, if necessary, before uploading to a YouTube channel. Additionally, videos are "embedded" on the abovementioned CP. As videos can be added relatively quickly in this way, new videos are created when required.

Watching videos requires the student to become accustomed to this form of media. There is a great temptation to consume passively. To provide students with an incentive to actively participate in the lessons, the teacher placed the following text on each CP in the VLE:

"The videos here are meant to help you throughout the chapter. They are not a substitute for the theory from the chapter, but you can learn with them just fine if you work actively. [followed by an instruction]"

Learning such a behaviour takes time, but months of experience show that students started watching videos in this way by themselves. The HPs are ready at the beginning of a period for the student, who can find his or her own way through the available study guide. During the period, videos will be added if there is demand.

The Control group is taught in a more traditional way. The same materials as for the Focus group are available to them on the VLE, but in this case, the lesson starts with a lecture to explain the concepts of the theory. After an assessment by doing part of an exercise together, the students are given the time to independently work through their exercises. During the last few minutes of the lesson, a part of an exercise is done together to reflect on what has been learned.

In Teacher A's phases ke4 and ke5, he implemented his design and tested it thoroughly. In the next section, we describe his data collection and the method of analysing it.

2.3. Data Collection and Analysis

In the context of triangulation, various sources have been used in the collection of data. We used student surveys, student scores, and teacher interviews. For this study, the teacher designed a series of 6 lessons, and we utilized a static-group comparison nonequivalent group design. The data were analysed using SPSS software.

2.3.1. Research Question 1: To What Extent Does Education by the FtCA Increase Student Motivation?

In order to be able to answer research question 1, a questionnaire is taken from the students of the Focus group and a teacher is interviewed. Table 2 provides justification of the survey questions, referring to the design requirements previously set.

Table 2. Justification of the questionnaire items referring to the design requirements (translated).

Nr	Question	Requirement						
		1	2	3	4	5	6	7
1	I enjoy math more now.	x						x
2	I find math more interesting now.	x						x
3	I am now more satisfied with my performance in math.	x				x		x
4	I now have more control over my performance.	x				x		
5	I have the option of using the software "It's Learning" for learning properly.	x	x	x				
6	I am now better prepared for the lessons.	x	x					
7	The questions at the beginning of the lesson help me to gain more insight into my understanding of the material.	x			x			
8	I now expect to perform better on my tests.	x				x		
9	I like math more now.	x						x
10	I am now more interested in math.	x						x
11	I now prepare better for the lessons.	x	x					
12	I am now more satisfied with how math is going.	x				x		
13	I now feel more responsible for my own achievements.	x				x		
14	I now feel more like I can do it.	x				x		
15	I now work together more in class.	x					x	
16	In which areas do you think this is an improvement?	x	x	x	x	x	x	x
17	In which areas do you think this is worse?	x	x	x	x	x	x	x
18	Would you like to add something else?	x	x	x	x	x	x	x

The response options that ranged from "completely disagree" to "completely agree" on a 5-point Likert scale are recoded from 1 to 5 for statistical analysis. Questions 16, 17, and 18 are open questions, the answers of which are classified in the categories of autonomy, competence, connectedness, intrinsic motivation, extrinsic motivation, amotivation, and other categories depending on the type of answer. The answers in these categories, together with the scores on the closed questions from the questionnaire, are meant to determine the degree and form of the motivation of students. To check the internal consistency of the requested constructs, Cronbach's alpha [40] is calculated via SPSS. Based on this, the relevant questions measuring intrinsic and extrinsic motivation are also combined into a new variable, which measures intrinsic and extrinsic motivation. We score the answers to the open questions with a + or a -: a positive when more comments are made or negative when less comments are made referring to the categories of intrinsic motivation (IM), extrinsic motivation (EM), autonomy (A), competence (C), connectedness (V), and possibility to prepare (M).

To assess whether the intervention is successful, a t-test (SPSS) analyses whether there is a significant difference between before and after the intervention.

Furthermore, a colleague of Teacher A, teaching a parallel class and involved in this study, is interviewed using the next topic list:

What improvements do you see by "flipping the classroom" with students in the area of:

- Performance
- Working position
- Fun
- Trust

What deterioration do you see by "flipping the classroom" in students in the area of:

- Performance
- Working position
- Fun
- Trust
- How could we solve that?

Where is room for further improvement?

- Within the lessons
- In the organization
- In other areas

Would you like to add something else?

As no short, closed answers are sought, an unstructured interview is used. To obtain as much information as possible from the interview, the topic list is used as a guideline. The interview opened with the reason for the interview. The researcher does not provide their own opinions (to avoid socially desirable answers). Answers are briefly written down during the interview according to Wiertzema and Jansen [41]. The summary of the interview is submitted to the person concerned for approval. The teacher's answers are categorized into positive and negative experiences with regard to intrinsic motivation (IM), extrinsic motivation (EM), autonomy (A), competence (C), connectedness (V), and ability to prepare (M).

2.3.2. Research Question 2: To What Extent Does Education by the FtCA Ensure Better Test Results for Students?

In order to answer the question "To what extent does education by FtCA ensure better test results for students?", an equal test is administered in the Focus group and in the Control group before and after the introduction of the FtCA. The scores of the students of both groups are entered in SPSS. The Shapiro–Wilk test determines whether the data is normally distributed [42]. To assess whether there is a significant difference in the mean scores between the control and experimental groups, a t-test is performed if the data is normally distributed or a Mann–Whitney U test if the distribution is abnormal.

In Teacher A's phase ke6 and ke7 of the Engeström cycle, he reflected on his design presented in the next section. Due to the pilot status of this study and because we only followed teacher A, we have chosen to combine the quantitative perspective for students with a qualitative perspective for teacher A as a mixed method study [42,43].

3. Results

3.1. Research Question 1

To make a statement about the change in intrinsic and extrinsic motivation, the answers to the questions of the questionnaire (Focus group) about intrinsic and extrinsic motivation are combined into one variable. If the new variable is a reliable measure of the form of motivation to be measured, a single variable provides more overview.

All questions about intrinsic motivation can be reliably combined into one variable ($\alpha = 0.827$). Omitting one of the questions does not increase the reliability of the measurement (maximum $\alpha = 0.835$). The same does not apply to all questions about extrinsic motivation ($\alpha = 0.403$). The low value of Cronbach's alpha makes aggregating all five questions an unreliable measurement. By removing questions, the mean of questions 2 and 5 does produce a reliable measurement ($\alpha = 0.723$). The questions are also reported separately.

A *t*-test is used to test whether the deviation from the mean (3 as neutral) is statistically significant. The findings can be found in Table 3. The intrinsic and extrinsic motivation show an increase of 0.36 and 0.88, respectively. Both differences are very significant: $t(24) = 3.57$, $p = 0.002$, and $t(24) = 8.37$, $p = 0.000$, respectively. Autonomy also shows a significant increase: a deviation of 0.76 from neutral with $t(24) = 8.72$, $p = 0.000$. Competency shows an increase of 0.60, which is also significant: $t(24) = 5.20$, $p = 0.000$. As a final part of

the degree of motivation, connectedness also shows an increase (0.32) that is significant: t(24) = 2.32, *p* = 0.029. With an average of 3.72, students also indicate that they are generally satisfied with the ability to prepare. The value is 0.72 above neutral, showing a significant difference: t(24) = 3.85, *p* = 0.001.

Table 3. Average per measured construct (Focus group) and deviation from neutral (=3).

Construct	M	SD	MD	*p*
Intrinsic motivation	3.36	0.50	+0.36	0.002
Extrinsic motivation	3.88	0.53	+0.88	0.000
Autonomy	3.76	0.44	+0.76	0.000
Competence	3.60	0.58	+0.60	0.000
Connectedness	3.32	0.69	+0.32	0.029
Possibility to prepare	3.72	0.94	+0.72	0.001

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, and the experimental conclusions that can be drawn.

The answers to the open questions are coded according to intrinsic motivation (IM), extrinsic motivation (EM), autonomy (A), competence (C), connectedness (V), and ability to prepare (M). In Table 4 we present the numbers of remarks made in the open questions.

Table 4. Number of comments regarding intrinsic motivation (IM), extrinsic motivation (EM), autonomy (A), competence (C), connectedness (V), and ability to prepare (M).

Code	+	−	Sub
IM	0	2	2
EM	0	1	1
A	18	0	18
C	11	2	13
V	0	0	0
M	0	0	0
total	29	5	34

Most answers relate to autonomy and competence. Table 4 shows that 18 pupils give a comment that indicates an increased sense of autonomy, while no comments indicate a decrease in this. Furthermore, 11 comments indicate a greater sense of competence, while four indicate a decreased sense of competence. Thus, in general, students seem to cite autonomy as the main advancement that the FtCA offers them.

To support the findings, the involved colleague of the teacher is interviewed. Table 5 shows the number of comments during the interview related to motivation. From this, it follows that, as far as the interviewee is concerned, an increase in autonomy, competence, and connectedness are the main points of improvement of the FtCA, but that the sense of competence also decreases in a group of students.

Table 5. Number of the remarks of the interviewee coded as intrinsic motivation (IM), extrinsic motivation (EM), autonomy (A), competence (C), connectedness (V), and ability to prepare (M).

Code	+	−	Sub
IM	1	0	1
EM	2	0	2
A	5	1	6
C	3	3	6
V	3	0	3
M	1	0	1
Total	15	4	19

The interview also shows that the interviewed colleague is predominantly positive about the use of the FtCA as a teaching method. The colleague in question has recently also switched to this method of teaching and is confident that in the long term it is an approach to teaching that will ensure that students perform better. A greater sense of autonomy, more cooperation, and, for some of the students, a greater sense of competence are also evident and can be cited as reasons for the expected growth in performance.

3.2. Research Question 2

The results of the tests before and after the intervention in the Focus group and the Control group show no significant deviation from a normal distribution according the Shapiro–Wilk test (Table 6).

Table 6. Shapiro–Wilk test before and after the intervention.

Group	Before Intervention		After Intervention	
	W	p	W	p
Focus Group ($n = 28$)	0.96	0.30	0.98	0.78
Control Group ($n = 29$)	0.96	0.38	0.97	0.67

Based on the results of the Shapiro–Wilk test, a paired t-test is used to compare the scores. The Focus group appeared to score higher on average on the second test than on the first test (see Table 7). However, this 0.45 improvement is not significant: t (27) = 1.25, $p = 0.224$. The Control group shows a lower average on the second test than on the first. However, this deterioration of 0.36 is not statistically significant: t (28) = −0.92, $p = 0.364$.

Table 7. Mean of test results before and after the intervention.

Group	# Students	Mean before		Mean after	
		M	SD	M	SD
Focus group	28	6.2	1.92	6.72	1.37
Control group	29	6.9	1.92	6.54	1.42

Furthermore, the difference between boys and girls was examined by splitting the data into a group of boys (13) and a group of girls (15). The results of the tests in the Focus group and the Control group show no significant deviation from a normal distribution according to the Shapiro–Wilk test (See Table 8). Based on the results of the Shapiro–Wilk test, a paired *t*-test is used comparing the scores of the Focus group.

Table 8. Shapiro–Wilk test before and after the intervention (boys vs. girls of Focus group).

Group	Before Intervention		After Intervention	
	W	p	W	p
Boys ($n = 13$)	0.95	0.61	0.92	0.26
Girls ($n = 15$)	0.91	0.16	0.98	0.99

The group of 15 girls scored higher (see also Table 9) after the intervention than before, but the difference of 0.20 is not statistically significant: t (14) = −0.52, $p = 0.615$. The group of 13 boys showed a stronger increase in the mean between before and after the intervention. However, the difference of 0.73 is also not statistically significant: t (12) = −1.15, $p = 0$.

Table 9. Mean results of boys and girls from the Focus group before and after intervention.

Group	# Students	Mean before		Mean after	
		M	SD	M	SD
Boys	13	5.35	1.92	6.08	1.37
Girls	15	7.07	1.54	7.27	1.25

In the last phase (ke7) of Teacher A's development process, we asked him to reflect on his project by asking five questions:

1. Does scientific literature play a major role in your daily teaching?

Teacher A: "When I did research on the effects of flipping the classroom on the performance and motivation of high school students, scientific research helped me shape my daily practice. The research design was met with great enthusiasm by the students and the initial results gave us the confidence to continue teaching this way. Therefore, even though not every decision I make is driven directly by scientific literature, it has played a major role in the way I teach math."

2. Why do you think it is important to align your teaching to scientific literature?

Teacher A: "It is the only way to make sure you are reaching the goals you have set for yourself. Creativity and intuition alone may come a long way in creating great lessons, but without proper research there is no way of knowing whether it's the best practice. Additionally, besides its possibility of making a difference between good teaching and great teaching, scientific literature is there to make it easier for teachers. To inspire you to change your daily teaching, to alter little details of your lessons, and simply copy everything that's proven to work. When you choose to align your teaching to scientific literature, you create better lessons with, sometimes, less work."

3. For whom is scientific literature important in school?

Teacher A: "Teachers should be the first to consider incorporating the outcomes of scientific literature in their teaching if we talk about research on specific teaching-related subjects for reasons described above. However, scientific research in general is not a constraint to teaching of course, which means everyone in the organization would be wise to take advantage of the often easily accessible scientific literature."

4. How well has your training prepared you to deal with evidence in your teaching profession?

Teacher A: "Findings from literature have almost always been provided to help students form opinions on certain subjects or help them complete their assignments. However, searching for the right literature yourself was given little attention, as was designing research yourself. Help was always there if you needed it, but it is not until the final phase of the study that you learn how to properly design a research in a more scientific way. This did help me to understand how to apply research techniques in my daily teaching setting. This way, I can also systematically evaluate interventions when trying out new things. I feel prepared enough to do research in a meaningful way, although I realize that I would need a little more experience and guidance to reach a more academic level."

5. What do you want to add?

Teacher A: "Research was not the direction I wanted to take after my studies. Additionally, being involved in research full time is still not what I want. However, I have seen that research such as I have now done can easily be combined with my work at school. In fact, it ensures that I look more critically at what I do and whether things could be improved. This openness to try new things is especially nice if you look for collaboration with colleagues. I noticed that I have started to look at lessons differently because we try things out and investigate in a systematic way. This ensures that we rise above the level where we mainly look from our feelings at which things work and which do not. This is not to say that from now on I will test everything scientifically and will never try something out spontaneously again, but a major intervention such as the introduction of FtCA as a teaching method is not only more useful and interesting, but also a lot more fun if you conduct substantiated research. If you know what to pay attention to and why you do everything. In that sense, I think I have gotten and will be better at setting up research and

doing larger interventions. Finally, I have found that it is very useful to look at changes from a different perspective. Sometimes a student or colleague immediately provides useful tips, or I can easily see how people really think about my lessons or changes in them. Asking the right questions is very important to find out what the other person's opinion is."

In this reflection we see that the teacher is very positive about "evidence-informed" education. He thinks using evidence is important and inspires him to change his everyday teaching. The project helped him to fill the gap of not knowing how to use research in his profession and stimulated him to continue using evidence during preparation for classes.

4. Discussion

To investigate "To what extent can evidence based on the 'flipping the classroom' approach improve the motivation and results of grade 8 preuniversity track students doing mathematics?", we examined by means of a questionnaire and a teacher interview the intrinsic and extrinsic motivation of the students. The Focus group showed a significant improvement in motivation.

To investigate to what extent intrinsic and extrinsic motivation has increased, a questionnaire was conducted in the Focus group. It also questioned a prerequisite for the success of the FtCA and the possibility of preparation and the feelings of autonomy, competence, and connectedness were measured, with which the degree of motivation can be assessed. On a five-point scale, intrinsic motivation scored 3.36, a slight improvement. Extrinsic motivation showed a stronger improvement at 3.88. An average of 3.72 for the ability to prepare shows that it can be said that students are able to prepare well. Autonomy (3.72), competence (3.60), and connectedness (3.32) also showed improvements. All improvements are statistically significant, so it can be concluded that the introduction of the FtCA has led to more intrinsic and more extrinsic motivation, and students are more motivated for the lessons.

An interview with a math teacher who teaches in a parallel class and who was not teaching according to the FtCA at the time of the experiment largely reveals a similar conclusion. Autonomy, competence, and cooperation, in particular, are noted as strong improvements, and the person concerned is confident that the performance on tests will also improve in the long term according to this method. However, it is argued that the new approach requires becoming accustomed to and that performance, sense of competence, and confidence may also decline among a group of students who have difficulty with math. However, this can be overcome by becoming accustomed to the system and implementing it correctly.

Further, the Focus group and Control group were both presented the same test before and after the introduction of FtCA. The Focus group scored on average 0.45 higher after the introduction, while the Control group scored on average 0.36 worse than on the test before the introduction. However, the differences were not statistically significant, which means that it cannot be concluded whether the FtCA in this environment improves the results. Additionally, the difference was examined between boys and girls of the Focus group. The girls were found to score an average of 0.20 higher after the introduction, while the boys showed an improvement of an average of 0.73. However, these differences also turned out not to be statistically significant. Comez and colleagues [20] had the same results with students on a different level, unlike the findings of Wei [28].

Methodological Limitations

Reasons for the lack of statistically significant differences between the Control and Focus group are not investigated, but because the motivation test was only held in the Focus group, we do not know if students were already intrinsically motivated before the study or not, and students in both groups had access to the same materials. Furthermore, we do not know the students' abilities to learn, and both groups had the same teacher. Future research must find reasons for the very little, insignificant improvement on the test

results. Although we would have liked students to perform better, we think the increase in motivation is an important result for mathematics classes in high schools and see the FtCA as a good possibility for educational change [5].

Investigating whether the impact on one STEM teacher's EIT approach using the FtCA improves the motivation of students in a Dutch setting (eight grade preuniversity level) during mathematics education and if this approach allows students to perform better is promising. The mixed methods approach was chosen because this study is a pilot. It makes it difficult to generalize the findings. In the follow-up in an Erasmus+ founded study, this idea is scaled up to four European countries (five universities) to investigate if EIT following the Engeström cycle [18] can help to narrow the gap between scientific research and the practice of teachers. Of course, this asks for more thorough design instruments (e.g., questionnaires and interview protocols) and reflects more on the teachers' process and progress. Additionally, it is necessary to investigate whether (apprentice) teachers still use an EIT approach when finished with their studies.

References

1. Gough, D. Knowledge Mobilisation in Education in England. In *The Impact of Research in Education*; Levin, B., Qi, J., Edelstein, H., Sohn, J., Eds.; The Policy Press: Bristol, UK, 2013.
2. Hulme, M. Change in Teacher Education and Educational Research. *Res. Intell.* **2013**, *121*, 10–11.
3. Parsons, D.; Burkey, S. Evaluation of the Teaching and Learning Research Programme (Second Phase). Horsham: HOST Policy Research. 2011. Available online: http://www.esrc.ac.uk/files/research/research-and-impact-evaluation/teaching-and-learning-research-programme/ (accessed on 23 March 2017).
4. Pollard, A. Exploring Strategies for Impact: Riding the Wave with the TLRP. In *Knowledge Mobilization and Educational Research: Politics, Languages and Responsibilities*; Fenwick, T., Farrell, L., Eds.; Routledge: London, UK, 2011.
5. Burner, T. Why Is Educational Change So Difficult and How Can We Make It More Effective? *Forskning Forandring.* **2018**, *1*, 122. [CrossRef]
6. DeJarnette, N. America's Children: Providing Early Exposure to STEM (Science, Technology, Engineering and Math) Initiatives. *Education* **2012**, *1*, 77–88.
7. Kulkarni, M.V. A study on secondary school teachers' attitude towards using new technologies in education. *Indian Streams Res. J.* **2012**, *2*, 1–6.
8. Mazur, E. Farewell, Lecture? *Science* **2009**, *323*, 50–51. [CrossRef]
9. Alvarez, B. Flipping the classroom: Homework in class, lessons at home. *Educ. Dig. Essent. Read. Condens. Quick Rev.* **2012**, *77*, 18–21.
10. Bergmann, J.; Sams, A. Before you flip, consider this: Leaders of the flipped classroom movement say each teacher will have a different experience, but securing school leadership support, time, and IT resources will be important to every effort. *Phi Delta Kappan* **2012**, *94*, 25. [CrossRef]
11. Fulton, K.P. 10 reasons to flip. *Phi Delta Kappan* **2012**, *94*, 20. [CrossRef]
12. Saunders, J. *The Flipped Classroom: Its Effect on Student Academic Achievement and Critical Thinking Skills in High School Mathematics*; Liberty University: Lynchburg, VA, USA, 2014; p. 936.
13. Huber, E.; Werner, A. A review of the literature on flipping the STEM classroom: Preliminary findings. In Proceedings of the ASCILITE 2016 Adelaide 2016—Show Me the Learning, Adelaide, Australia, 27–30 November 2016; pp. 267–274.

14. Karabulut-Ilgu, A.; Jaramillo Cherrez, N.; Jahren, C.T. A systematic Review of Research on the Flipped Learning Method in Engineering Education. *Br. J. Educ. Technol.* **2017**. Available online: https://doi-org.proxy.library.uu.nl/10.1111/bjet.12548 (accessed on 5 March 2020). [CrossRef]

15. Seery, M.K. Flipped learning in higher education chemistry: Emerging trends and potential directions. *Chem. Educ. Res. Pract.* **2015**, *16*, 758–768. [CrossRef]

16. Abeysekera, L.; Dawson, P. Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research. *Higher Educ. Res. Dev.* **2015**, *34*, 1–14. [CrossRef]

17. Lo, C.K. Grounding the flipped classroom approach in the foundations of educational technology. *Educ. Tech. Res. Dev.* **2018**, *66*, 793–811. [CrossRef]

18. Engeström, Y. Activity theory and individual and social transformation. In *Perspectives on Activity Theory*; Engeström, Y., Miettinen, R., Punamäki, R.-L., Eds.; Cambridge University Press: Cambridge, UK, 1999.

19. Ryback, D.; Sanders, J. Humanistic versus traditional teaching styles and student satisfaction. *J. Humanist. Psychol.* **1980**, *20*, 87–90. [CrossRef]

20. Gómez-Carrasco, C.-J.; Monteagudo-Fernández, J.; Moreno-Vera, J.-R.; Sainz-Gómez, M. Effects of a Gamification and Flipped-Classroom Program for Teachers in Training on Motivation and Learning Perception. *Educ. Sci.* **2019**, *9*, 299. [CrossRef]

21. Bloom, B.; Engelhart, M.; Furst, E.; Hill, W.; Krathwohl, D. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*; David McKay Company: New York, NY, USA, 1956.

22. Brame, C. Flipping the Classroom. Vanderbilt University Center for Teaching. 2013. Available online: http://cft.vanderbilt.edu/guides-sub-pages/flipping-the-classroom/ (accessed on 8 March 2020).

23. Bransford, J.D.; Brown, A.L.; Cocking, R.R. *How People Learn*; National Academy Press: Washington, DC, USA, 2000.

24. Anderson, L.W.; Krathwohl, D.R.; Airasian, P.W.; Cruikshank, K.A.; Mayer, R.E.; Pintrich, P.R.; Wittrock, M.C. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*; Longman: New York, NY, USA, 2001.

25. Deslauriers, L.; Schelew, E.; Wieman, C. Improved Learning in a Large-Enrollment Physics Class. *Science* **2011**, *332*, 862–864. [CrossRef] [PubMed]

26. Sedig, K. From play to thoughtful learning: A design strategy to engage children with mathematical representations. *J. Comput. Math. Sci. Teach.* **2008**, *27*, 65–101.

27. Jeong, J.S.; Cañada-Cañada, F.; González-Gómez, D. The Study of Flipped-Classroom for Pre-Service Science Teachers. *Educ. Sci.* **2018**, *8*, 163. [CrossRef]

28. Wei, X.; Cheng, I.L.; Chen, N.S.; Yang, X.; Liu, Y.; Dong, Y.; Zhai, X. Effect of the flipped classroom on the mathematics performance of middle school students. *Educ. Tech Res. Dev* **2020**, *68*, 1461–1484. [CrossRef]

29. Ryan, R.; Deci, E. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* **2000**, *55*, 68–78. [CrossRef] [PubMed]

30. De Charms, R. *Personal Causation*; Academic Press: New York, NY, USA, 1968.

31. Deci, E. *Intrinsic Motivation*; Plenum: New York, NY, USA, 1975.

32. Harter, S. Effectance, motivation reconsidered: Toward a developmental model. *Hum. Dev.* **1978**, *1*, 661–669. [CrossRef]

33. White, R. *Ego and Reality in Psychoanalytic Theory*; International Universities Press: New York, NY, USA, 1963.

34. Baumeister, R.; Leary, M. The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* **1995**, *117*, 497–529. [CrossRef] [PubMed]

35. Reis, H. Domains of experience: Investigating relationship processes from three perspectives. In *Theoretical Frameworks for Personal Relationships*; Erber, I.R., Gilmour, R., Eds.; Erlbaum: Hillsdale, NJ, USA, 1994.

36. Deci, E.; Connell, J.; Ryan, R. Self-determination in work organization. *J. Appl. Psychol.* **1989**, *74*, 580–590. [CrossRef]

37. Deci, E.; Ryan, R.; Gagne, M.; Leone, D.; Usonov, J.; Kornazheva, B. Need satisfaction, motivation, and well-being in the work organization of a former Eastern Bloc country. *Personal. Soc. Psychol. Bull.* **2001**, *27*, 930–942. [CrossRef]

38. Reeve, J.; Deci, E.; Ryan, R. Self-determination theory: A dialectical framework for understanding the sociocultural influences on student motivation. In *Research on Sociocultural Influences on Motivation and Learning*; McInerney, D., McInerney, I.D., van Etten, S., Eds.; IAP: Greenwich, CT, USA, 2004; Volume 1, pp. 183–203.

39. Missildine, K.; Fountain, R.; Summers, L.; Gosselin, K. Flipping the Classroom to Improve Student Performance and Satisfaction. *J. Nurs. Educ.* **2013**, *52*, 597–599. [CrossRef]

40. Cronbach, L. Coefficient Alpha and the internal structure of tests. *Psychometrika* **1951**, *16*, 297–334. [CrossRef]

41. Wiertzema, K.; Jansen, P. *Basisprincipes van Communicatie*; Pearson: London, UK, 2016.

42. Shapiro, S.; Wilk, M. An analysis of variance test for normality (complete samples). *Biometrika* **1965**, *52*, 591–611. [CrossRef]

43. Denscombe, M. *The Good Research Guide: For Small-Scale Social Research Projects*, 5th ed.; Open University Press: Maindenhead, UK, 2014.