Computers in Human Behavior 108 (2020) 106318

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: http://www.elsevier.com/locate/comphumbeh

Full length article

Effects of self-regulated learning prompts in a flipped history classroom

David C.D. van Alten *, Chris Phielix, Jeroen Janssen, Liesbeth Kester

Utrecht University, the Netherlands

ARTICLE INFO

Keywords:
Self-regulated learning
Prompts
Flipping the classroom
Flipped learning
Secondary education
History education

ABSTRACT

Flipping the classroom (FTC) is a didactical approach aimed at letting students come to class prepared and apply the learning material actively during class. As FTC places a higher demand on students' self-regulated learning (SRL) skills, our goal in the current study was to research the effects of SRL support in a flipped classroom on students' SRL (self-reports and online activities) and learning outcomes. Previous research showed that video embedded SRL prompts enhances students' SRL and learning outcomes. We measured the effects of SRL prompts with a quasi-experimental design in six flipped History classrooms in secondary education where 154 students were randomly assigned to the SRL prompts or no SRL prompt condition. We found positive effects of the SRL prompts for the completion rate of the instructional videos (i.e., students in the SRL prompts condition watched more videos), but not for other indicators of SRL or learning outcomes. Thus, in contrast to previous research from higher education, our results show that implementing SRL prompts in a flipped classroom is not directly effective in secondary education. We address potential explanations for the absence of effects of the SRL prompts for theory and practice into SRL support in flipped classrooms.

In a flipping the classroom (FTC) approach, students study learning materials before class (e.g., by watching instructional videos) and apply the content of the learning materials during class. Meta-analyses in which the effect of flipped classrooms was compared to traditional classrooms showed small positive effects on learning outcomes (Chen et al., 2018; Cheng, Ritzhaupt, & Antonenko, 2018; Låg & Sæle, 2019; van Alten, Phielix, Janssen, & Kester, 2019). The meta-analyses also showed substantial and significant variation in the effects of FTC on learning outcomes. These studies revealed possible moderating variables affecting the effectivity of FTC, such as adding quizzes to the flipped classroom, not reducing face-to-face time, and academic subject.

What is not yet known about FTC, is the role that students' self-regulation plays while they apply flipped learning. This is surprising, because self-regulated learning (SRL) seems to be a prerequisite for the success of FTC as the individual learning phase before class places a higher demand on students' SRL skills and self-discipline (He, Holton, Farkas, & Warschauer, 2016). In addition, a meta-analysis of Chinese studies (where effects of FTC on SRL are measured more frequently) found that students' SRL skills improved significantly better in flipped classrooms compared to traditional classrooms (Tan, Yue, & Fu, 2017). As flipped classrooms seem to both depend on students' SRL skills and potentially also enhance them, it is worthwhile to investigate to what extent SRL skills should explicitly be supported in a flipped classroom, in which educational contexts, and what kind of support is effective.

In general, supporting SRL is a well-established method to improve student learning outcomes (Dent & Koenka, 2016). Providing students with SRL *prompts* in a computer-based learning environment is one example of an effective strategy to improve students' SRL (Devolder, van Braak, & Tondeur, 2012; Wong et al., 2019). In the context of flipped classrooms in higher education, embedded SRL prompts in an instructional video resulted in better learning outcomes and students performing more SRL activities, compared to students who did not received SRL prompts (Moos & Bonde, 2016). For secondary education students, it is known that providing them with SRL prompts could lead to the internalization of SRL activities and, as a result, improve learning (Greene et al., 2015). In fact, as SRL skills increase at higher ages, it is highly recommended to support secondary education students' SRL to gradually develop their skills (Ramdass & Zimmerman, 2011; Wigfield, Klauda, & Cambria, 2011). However, it is not yet fully known if SRL prompts in flipped classroom videos are as beneficial for secondary school students as for higher education students. Therefore, the current study further investigates the effects of SRL prompts in flipped classroom videos, and aims to reproduce earlier found effects from primary and higher education in the context of secondary education.

* Corresponding author. Department of Education, Utrecht University, P.O. Box 80140, 3508, TC Utrecht, the Netherlands.
E-mail address: d.c.d.vanalten@uu.nl (D.C.D. van Alten).

https://doi.org/10.1016/j.chb.2020.106318
Received 11 October 2019; Received in revised form 16 January 2020; Accepted 24 February 2020
Available online 25 February 2020

1.1. Self-regulated learning in flipped classrooms: theoretical perspectives

The common ground of SRL definitions is that students are actively regulating and monitoring their cognition, metacognition, learning behavior, and motivation while learning, and that this involves strategies and goal processes (Panadero, 2017; Puustinen & Pulkkinen, 2001). Metacognition involves knowledge, awareness, and regulation of one's thinking and is one of the central aspects of SRL (Zimmerman & Moylan, 2009). Self-regulated learners have the ability and motivation to apply metacognitive processes such as planning, monitoring, and controlling their own learning and reflect thereon. In the theoretical model we use, SRL is presented as a cyclic model in which three phases (forethought, performance, self-reflection) are distinguished and learners apply metacognitive strategies such as goal setting and planning in the forethought phase, monitoring and self-control in the performance phase, and evaluation and adaption in the self-reflection phase (Zimmerman & Moylan, 2009).

SRL is a relevant addition to current research on FTC, as a flipped classroom places a strong demand on students' SRL skills (He et al., 2016) and self-directness (J. Lee & Choi, 2019) when studying outside the classroom. First, general meta-analytical research on SRL has shown that supporting SRL in primary and secondary education resulted in better learning outcomes (Dent & Koenka, 2016; Dignath & Büttner, 2008; Donker, de Boer, Kostons, Dignath van Ewijk, & van der Werf, 2014). In addition, there is plenty of evidence from Hypermedia Learning and Multimedia Learning that shows that receiving SRL support enhances students' learning outcomes (Azevedo & Cromley, 2004; Azevedo & Hadwin, 2005; Bannert, Hildebrand, & Mengelkamp, 2009; Bannert, Sonnenberg, Mengelkamp, & Pieger, 2015; Kauffman, 2004). Therefore, as most flipped classrooms contain technology-based learning environments, it seems worthwhile to add SRL support in flipped classrooms to enhance learning.

Second, students appear to utilize SRL strategies in online learning contexts better than in traditional learning contexts (Lee & Tsai, 2011; Sletten, 2017). On the positive side, this implies advantages for students in a flipped classroom, as students are able to control the frequency and pace of the instructional videos before class. This could reduce cognitive load which could be beneficial for learning (Clark, Nguyen, & Sweller, 2011) and increase student's autonomy in their own learning process which could be beneficial for motivation (Hsu, Wang, & Levesque-Bristol, 2019) in comparison to learning material that is presented in a non-manipulable lecture format. In addition, research has found that students who are well able to regulate their own learning process benefit more from a flipped classroom in terms of higher learning outcomes than student's who lack these SRL skills (Lai & Hwang, 2016; J.; Lee & Choi, 2019; Lim & Morris, 2009; Shibukawa & Taguchi, 2019). On the negative side, however, it could be challenging for students to regulate their own learning while preparing the instructional material before class (Sletten, 2017; Wolters, Pintrich, & Karabenick, 2005). For example, Butzler (2016) illustrated that most students in a flipped classroom lacked SRL skills (such as reflecting on their own learning) to fully benefit from the flipped classroom structure. Other researchers also found that the quality of students' preparations before class, and with it the engagement of the students during in-class activities, depends on self-regulated and self-directed skills (e.g., time management and self-discipline; He et al., 2016; Lai & Hwang, 2016). SRL skills of students can thus also be seen as an important prerequisite skill that students need to acquire and apply to benefit from the advantages of FTC.

Third, SRL skills of students who participate in a flipped classroom seem to improve, because FTC places a higher demand on students' SRL skills (Tan et al., 2017). In contrast, Lape et al. (2014) found no significant gains in SRL skills for students in the flipped classroom compared to a traditional classroom. Thus, these inconsistencies in the literature raise the question to what extent teachers and curriculum designers should explicitly support SRL in flipped classrooms in order to make them more effective, in terms of increased SRL skills and better learning outcomes.

1.2. Previous research on SRL prompts in FTC

As it is demonstrated that FTC places a higher demand on students' SRL skills while they learn at home, we aim to support their SRL at that particular learning phase (i.e., during the online videos). One effective way to effectively support SRL is to make use of prompts (Zheng, 2016). Prompts are usually given to students in the form of questions and encouragements (e.g., *What have you learned? Explain what you have learned.*) to incite students to engage in SRL activities. Providing students with SRL prompts in a computer-based learning environment is an effective strategy to improve students' SRL and learning outcomes (Devolder et al., 2012; Sitzmann, Bell, Kraiger, & Kanar, 2009; Wong et al., 2019; Zheng, 2016). As online instructional videos are frequently used in a flipped classroom, SRL prompts in these videos could increase students' SRL activity which - in turn - could lead to higher learning outcomes in a flipped classroom.

It should be noted, however, that there could be possible moderators which means that different students groups benefit differently from SRL prompts. For example, research has shown that effects of SRL prompts can be different according to students' cognitive ability (Sitzmann et al., 2009; Wong et al., 2019). It is also known that students with certain cognitive or metacognitive deficits experience difficulties in completing homework (Bryan, Burstein, & Bryan, 2001). Yeh, Chen, Hung, and Hwang (2010) found an interaction effect between cognitive ability and prompt type, indicating that higher and lower-knowledge learners benefit differently from SRL prompts.

Only a limited number of studies have tested the implementation of SRL support in a flipped classroom. Butzler (2016) found that students worked inefficiently and felt incompetent (e.g., not knowing how to take notes from instructional videos) when SRL was not explicitly supported. Prompts with specific learning strategies at each learning phase in a flipped classroom could help students to actively plan, monitor, and reflect on their learning (Butzler, 2016; Jovanović, Gašević, Dawson, Pardo, & Mirriahi, 2017).

A study conducted by Moos and Bonde (2016) found that embedded SRL prompts in an instructional video resulted in better learning outcomes and students performing more SRL activities when learning from the video, in contrast to students who did not receive SRL prompts. In addition, no effect of prompts was found on students' perceived mental effort. Moos and Bonde (2016) prompted students in the forethought phase to set goals and strategically plan their learning, in the performance phase to monitor and control their learning, and in the self-reflection phase to evaluate their own learning and set goals. A typical forethought prompt was, for instance: *What strategies do you think will be effective while learning about* [the topic of this video]*?*; a performance prompt: *What information have you learned so far?*; and a self-reflection prompt: *What could you have done differently while learning about* [the topic of this video]*?* With these SRL prompts, both the quantity and quality of SRL activities were enhanced by informing students about effective strategies and prompting them at the forethought, performance and self-reflection phase of learning (Moos & Bonde, 2016; Zimmerman & Moylan, 2009).

In addition, two studies showed that students achieved higher learning outcomes when they had access to an additional SRL support learning system, in contrast to students who did not have access. First, Lai and Hwang (2016) compared a flipped classroom in which primary students were supposed to use a self-regulated monitoring system to set goals and evaluate their own learning with a regular flipped classroom. The students in the SRL supported flipped classroom improved significantly in terms of self-efficacy, strategies of planning study time, and

learning outcomes, in contrast to students in the regular flipped classroom. Second, Shyr and Chen (2018) showed that a technology-enhanced language learning system that scaffolded students' SRL was able to increase students' learning outcomes and self-reported SRL as compared to students in a regular flipped classroom. Students in the intervention group who had access to this language learning system received SRL prompts while they selected learning tasks, and by scaffolding, students were supported on orientation, goal-setting, planning, monitoring, and assessing their own learning tasks.

For the following reasons, our study extends the previous research. First, earlier studies that yielded positive effects of SRL support in a flipped classroom compared two flipped groups, where the intervention group had access to an additional SRL support system (e.g., Lai & Hwang, 2016; Shyr & Chen, 2018). However, it is possible that the results they found were due to differences in time-on-task between the groups with or without the SRL support caused by the additional support system. Therefore, in this study, we investigated the effects of SRL support while keeping *time-on-task* between groups as equal as possible.

Second, there is a lack of data regarding FTC in *secondary education* as most research on SRL support in FTC was conducted in primary or higher education. It is important to know if earlier results of explicit support of SRL by prompts also apply to secondary education. It is clear that students of different age differ in SRL skills, and that SRL skills increase at higher ages (Dent & Koenka, 2016; Wigfield et al., 2011; Zimmerman & Martinez-Pons, 1990). The fact that students' confidence in their SRL skills from primary to secondary education decreases (Wigfield et al., 2011), indicates extra importance to apply SRL support in secondary education. Similarly, Ramdass and Zimmerman (2011) argued that it is important that students from all ages gradually learn to develop SRL by repeated practice, as it enhances learning. Therefore, we conducted our study in secondary education as it seems even more important that SRL of younger students in flipped classrooms is supported.

Third, the current study took place in an ecologically valid classroom setting and aimed to investigate effects over the course of six weeks. Previous research often focuses on short term interventions of SRL prompts and sometimes in an out-of-school, laboratorial learning environment with a restricted learning time (cf. Bannert et al., 2015; H. W.; Lee, Lim, & Grabowski, 2010; Moos & Bonde, 2016).

1.3. Present study

In the present study, we investigated the effect of explicit support of SRL with prompts in a flipped classroom, on students' SRL self-reports, online SRL activity, and learning outcomes. We answer the following research questions:

1. What is the effect of SRL prompts in a flipped classroom on secondary education students' (a) SRL self-reports and (b) online SRL?
2. What is the effect of SRL prompts in a flipped classroom on learning outcomes of students in secondary education?
3. Does students' performance level moderate the effects of SRL prompts?

With regard to the first research question, we operationalized SRL in two ways: (1) as a self-reported construct and (2) as online log data measures of SRL activities (Panadero, Klug, & Järvelä, 2016; Rovers, Clarebout, Savelberg, de Bruin, & van Merriënboer, 2019). Thus, we measured SRL in two ways: (1) by asking students how they perceived their own SRL and (2) by investigating SRL activities that are traceable in online log data. Jovanović et al. (2017) showed that it is possible to identify actions in students' learning behavior as manifestations of their learning strategies. This is considered to be an important addition to self-reports, as it aims to measure interaction traces of students' SRL behavior in a natural setting (Jovanović et al., 2017; Panadero et al., 2016; Winne, 2014). For the second research question, we defined learning outcomes as students' score on a learning outcome test similar to tests that are regularly used in secondary education to assess their learning. With regard to research question three, we used *performance level* in a moderator analysis to investigate a possible interaction effect between performance level and condition. We defined performance level as the students' average scores for History that current schoolyear. An interaction between performance level and the research condition would mean that SRL prompts work better for students with a higher, or lower, performance level. Lastly, we evaluated students' answers on the satisfaction questionnaire and their answers on a semi-structured focus group interview to better interpret our findings with regards to their experiences in our research intervention.

All in all, our hypothesis is that students who are regularly prompted to perform SRL activities during their learning process, will (a) report higher SRL self-report values, (b) show more SRL activities during online learning, and (c) will achieve higher learning outcomes, in comparison with students that are not provided with SRL prompts.

2. Method

2.1. Participants

A total of 169 second year students of (1) senior general and (2) pre-university level within Dutch secondary education were asked to participate in this study (including opt-out consent from their parents). The Netherlands has a tracked secondary education system (from grade 7 onwards) and these are the two highest tracks. Ethical approval was obtained from the Faculty Ethics Review Board in April 2018. The study took place in one school and the intervention was part of the students' regular school curriculum. Students who decided not to participate in this research still had to attend the lessons and complete the final test, but their results were not analyzed by the researchers. One voucher worth of 10 euros was raffled amongst the students in each participating class, while every student in every classroom received a small gift at the end of the study. Eight students were excluded because consent was not obtained and seven students were excluded because they missed more than half of the lessons and/or they were absent when the measurements were taken. Consequently, our sample consisted of 154 students (M_{age}=13.49 years, $SD_{age} = 0.50$; 84 boys and 70 girls). The sample came from six pre-existing classrooms: two senior general level classrooms ($n = 44$) and four pre-university level classrooms ($n = 110$). A priori power analyses indicated that, for sufficient power of 80% and detecting medium to large effects as expected by previous research, we needed to have a minimum sample size of 152 students. Thus, our final sample size of 154 students was adequate.

The first author was one of the four teachers. The other three teachers volunteered to participate and each received a voucher worth of 25 euros at the end of the study. The average teaching experience of the participating teachers was 4 years, and their experience with teaching flipped classrooms was comparably little.

2.2. Design

The quasi-experiment used a between-subjects design with embedded SRL prompts in the instructional videos in the intervention group, and no embedded SRL prompts in the control group. The experiment consisted of eight lessons (and seven videos) over the course of six weeks. Stratified randomization was used to allocate students within every class to the SRL prompts condition ($n = 74$) or the no SRL prompts condition ($n = 80$). The students' average scores for History (recalculated in a standardized rank score within each classroom) were used to match students and make sure both groups were comparable in terms of performance level. The rationale for this stratified randomization within each classroom was to control for possible differential effects on the classroom level (e.g., teacher effect, cf. Nye, Spyros, & Hedges,

2004).

2.3. Materials

2.3.1. Online learning environment

Edpuzzle (https://edpuzzle.com/) was used as online learning environment to share the videos with the students. We provided each student with a personal login that was linked to their corresponding research condition (video with or without SRL prompts). The instructional videos were identical for both groups, had an average length of 7 min and were developed and recorded by the first author. The school's digital homework system was used to announce which video had to be viewed in preparation before class.

2.3.2. In class learning materials

The subject of the lessons developed for this intervention was the industrial revolution during the late 19th and early 20th century, a compulsory part of the curriculum at the senior general and pre-university educational levels. The in-class learning materials were mostly based on the schoolbooks that are used for teaching History in the participating school. In addition, every student was provided with a physical workbook which included instructions about Edpuzzle, lesson procedures, and three to four corresponding cognitive questions (in both conditions) for each instructional video that was required to complete before class. Teachers discussed the cognitive questions at the start of each lesson to recall the content of the instructional videos. Additionally, we developed assignments in which students were actively engaged to apply the content of the learning materials (e.g., analyzing historical sources and creating advertisement leaflets to explain the benefits of 19th century industrial inventions). Participating teachers were able to give feedback on all the learning materials before the intervention started. As the students were allocated to research conditions within classrooms, all the in-class learning materials were exactly the same for both research conditions.

2.3.3. SRL prompts

The students in the SRL prompts condition received three to four SRL prompts per video (similar to Moos & Bonde, 2016) according to the three SRL phases of forethought, performance, and self-reflection (Zimmerman & Moylan, 2009). For example, in the forethought phase one of our SRL prompts was: *How do you make sure that you will be able to recall all the learning material from this instructional video?* The practical example that students received after answering this prompt was: *For example, you can regularly pause the video and make notes to summarize the learning material.* An example of a SRL prompt in the performance phase is: *Do you understand everything you heard so far? If not, what can you do to fix it?* Lastly, a typical SRL prompt in the self-reflection phase at the end of the instructional video looked like this: *Did you achieve your goal(s) you set at the beginning of this instructional video? If not, what do you need to complete it?* An overview of all the SRL prompts per video, focusing on the whole process of SRL in each SRL phase is presented in Appendix A (cf. Zheng, 2016 on supporting each aspect of SRL instead of one specific learning phase). The prompts were presented to students as forced open question stops within the video. Students were required to think about the SRL prompts and answer them to continue the video. In around half of the SRL prompt, a specific hint popped up after a student's answer to show a practical example of that SRL behavior (see Appendix A). Students in both research conditions were not able to fast forward the video, but were allowed to rewind and re-watch the videos (or portions of the videos) unlimited.

2.4. Measurements

2.4.1. SRL log data

Edpuzzle was also used to collect SRL log data. We created the following four variables as indications of SRL activity: Students' *video completion rate*, *video watch time*, *video portions viewed*, and students' *rewind actions* (i.e., how many times a student reviewed a portion of the video). *Video completion rate* and *video watch time* can be seen as traces of SRL activity, as Cicchinelli et al. (2018) showed that students with high SRL accessed online learning platforms more and spent more time on it (cf. Sitzmann & Ely, 2010). *Video portions viewed* and *rewind actions* can be seen as manifestations of learning regulation activities, as students take action to acquire or reinforce knowledge (Bannert, Reimann, & Sonnenberg, 2014; Cicchinelli et al., 2018; Kizilcec, Pérez-Sanagustín, & Maldonado, 2017; Maldonado-Mahauad, Pérez-Sanagustín, Kizilcec, Morales, & Munoz-Gama, 2018).

In addition, to evaluate if the students in the SRL-prompt condition properly worked with the prompts, we collected their answers that they had to provide to continue the video. We counted how many valid (e.g., serious) answers were given by students who watched the videos and calculated an average per video to assess SRL-prompt compliance.

2.4.2. SRL questionnaire

The first method we used to measure SRL activity was a contextualized version of the revised version of the self-regulated online learning questionnaire (SOL-Q-R) with a pretest-posttest design (Jansen, van Leeuwen, Janssen, & Kester, 2018; Jansen, van Leeuwen, Janssen, Kester, & Kalz, 2017). The revised version of this questionnaire contains the following seven scales: *metacognitive activities before, during*, and *after* learning, *time management, environmental structuring, persistence,* and *help seeking*.

We translated the SOL-Q-R that was originally constructed for online learning in Dutch, with a student population aged 14–15 years in mind. Potential difficult words (such as: *to reflect*) were explained at the bottom of the questionnaire. We substituted references to *online course* in the questionnaire into *homework for History* (e.g., I think about what I have learned after I finish working on homework for History), to adjust it to the educational context and to be able to use it as pretest as well. As can be seen in Table 1, the scales of our contextualized version of the SOL-Q-R were reliable, and, in addition, the reliabilities were very similar to the three datasets used to validate the original questionnaire in the context of blended learning in higher education (Jansen et al., 2018).

2.4.3. Motivation questionnaire

Motivation was measured as a pretest only to investigate if the students in the research conditions were equal in terms of motivation. We used four relevant motivation scales from the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1991). We also contextualized this questionnaire to the context of completing homework for History, to be able to measure students' motivation in this particular context as precise as possible. *Intrinsic goal orientation*, *extrinsic goal orientation*, and *task value* were used to measure how students valued and perceived their engagement into a particular learning task (i.e., homework for History). *Self-efficacy for learning* was included to measure students' expectancy beliefs about how they complete their History homework. As can be seen in Table 2, the scales of our translated and contextualized version were acceptable except for the

Table 1
Internal-Consistency Reliabilities of the Contextualized SOL-Q-R Scales.

Scale	pretest			posttest		
	n	Items	*α*	*n*	Items	*α*
Metacognitive activities before	146	7	.779	131	7	.804
Metacognitive activities during	149	7	.735	136	7	.798
Metacognitive activities after	142	6	.833	133	6	.846
Persistence	142	7	.896	136	7	.903
Help seeking	149	6	.744	130	6	.821
Environmental structuring	151	4	.773	133	4	.791
Time management	146	5	.740	129	5	.672

Table 2
Internal-Consistency Reliabilities of the Contextualized MSLQ Scales.

Scale	*n*	Items	*α*
Intrinsic goal orientation	151	4	.760
Extrinsic goal orientation	149	4	.629
Task value	141	6	.911
Expectancy: self-efficacy for learning	141	8	.890

extrinsic goal orientation scale. We nevertheless kept this scale because the value is still above 0.5 (which would be unacceptable) and the scale is only used to check for research group equivalence.

2.4.4. Learning outcome test

The researchers and involved teachers developed a learning outcome test based on the learning materials that are used for teaching History in that school and comparable to what students regularly have to complete during the schoolyear. Every participating student received the same test. The test consisted of 25 items, of which five were multiple choice (four answer options each). The 20 open questions contained both recall (i.e., reproduction), comprehension (i.e., understanding), and transfer (i.e. applying) assignments. For example, a recall question was: *Explain the following concepts: Cottage industry, urbanization and capitalism*. A comprehension question was: *Explain why a rapid need arose for the Spinning Jenny during the Industrial Revolution.* A transfer level question was, for instance: *Give an argument why the following source represents the working conditions in the 19*th *century*. In general, the test was found to be reliable (25 items; $\alpha = 0.79$).

2.4.5. Student satisfaction

We evaluated how the students valued the flipped classroom method as a means to better interpret our findings. We added seven Likert-items (range from 1 to 7) to the posttest as an anonymized satisfaction questionnaire (e.g., I think the instructional videos are boring; I think the instructional videos are instructive; I believe that the instructional videos are designed in a clear and understandable way). The *satisfaction* questionnaire was found to be reliable (7 items; $\alpha = 0.87$). Besides, we asked students suggestions for improvements in the videos, and if they want to share any other compliment or complaint about the FTC method, in two open answer questions.

In addition, we also conducted small-scale semi-structured focus group interviews to have a better understanding of experiences of students and to aid the interpretation of our results. The first author conducted three semi-structured interviews with three randomly chosen students (on voluntary basis). Two interview sessions were done with three students from the SRL prompts condition each, and one session with three students from the no SRL prompts condition). Each focus group interview lasted around 10 min and students were asked to evaluate their experiences in this flipped classroom, the in-class materials, and the homework (instructional videos and accompanying questions). In addition, students from the SRL prompts condition were specifically asked how they valued the SRL prompts.

2.5. Procedure

Before the flipped intervention started, the students completed the pre-test questionnaire with the SRL and motivation items during regular class time. At the start of the intervention, students received a brief instruction about working in a flipped classroom and a small rationale behind it, as for most students it was their first experience. In addition, students received a classroom instruction on how to access the videos at home.

Students were informed that their own teacher could track their progress. Teachers were instructed not to change their usual teaching interaction with their students, and received instructions and training on how to apply the learning materials each lesson. The teachers used the school's digital homework system to inform students of their own class about the required preparations for each lesson (one instructional video and the accompanied content questions in their workbooks). Each lesson of 65 min started with retrieval practice activities about the content of the videos and a discussion of the content questions from the workbooks. This was usually followed by a complementary micro lecture of approximately 10 min to discuss not yet fully understood parts and more in-depth aspects of the learning materials. The largest and remaining part of the lesson was dedicated to engaging learning activities to apply the learning materials. After eight lessons, students completed the posttest SRL and satisfaction questionnaire during regular class time. In the subsequent lesson, students had 65 min to complete the learning outcome test.

2.6. Data-analysis

For the SOL-Q-R and satisfaction questionnaires we calculated scale means of the relevant items, where all students were included if they filled in at least $n - 1$ items of that scale. Log data from Edpuzzle was used to calculate four SRL activity variables. For *completion rate* of the videos, we calculated the mean of the video completion rate of each individual student. For students' *video watch time* on and *video portions viewed*, we added up the individual scores of each student on these variables. For students' *rewind actions*, we counted and added up all rewind actions from all videos.

We collected the students' average scores for History that current schoolyear as indication of performance level and recalculated that in a standardized rank score within each classroom. To include performance level as covariate in our analyses, we recalculated rank scores into standardized Z-scores within each classroom to be able to make a fair comparison between performance level of individual students.

For the learning outcome test, a comprehensive assessment scheme was validated by the researchers and teachers to score the students' tests. Each teacher scored the tests of their own students and we created a moment early in the scoring phase in which the teachers discussed cases where they doubted their scoring of an answer to an open question. To assess the interrater reliability of the open question scores, the first author independently scored a random sample of five unscored tests from the other classrooms (20 tests in total). We compared the scores of the first author (400 items in total) with the scores of the corresponding teacher from that classroom. After evaluating the Cohen's κ for each item, the agreement percentage, and the item total correlation, we deleted five items because of low reliability (i.e., $\kappa < 0.55$). This resulted in 15 remaining open questions with an average total agreement of 85%, and a moderate to substantial interrater-reliability with an average Cohen's κ of 0.74 (range $\kappa = 0.55$ to 1, $p < .001$). Finally, we created a sum score of the 15 remaining open answer questions and the 5 multiple choice questions, resulting in a maximum score of 24 points for the learning outcome test (as some open answer questions had a greater weight on the test).

Finally, to answer the first two research questions (i.e., is there a difference between the SRL prompts and no SRL prompts conditions on SRL self-reports, SRL activities, and learning outcome), we checked the data for normality and outliers. Due to the linear relationships between the dependent variables, as shown in the correlation matrix in Table C1 (Appendix C), we analyzed the differences between both research conditions by two separate MANOVAs (i.e., SRL self-report scales and SRL log data) and one ANOVA (i.e., learning outcomes) in SPSS. As the SRL log data variables were not normally distributed, we performed a bootstrapping method (1000 Bootstrap samplings with replacement, cf. Field & Wilcox, 2017). In addition, we included educational level (i.e., senior general or pre-university level) in the model as controlling factor.

To answer the third research question and examine if performance level moderated the findings, we added students' performance level as covariate and checked for possible interaction effects between performance level and condition, which would mean that the SRL prompts had

a differential effect according to how well students usually performed in History. We also included *educational level* in this model as controlling factor.

Lastly, to better interpret our findings, we compared the answers on the satisfaction questionnaire with the answers on the semi-structured focus group interviews. These interviews were transcribed and subsequently analyzed by recurring themes such as students' experiences of the flipped classroom in terms of advantages and disadvantages, learning from the videos before class, and, for the SRL prompts condition, how they valued the prompts.

3. Results

Independent samples *t*-tests revealed no significant difference ($p > .05$) between the students in the SRL prompts and no SRL prompts conditions *before* the intervention for the pretest SOL-Q-R scales, the MSLQ scales, and students' performance level (standardized average of students' current History scores).

In our flipped classroom and research design, videos were used to convey important subject matter before class. Students were instructed to independently watch these videos out of the classroom. It is important to know to what extent the students actually watched these videos. First, in order to know the amount of subject matter they were exposed to before class. Second, and more importantly, to evaluate the effect of non-compliance to the intervention on our results (cf. He et al., 2016; Müller & Seufert, 2018). For example, non-compliance of participants in a research intervention could hamper statistical power to detect intervention effects (Jo, 2002). In our case, a low *video completion rate* for the students in the SRL-prompts condition means that they were not substantially exposed to our intervention. In fact, it would mean that the students in the SRL-prompt condition who did not watch the videos are more similar to students in the control group in the sense that they also did not receive the SRL prompts. To summarize, students who had a low *video completion rate* could hinder a fair comparison between both conditions.

The *video completion rate* variable showed that, in total, 17 students did not watch any videos before class, 67 students watched all videos, and 105 students (77%) watched at least five out of seven videos. We decided to answer our research questions and perform our analyses on the group of students (both research conditions) who watched a substantial amount of the videos (more than two-thirds of the videos). We chose five out of seven videos as our threshold as this is comparable to the 80% threshold in other research dealing with non-compliance (Armijo-Olivo, Warren, & Magee, 2009). In Table 3, we present the descriptive statistics for all the dependent variables for this group of students. Because the exclusion of non-compliance students could introduce a selection bias (Armijo-Olivo et al., 2009; Fergusson, Aaron, Guyatt, & Hébert, 2002), we present our results on the entire group of students in Appendix B for transparency purposes.

Lastly, we assessed SRL-prompt compliance by the average amount of valid answers students provided in response to the prompts in order to continue the video. The average valid answers per video was 90% and ranged from 85% (video 7) to 95% (video 5). This indicates that students in general properly worked with the prompts.

3.1. Effects of SRL-prompts

3.1.1. SRL self-reports

First, A MANOVA that included condition (SRL prompts versus no SRL prompts) and educational level (senior general versus pre-university) as factors and the seven SOL-Q-R scales as dependent variables was used to investigate if students who received SRL prompts reported higher SRL self-reports than students who did not receive SRL prompts (RQ1a). We found no significant main effect of conditions on the SOL-Q-R scales, $F(7, 91) = 2.01$, $p = .062$; Wilk's $\Lambda = 0.866$, partial $\eta^2 = 0.13$. We are aware that this omnibus test is not significant, but we decided to present the results of the follow up ANOVA for completeness reasons. Due to the non-compliance exclusion, our analyses lost some power and these post-hoc tests have more power to detect differences. These results, nevertheless, should be interpreted with caution and seen as explorative result interesting for future research. ANOVA revealed a significant difference between conditions on the SOL-Q-R *help seeking* scale, $F(1, 97) = 6.22$, $p = .014$, partial $\eta^2 = 0.06$. Students who received SRL prompts reported that they performed more help seeking activities while completing their homework than students who did not receive the prompts (see Table 3 for the descriptive findings).

Table 3
Mean and Standard Deviation for the Dependent Variables (Included Students who Watched at least 5 out of 7 Videos).

Dependent variables (posttest)[a]	SRL prompts condition			No SRL prompts condition		
	n	M	SD	n	M	SD
SRL self-report scales ($p = .062$)						
Metacognitive act. before	51	3.51	1.17	51	3.45	1.13
Metacognitive act. during	51	3.69	1.17	52	3.50	1.14
Metacognitive act. after	51	3.10	1.23	52	2.80	1.17
Persistence	51	4.87	1.18	52	4.70	1.34
Help seeking	51	3.97	1.19	52	3.35	1.14
Environmental structuring	51	5.37	1.40	51	5.41	1.41
Time management	51	4.92	0.96	52	4.64	1.04
SRL online activity ($p = .037$)						
Video completion rate	51	95.66	7.37	54	92.67	10.44
Video watch time	51	68.24	22.56	54	66.35	21.46
Video portions viewed	51	92.27	29.80	54	85.07	27.86
Rewind actions	51	25.18	27.96	54	20.17	24.66
Learning outcomes ($p = .569$)						
Learning outcome test	51	16.72	3.87	54	16.13	3.91

[a] *Note*: Provided *p*-values are from the comparison of SRL-prompts condition in the separate MANOVA on *SRL self-report scales*; MANOVA on *SRL online activity*, which included video completion rate and rewind actions; ANOVA on *learning outcomes*.

Furthermore, we found a significant effect of educational level on the SOL-Q-R scales, $F(7, 91) = 3.31$, $p = .003$; Wilk's $\Lambda = 0.797$, partial $\eta^2 = 0.20$. Follow up ANOVAs showed that for the SOL-Q-R scales *help seeking* ($p = .014$, partial $\eta^2 = 0.04$) and *time management* ($p = .002$, partial $\eta^2 = 0.09$), students in the pre-university level scored significantly higher (*help seeking*: $M = 3.80$, $SD = 1.22$; *time management*: $M = 4.93$, $SD = 0.99$, $n = 80$) than students in the higher senior level (*help seeking*: $M = 3.17$, $SD = 1.03$; *time management*: $M = 4.16$, $SD = 0.91$, $n = 21$).

3.1.2. SRL online activity

Second, a MANOVA with condition and educational level as factors and *video completion rate* and *rewind actions* as online SRL activity dependent variables was used to investigate if students who received SRL prompts performed more online SRL activities than students who did not receive SRL prompts (RQ1b). We dropped *video watch time* and *portions viewed* from the analysis because of a very high correlation ($R > 0.96$, $p < .01$, see Appendix C) and the problem of multicollinearity. We found a significant main effect of conditions on online SRL activity, $F(2, 100) = 3.41$, $p = .037$; Wilk's $\Lambda = 0.936$, partial $\eta^2 = 0.06$. Results of the follow up ANOVAs determined a significant difference for the SRL prompt and no SRL prompt conditions on the online SRL activity variable *video completion rate*, as the students in the SRL prompts had higher *video completion rates* than students who did not receive SRL prompts, $F(1, 101) = 6.88$, $p = .010$, partial $\eta^2 = 0.06$. We found no significant main effect of educational level on online SRL activity, $F(2, 100) = 1.35$, $p = .264$; Wilk's $\Lambda = 0.974$, partial $\eta^2 = 0.03$.

3.1.3. Learning outcomes

Third, an ANOVA on the learning outcome test scores as dependent variable with condition and educational level as factors was used to investigate if students who received SRL prompts obtained higher

Table 4
Results of the open answers from the satisfaction questionnaire, n = 154 students. In parentheses is given how often a comparable answer was given.

1. Is there something that can be improved in the videos? (e.g., anything disturbing, unclear?)	2. Do you have any other compliments or complains?
Too long and/or boring (18)	Videos were instructive and/or comprehensible (12)
Face cam distracting (14)	Better than usual teaching method (12)
Negative comment about the prompts (9)	Nice method (10)
Lacks entertainment (7)	Questions accompanying the videos help me to focus (2)
Quality of the microphone (5)	Method motivates me to complete my homework (2)
Simpler language (3)	Prefer the usual teaching method (2)
Not able to fast forward videos (2)	Well organized (2)
	Method requires less time to learn (2)

learning outcomes than students who did not receive SRL prompts (RQ2). We found no significant main effect of condition on learning outcome; $F(1, 101) = 0.33$, $p = .569$, partial $\eta^2 = < 0.01$, but we did find a significant main effect of educational level on learning outcome; $F(1, 101) = 7.71$, $p = .007$, partial $\eta^2 = < 0.07$. Students in the senior general level scored 12.92 points on average (12 points required to get a passing grade, $SD = 3.40$) and students in pre-university level scored on average 16.20 points (13.2 points required for a passing grade, $SD = 4.07$). The total fail rate of the test was 30%, which indicates that the test was not too difficult nor too easy (e.g., an easy test could possibly result in a ceiling effect).

3.2. Moderation of the effects of SRL-prompts by performance level

While educational level contains information about students' abilities on the classroom level, performance level (i.e., the students' average scores for History that current schoolyear) contains information about students' abilities on the individual level for a particular subject matter. Due to matching, both research conditions contained students of equal performance levels. The possible interaction effect between performance level and condition on the dependent variables would indicate that the SRL prompts have different effects for students with a higher or lower performance level.

Thus, we performed two separate MANCOVAs and one ANCOVA with performance level as a covariate, condition type and educational level as factors, and the SRL questionnaire scales, the SRL online activity variables, and the learning outcome test as dependent variables to investigate if performance level moderated the effects of SRL prompts (RQ3). In Table C2 (Appendix C) we present the correlations between performance level and the dependent variables. We found no significant interaction effects between performance level and condition on the SRL self-reports, $F(7, 90) = 1.06$, $p = .394$; Wilk's $\Lambda = 0.924$, partial $\eta^2 = 0.08$; no significant interaction effects between performance level and condition for the SRL online activities, $F(2, 99) = 0.34$, $p = .713$; Wilk's $\Lambda = 0.993$, partial $\eta^2 = 0.01$; and no significant interaction effects between performance level and condition for learning outcome; $F(1, 100) = 1.33$, $p = .252$, partial $\eta^2 = < 0.01$. This shows that the effect of the SRL prompts is not different for students' performance level.

3.3. Student satisfaction in the flipped classroom

The effectiveness of the SRL prompts is related to the design of the flipped classroom and how students valued it in general. Therefore, we evaluated the results from the satisfaction questionnaire to interpret students' evaluations of our flipped classroom implementation, as negative evaluations of our research setting could potentially affect the outcomes.

We decided to include all students (including the non-compliant students) in the analysis of satisfaction, as it could be that *video completion rate* was dependent on students' satisfaction about the flipped classroom. The correlation between *video completion rate* and the satisfaction questionnaire was $r(140) = 0.23$, $p = .006$. We performed an explorative ANCOVA with condition and educational level as factors and controlling for *video completion rate* as covariate on the total score of the satisfaction questionnaire. After controlling for *video completion rate*, the ANCOVA showed no main effect of condition on satisfaction, $F(1, 136) = 0.04$, $p = .833$, partial $\eta^2 < 0.01$, but a significant main effect of educational level on satisfaction, $F(1, 136) = 17.59$, $p < .001$, partial $\eta^2 < 0.12$. Students in the pre-university level ($M = 36.17$, $SD = 8.23$, $n = 103$) valued the flipped classroom and the instructional videos significantly more positively than the students in the senior general level ($M = 28.62$, $SD = 8.47$, $n = 37$), with a maximum score of 49. All in all, students were generally positive about the flipped classroom. This is further illustrated with the findings presented in Table 4. It shows the results of the open questions in which all the students were asked for improvements in the videos, or provide any other compliment or complaint about the FTC method. The majority of students answered that they had no specific feedback and relatively few students provided feedback.

Finally, the semi-structured focus group interviews showed that students in general were satisfied with the flipped course and would like to use this method more often. For example, they valued the variety this method adds to regular teaching methods and the quality and organization of the course. Some students expressed that the videos helped them understand the learning material better, and that it reduced the time they usually spend to History homework. In contrast, another recurring theme was the recommendation to shorten the videos, and add some variety in the style of the videos. When asked about their experience of the accompanying lessons, some students remarked that the lessons contained a high amount of learning activities where they had to actively apply the learning material, but that these activities were also present in the usual lessons before the flipped course. Students who admitted that they did not watch (all of the) online instruction videos gave as reason that they usually lack motivation to complete homework, also for their homework of other subjects. Some students from the SRL prompts condition complained about the SRL prompts. They felt that these prompts distracted them from learning the content. When asked for clarification, they said that they are not used to think about regulating their own learning, and that they do not see the added value of this activity. The content questions in their workbooks were positively valued, as they liked to use them to make sure if they understood the learning content.

4. Discussion

The main purpose of the current study was to investigate the effect of video embedded SRL prompts on students' SRL (self-reports and online activities) in a flipped history classroom over a period of six weeks. We found that students provided with SRL prompts did not significantly improve in terms of SRL and learning outcomes in comparison with students who were not provided with the SRL prompts. These results could have implications for both theory and practice, as previous research in primary and higher education found large effects on students' SRL and learning outcomes by implementing SRL support (e.g., with video embedded prompts) in FTC (cf. Moos & Bonde, 2016). However, our results show that SRL support does not yield similar

results in the context of secondary History education. We will address possible explanations to explain these findings related to, for example, the research context, type of prompt implementation, methodological aspects, or a combination of them.

4.1. Did the SRL-prompts affect SRL self-reports and SRL activity?

4.1.1. SRL self-reports

In contrast to our expectations based on previous research (Lai & Hwang, 2016; H. W.; Lee et al., 2010; Moos & Bonde, 2016; Shyr & Chen, 2018), the results on the SOL-Q-R posttest showed no significant effects of the prompts. Although we need to be cautious to interpret this finding, as the MANOVA omnibus test was not statistically significant ($p = .06$), we saw a significant difference between the SRL and no SRL prompts conditions on the *help seeking* scale. It can therefore not be ruled out that this finding can be seen as a false positive. However, as we lost some power due to the exclusion of 23% of the students due to non-compliance, it is worth to at least discuss this difference to direct further research. For example, Sun, Wu, and Lee (2016) investigated the effect of FTC (in contrast to non-FTC) on SRL, and also only found an effect on help seeking. In our intervention, it would mean that students in the SRL prompts condition were more inclined to ask questions to others during their learning process than students without the prompts. As can be seen in the overview of prompts (Appendix A), a few prompts explicitly prompted students to ask questions to others when they felt they needed help. It could be possible that students without the SRL prompts paradoxically perceived more control and competence while learning from the videos, which gave them less urge to ask questions. It is also possible that due to the Dunning–Kruger effect students in the no SRL prompts group overestimated their own learning (and asked less questions), because they were less aware of their own learning process as they did not receive SRL prompts. Students in the SRL prompts condition, in contrast, were explicitly prompted to think about their learning process and were therefore probably more aware of what they did not understand, and possibly applied the help seeking strategy more often.

4.1.2. Online SRL activity

The results for students' online SRL activity only showed a significant difference in *video completion rate*. This means that students in the prompts condition watched more videos than students in the no prompts condition. This effect could indicate that the SRL prompts made students more aware of their learning process and that they therefore completed their homework more consistently than students without the SRL prompts. This could, for instance, be caused by SRL prompts that explicitly connected the coherence across the homework videos.

However, we also expected to find differences in the other SRL activity indications we measured (such as *video watch time*, *video portions viewed*, and *rewind actions*; Bannert et al., 2015). In previous research, for example, Sitzmann and Ely (2010) found that positive effects of SRL prompts on learning outcomes were mediated via the total time students spent on learning. Our results did not show any significant difference for *rewind actions*. Our SRL prompts explicitly encouraged students to monitor their understanding of the content, and were frequently prompted to rewind video portions (there were 70 video portions to be viewed in total) when they would encounter difficulties. In our results, we see high standard deviations for the *rewind actions* for both conditions, and that students with SRL prompts on average rewound five more video portions than students without the prompts. This indicates that, independent from the research condition, there is a group of students who hardly rewind video portions, in contrast to a group of students who rewinds video portions multiple times. Although we did not find any statistically significant effects of the SRL prompts on these data, we can carefully draw the conclusion that our FTC intervention (even without the SRL support) stimulated (at least a part of) the students to monitor their own learning and apply SRL. It is worthwhile for future research to investigate this contrast in learning behavior on the student level, to investigate potential explaining factors at this level.

It should also be noted that *video watch time*, *video portions viewed*, and *rewind actions* are very highly correlated $R > .96$ (see Appendix C). This can be explained as the video watch time is only increased if a student rewound a portion of a video. The fact that we had to drop two of these variables from the analysis due to multicollinearity seems a minor problem, as it appears to measure one underlying SRL concept of deliberately acquire or reinforcing knowledge (Bannert et al., 2014).

4.2. Did the SRL-prompts enhance learning outcomes?

In contrast to previous research (Bannert et al., 2009; Lai & Hwang, 2016; Moos & Bonde, 2016; Sitzmann & Ely, 2010; Sitzmann et al., 2009), we did not find differences in learning outcomes between the students in the conditions with and without SRL prompts. First, Müller and Seufert (2018) suppose that the lack of effects on learning outcomes could be due to student compliance. As example, they show that student compliance is crucial for the effects of prompts, and, that this compliance can be increased when students become more involved (e.g., they can choose the best applicable prompts in their situation) in the design of the prompts (Bannert et al., 2015). We cannot rule out that this could have affected our study results, as our students all received the same prompts at the same time. Besides, in our subsample of students who completed most of the videos, we can only ascertain that students at least saw the SRL prompts and responded to them, but we did not have control over the extent to which they really applied the SRL prompts in their learning process.

Second, it could be the case that the effect of SRL prompts is dependent on the type of learning outcome that is measured. Previous SRL prompt research, where a clear distinction is made between recall, comprehension, and transfer (i.e., applying the learning material to new contexts), usually find significant differences in learning outcomes for transfer (Bannert et al., 2015, 2009; Müller & Seufert, 2018). In our present study, we made use of an ecologically valid and generic test, in which this categorically difference is hard to make for individual items.

4.3. Did performance level moderate the effect of the SRL prompts?

Our last research question was if students' performance level (i.e., a students' average score for History that current schoolyear) moderated the effects of SRL prompts. Although we matched students according to their individual performance level to make comparable research conditions, we hypothesized that students with different performance levels within the research conditions could experience differential benefits of SRL prompts. In contrast to previous research (Sitzmann et al., 2009; Wong et al., 2019; Yeh et al., 2010), we did not find that the SRL prompts had different effects for students with different performance levels in our context. It could be possible, however, that other measurements (e.g., intelligence, pre-knowledge) are more suitable to investigate individual differences on cognitive abilities. For example, research has found that students with higher prior knowledge perform more online SRL activities than students with lower prior knowledge (Taub, Azevedo, Bouchet, & Khosravifar, 2014).

However, we did find differential effects for the different educational levels on SRL self-reports and learning outcomes. While performance level indicates an individual student's performance in History, educational level is a more generic indication of a student's ability. We found that students in the pre-university level scored higher on the help seeking and time-management scale, on the learning outcome test, and were more satisfied with the flipped environment than students in the senior general level. We also found that students in the senior general level on average watched 55% homework videos compared to 80% of the videos watched by pre-university students. The findings are in line with research that shows that SRL and motivation are important predictors of school achievement for the current group of students (Vukman

& Licardo, 2010; Pintrich & De Groot, 1990). It shows that students in the highest educational level not only perform better in the learning test because of higher cognitive abilities, but it also explains the higher SRL self-report scores and the higher persistency and motivation to complete the homework videos. Furthermore, as these higher achievers have more SRL skills, they are likely to be more satisfied in a flipped classroom that makes greater use of their ability to self-regulate their learning. We recommend to take this difference in educational level into account in future research on SRL prompts in a flipped classroom setting.

4.4. Strengths and limitations: implications for theory and practice

First, an important difference between our study and previous research is the educational context (student level and age). Previous research in primary and higher education showed positive effects of SRL prompts in (FTC) learning environments, but we did not find such effects in secondary education. We know that students' age is an import factor in the development of SRL skills (Wigfield et al., 2011; Zimmerman & Martinez-Pons, 1990). It could be possible that some of the younger students lack certain metacognitive skills, or do not have the motivation to apply them. It could be possible that additional SRL strategy instruction is needed for these students in secondary education to effectively work with the SRL prompts. For example, nine students from the SRL prompt condition expressed negative comments about the prompts when asked about disadvantages in the satisfaction questionnaire. In the focus group interviews, some students from the SRL prompts condition explained that they sometimes felt that these prompts distracted them from learning the content and that they do not see the added value of this activity. It could be that they are not aware that they are actually applying SRL methods, do not fully understand the connection with the prompts, and sometimes even see them as additional (and distracting) workload. The fact that the interviewed students are clearly positive about the content questions accompanying the videos in their workbook, and a large group of students in general rewound video parts relatively often, illustrates that students did put effort in monitoring their learning. This implies that making the *why* of the SRL prompts explicit to students (e.g., by adding instruction or training) could be an important point of improvement for future research (Jansen, van Leeuwen, Janssen, Conijn, & Kester, 2020). The introduction about working in a flipped classroom in the current intervention could, for instance, be supplemented with an in-class training on SRL skills to make sure the students have a basic level and understanding of SRL skills in this particular setting (Bannert et al., 2009; Bol, Campbell, Perez, & Yen, 2016; Kistner et al., 2010).

Second, the analysis of students' satisfaction showed that there was a difference for educational level on the satisfaction questionnaire, as students in the higher educational level valued the flipped classroom learning environment more. The average satisfaction of the students from the senior general level, however, was still acceptable (with a mean score of 29 out of 49). Given the positive results of the satisfaction questionnaire and the focus group interviews, we conclude that students' general satisfaction with the flipped classroom have not been the cause of the absence of effects of the SRL prompts.

Third, our methodological design contains both strengths and weaknesses that could have affected the results. In contrast to previous research, we developed an intervention of six weeks in an ecologically valid classroom setting, to further test effects that were found in short term research designs. The advantage of randomization within groups was a likely elimination of between group differences and confounders (e.g., teacher, classroom, educational level, and other circumstantial effects). A possible limitation of this is that we had less control over the contact participants from different research conditions possibly had with each other. However, this can never be fully prevented in quasi-experiments of multiple weeks in an ecologically valid environment. Ideally, future research will also include designs that focus on improving students' SRL during a semester or school year to evaluate longer term effects.

In addition, measurement type of SRL could also have been a factor. By measuring SRL with pretest and posttest questionnaires and by collecting online expressions of SRL activities, we followed recent recommendation to measure different levels of SRL and including measurements that focus on both the quality and quantity of SRL (Dent & Koenka, 2016; Panadero et al., 2016; Rovers et al., 2019). However, a recent meta-analysis found that effects of interventions on SRL depend on the measurement type, as effects were stronger when SRL activities were measured with a quantitative measurements (e.g., Moos & Bonde, 2016), compared to questionnaires (Jansen, van Leeuwen, Janssen, Jak, & Kester, 2019).

Fourth, while we used power analysis to calculate the necessary sample size beforehand to find effect sizes comparable to previous research, we lost some power due to non-compliance of 23% of the students. We decided to exclude non-compliance students who watched less than five (of seven) videos (from both conditions), because it means that students in the SRL prompt condition were not sufficiently exposed to the SRL prompts and are in fact more comparable to students in the control group. Therefore, conclusions on the non-significant results should be interpreted with caution. As we are aware that this could introduce a selection bias to our results, we also provided the results of the analyses on all the students in Appendix B. It shows that the results from the analyses on all the students do not lead to substantially different conclusions.

Fifth, the online SRL measurement *video watch time* is our best indication of time-on-task. It should be noted, however, that this measurement only measured actual playtime of the video: when students paused the video, the *video watch time* did not increase. We could also not measure the time that students spent on learning and applying SRL strategies when they learned material at home before and after the video. The average SRL-prompt compliance rate of 90% indicates that students in the SRL condition generally followed the instruction. The actual rate could be slightly higher or lower, as it does not mean that when a student did not provide a valid answer (e.g., blank) on a particular prompt, the student did do something outside of the online learning environment with the SRL instruction (and vice versa).

4.5. Conclusion

When curriculum developers/teachers decide to apply FTC, they should ask the question if it is worth to invest in the support of SRL. General consensus from previous research shows that supporting SRL could be an important addition or even precondition for a FTC. One effective way to do this in higher education is embedding SRL prompts in instructional videos. Our study shows that this is not directly applicable in the educational context of a History course in secondary education. Our findings indicate that the current flipped intervention in general stimulates students to perform SRL activities while learning from the videos, and we found an effect of the SRL prompts on the amount of videos the students completed. Future research should investigate how SRL can be optimally supported in flipped secondary education. The results from the small scale focus group interviews suggest that making explicitly clear to students why SRL prompts are useful for their learning could be an important addition to improve the effect of SRL prompts.

Credit author statement

David C. D. van Alten: Conceptualization, Methodology, Investigation, Resources, Writing - Original Draft, Visualization, Chris Phielix: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Jeroen Janssen: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Liesbeth Kester: Conceptualization, Methodology, Writing - Review & Editing, Supervision.

Acknowledgments

The conduct of this research was supported by funding from the Dutch Ministry of Education, Culture and Science, (grant number OCW/PromoDoc/1065001). Declarations of interest: none.

Appendix A

Overview of all the video embedded SRL prompts for each video (knowledge clip, KC) and each SRL phase. Hints (in italics) were provided after students answered a that particular SRL prompt.

KC	**Forethought prompts** (start of the knowledge clip)	**Performance prompts** (middle of the knowledge clip)	**Self-reflection prompts** (end of knowledge clip)	
1	What do you think will be the best way to learn from this knowledge clip? *Hints: You can think of answering the quizzes and making notes in your workbook. And how will you make sure you are concentrated to learn?*	Which information have you learned up until now about the agricultural revolution?	(1) Did your approach to learn from this knowledge clip work for you? Or is it necessary to make some changes in your strategy? *Hint: Do you use your workbook to make notes? Do you eliminate distraction when you learn?*	(2) Do you understand the learning content from the video? If not, will you watch certain parts again?
2	What do you already know about the emergence of factories?	Is there any information so far that you did not understand? *Hint: to help you, you could rewind difficult parts of the video, or read paragraph 5.2 in your textbook as aid.*	What have you learned about the emergence of factories?	
3	Answering the quizzes in your workbook helps you to understand the learning content. How do you make sure you will also understand the content outside the quizzes? *Hint: Make notes, pause regularly and summarize the content for yourself, keep asking yourself questions if you understand the learning content.*	What have you learned about steam power? What can you do if you experience difficulties?	If you think about the way you are learning from the videos right now, do you need to adjust your strategies? *Hint: Make your current learning strategy explicit for yourself. What else could work good for you? How can you know if it works for you?*	
4	Before you start this knowledge clip, do you already have questions about the working conditions of the working class in the factories? *Hint: Try to ask yourself 2–3 questions and put them here. For example: are the working conditions the same in the 17*th *and 19*th *century?*	Is there information about the working environments that you did not understand? If yes, what could you do now? *Hint: Do you ask questions to someone nearby, or a classmate, or your teacher? Are there other options for help that you approach (for example your textbook)?*	Is it clear for you what the most important learning content was in this video? *Hint: Do you look at the heading titles in the video to understand this? Or do you use the quizzes in your workbook to evaluate that?*	
5	(1) Do you set goals for yourself that would help you to learn the learning content in this video? Why? *Hints: for example, you could set goals like:* *-I want to be able to explain difficult concepts in my own words* *-I want to answer the quizzes correctly* *-I want to spend 15 min to my history home work.* (2) This knowledge clip is about urbanization: cities that will grow larger and larger. Before you watch the video, do you already have ideas what this has to do with factories?	Are you able to learn from this video in an effective way? How is your concentration, do you experience distraction? Do you work on a quiet place? If not, what can you change?	Do you know why it is more effective to answer the quizzes thoughtfully, and to make notes in your workbook?	
6	*Hint: this video is about the question who will stick up for the working class. Why is this necessary if you think about the previous videos?*	The next section of this video will be about three different political views. It is important to know the differences and similarities. How will you make sure to understand and remember them?	Did you understand all the information about the social issue? If not, will you rewind the most important parts?	
7	Think about two goals that you want to accomplish by watching this knowledge clip. For example about your planning, or the way you work.	–	Did you accomplish the goals you have set at the beginning of this knowledge clip? If not, what can you still do to reach them?	

Appendix B

Presentation of the analyses on data from all the students (including non-compliance). In Table B1, we present the descriptive statistics for all the dependent variables.

B.1. Effects of SRL-prompts

B.1.1. SRL self-reports

A MANOVA that included condition (SRL prompts versus no SRL prompts) and educational level (senior general versus pre-university) as factors and the seven SOL-Q-R scales as dependent variables was used to investigate if students who received SRL prompts reported higher SRL self-reports than students who did not receive SRL prompts (RQ1a). We found no main effect of condition on the SRL self-reports scales, $F(7, 132) = 0.44$, $p = .879$; Wilk's $\Lambda = 0.977$, partial $\eta^2 = 0.02$, but we did find a main effect of educational level on the SRL self-reports scales, $F(7, 132) = 3.25$, $p = .003$; Wilk's

$\Lambda = 0.853$, partial $\eta^2 = 0.15$. Follow up ANOVAs showed that the pre-university level ($M = 4.73$, $SD = 1.11$, $n = 104$) scored significantly higher ($p = .005$, partial $\eta^2 = 0.06$) on the *time management* scale than the senior general level ($M = 4.16$, $SD = 0.84$, $n = 38$).

B.1.2. SRL online activity

We dropped *video watch time* and *portions viewed* from the analysis because of very high correlations with *rewind actions* and the problem of multicollinearity. A MANOVA with condition and educational level (pre-university, senior general) as factors and the *completion rate* and *rewind actions* as online SRL activity variables as dependent variables was used to investigate if students who received SRL prompts performed more online SRL activities than students who did not receive SRL prompts (RQ1b). We found no significant main effect for condition on online SRL activity, $F(2, 148) = 0.18$, $p = .837$; Wilk's $\Lambda = 0.998$, partial $\eta 2 < 0.01$. We found a significant main effect for educational level on online SRL activity, $F(2, 148) = 9.42$, $p = < .001$; Wilk's $\Lambda = 0.887$, partial $\eta^2 = 0.11$. Follow up ANOVAs showed a significant difference in means ($p < .001$, partial $\eta^2 = 0.11$) of the online SRL activity variable *completion rate* between educational levels: students in the senior general level on average watched 55% of the videos ($SD = 30.49$, $n = 44$) and students in pre-university on average 80% ($SD = 30.49$, $n = 109$).

B.1.3. Learning outcomes

An ANOVA on the final recall and comprehension test with condition and educational level as factors was used to investigate if students who received SRL prompts obtained higher learning outcomes than students who did not receive SRL prompts (RQ2). We found no significant effect of condition on learning outcome; $F(1, 148) = 0.45$, $p = .505$, partial $\eta^2 = < 0.01$, but we did find a significant effect of educational level on learning outcome, $F(1, 148) = 21.30$, $p = <.001$, partial $\eta^2 = < 0.13$.

Table B.1
Mean and Standard Deviation for the Dependent Variables (all Students Included)

Dependent variables (posttest)[a]	**SRL prompts condition**			**No SRL prompts condition**		
	n	M	SD	n	M	SD
SRL self-report scales ($p = .879$)						
Metacognitive act. before	71	3.50	1.18	74	3.50	1.15
Metacognitive act. during	71	3.58	1.17	75	3.63	1.11
Metacognitive act. after	71	2.99	1.21	74	2.96	1.17
Persistence	70	4.55	1.31	75	4.59	1.29
Help seeking	70	3.84	1.14	75	3.53	1.14
Environmental structuring	71	5.26	1.43	74	5.36	1.34
Time management	71	4.55	1.10	75	4.59	1.04
SRL online activity ($p = .837$)						
Video completion rate	73	73.89	36.43	80	72.25	34.06
Video watch time	73	52.40	32.00	80	52.13	30.03
Video portions viewed	73	70.34	43.35	80	65.91	39.02
Rewind actions	73	18.51	25.63	80	15.26	22.71
Learning outcomes ($p = .505$)						
Learning outcome test	74	15.52	4.22	78	15.03	4.11

[a] *Note*: Provided *p*-values are from the comparison of SRL-prompts condition in the separate MANOVA on *SRL self-report scales*; MANOVA on *SRL online activity*, which included video completion rate and rewind actions; ANOVA on *learning outcomes*.

B.2. Moderation of the effects of SRL-prompts by performance level

We performed two separate MANCOVAs and one ANCOVA with performance level (i.e., the students' average scores for History that current schoolyear) as a covariate, condition type and educational level as factors, and the SRL questionnaire scales, the SRL online activity variables, and the learning outcome test as dependent variables to investigate if performance level moderated the effects of SRL prompts (RQ3). We found no significant interaction effects between performance level and condition for the SRL self-reports, $F(7, 131) = 1.27$, $p = .269$; Wilk's $\Lambda = 0.936$, partial $\eta^2 = 0.06$; no significant interaction effects between performance level and condition for the SRL online activities, $F(2, 147) = 0.44$, $p = .645$; Wilk's $\Lambda = 0.994$, partial $\eta^2 = 0.01$; and no significant interaction effects between performance level and condition for learning outcome; $F(1, 147) = 0.52$, $p = .470$, partial $\eta^2 = < 0.01$. This shows that the effect of the SRL prompts is not different for students' performance level.

Appendix C

Table C.1
Pearson Correlation Matrix for all the Dependent Variables (N between 102 and 105, only students who watched at least five out of seven videos)

Dependent variable	MAb	MAd	MAa	PER	HS	ES	TM	VCR	VWT	VPV	RA
SRL self-report scales											
Metacognitive act. before (MAb)	–										
Metacognitive act. during (MAd)	.715**	–									
Metacognitive act.after (MAa)	.768**	.799**	–								
Persistence (PER)	.357**	.382**	.350**	–							
Help seeking (HS)	.432**	.485**	.362**	.160	–						
Environmental structuring (ES)	.200*	.413**	.198*	.299**	.298**	–					
Time management (TM)	.248*	.334**	.294**	.684**	.281**	.422**	–				
SRL online activities											
Video completion rate (VCR)	.070	.048	.076	.183	.014	-.081	.151	–			
Video watch time (VWT)	.021	.027	.073	.088	.050	-.108	.053	.481**	–		

(*continued on next page*)

Table C.1 (*continued*)

Dependent variable	MAb	MAd	MAa	PER	HS	ES	TM	VCR	VWT	VPV	RA
Video portions viewed (VPV)	.026	.035	.085	.086	.078	-.110	.076	.496**	.987**	–	
Rewind actions (RA)	.014	.027	.075	.053	.083	-.101	.048	.306**	.968**	.978**	–
Learning outcomes											
Learning outcome test (LO)	.176	.164	.156	.328**	.104	.209*	.374**	.303**	.217*	.227*	.175

**p < .01 (1-tailed).
*p < .05 (1-tailed).

Table C.2
Pearson Correlation Matrix for Performance Level and Dependent Variables (N between 102 and 105, only students who watched at least five out of seven videos)

Dependent variable	Performance level
Metacognitive activities before	.180*
Metacognitive activities during	.056
Metacognitive activities after	.185*
Persistence	.391**
Help seeking	-.008
Environmental structuring	.169*
Time management	.357**
Video completion rate	.074
Video watch time	.069
Video portions viewed	.071
Rewind actions	.060
Learning outcome test	.556**

**p < .01 (1-tailed).
*p < .05 (1-tailed).

References

van Alten, D. C. D., Phielix, C., Janssen, J., & Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educational Research Review, 28*, 1–18. https://doi.org/10.1016/j.edurev.2019.05.003.

Armijo-Olivo, S., Warren, S., & Magee, D. (2009). Intention to treat analysis, compliance, drop-outs and how to deal with missing data in clinical research: A review. *Physical Therapy Reviews, 14*(1), 36–49. https://doi.org/10.1179/174328809X405928.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*(3), 523–535. https://doi.org/10.1037/0022-0663.96.3.523.

Azevedo, R., & Hadwin, A. F. (2005). Scaffolding self-regulated learning and metacognition - implications for the design of computer-based scaffolds. *Instructional Science, 33*(5–6), 367–379. https://doi.org/10.1007/s11251-005-1272-9.

Bannert, M., Hildebrand, M., & Mengelkamp, C. (2009). Effects of a metacognitive support device in learning environments. *Computers in Human Behavior, 25*(4), 829–835. https://doi.org/10.1016/j.chb.2008.07.002.

Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning, 9*(2), 161–185. https://doi.org/10.1007/s11409-013-9107-6.

Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior, 52*, 293–306. https://doi.org/10.1016/j.chb.2015.05.038.

Bol, L., Campbell, K. D. Y., Perez, T., & Yen, C. J. (2016). The effects of self-regulated learning training on community college students' metacognition and achievement in developmental math courses. *Community College Journal of Research and Practice, 40*(6), 480–495. https://doi.org/10.1080/10668926.2015.1068718.

Bryan, T., Burstein, K., & Bryan, J. (2001). Students with learning disabilities: Homework problems and promising practices. *Educational Psychologist, 36*(3), 167–180. https://doi.org/10.1207/S15326985EP3603_3.

Butzler, K. B. (2016). The synergistic effects of self-regulation tools and the flipped classroom. *Computers in the Schools, 33*(1), 11–23. https://doi.org/10.1080/07380569.2016.1137179.

Cheng, L., Ritzhaupt, A. D., & Antonenko, P. (2018). *Effects of the flipped classroom instructional strategy on students' learning outcomes: A meta-analysis*. Educational Technology Research and Development. https://doi.org/10.1007/s11423-018-9633-7.

Chen, K. S., Monrouxe, L., Lu, Y. H., Jenq, C. C., Chang, Y. J., Chang, Y. C., et al. (2018). Academic outcomes of flipped classroom learning: A meta-analysis. *Medical Education, 52*(9), 910–924. https://doi.org/10.1111/medu.13616.

Cicchinelli, A., Veas, E., Pardo, A., Pammer-Schindler, V., Fessl, A., Barreiros, C., et al. (2018). Finding traces of self-regulated learning in activity streams. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge - LAK, '18*, 191–200. https://doi.org/10.1145/3170358.3170381.

Clark, R. C., Nguyen, F., & Sweller, J. (2011). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. John Wiley & Sons.

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review, 28*(3), 425–474. https://doi.org/10.1007/s10648-015-9320-8.

Devolder, A., van Braak, J., & Tondeur, J. (2012). Supporting self-regulated learning in computer-based learning environments: Systematic review of effects of scaffolding in the domain of science education. *Journal of Computer Assisted Learning, 28*(6), 557–573. https://doi.org/10.1111/j.1365-2729.2011.00476.x.

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning, 3*(3), 231–264. https://doi.org/10.1007/s11409-008-9029-x.

Donker, A. S., de Boer, H., Kostons, D., Dignath van Ewijk, C. C., & van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review, 11*, 1–26. https://doi.org/10.1016/j.edurev.2013.11.002.

Fergusson, D., Aaron, S. D., Guyatt, G., & Hébert, P. (2002). Post-randomisation exclusions: The intention to treat principle and excluding patients from analysis. *British Medical Journal, 325*(7365), 652–654.

Field, A. P., & Wilcox, R. R. (2017). Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers. *Behaviour Research and Therapy, 98*, 19–38. https://doi.org/10.1016/j.brat.2017.05.013.

Greene, J. A., Bolick, C. M., Caprino, A. M., Deekens, V. M., McVea, M., Yu, S., et al. (2015). Fostering high-school students' self-regulated learning online and across academic domains. *High School Journal, 99*(1), 88–106. https://doi.org/10.1353/hsj.2015.0019.

He, W., Holton, A., Farkas, G., & Warschauer, M. (2016). The effects of flipped instruction on out-of-class study time, exam performance, and student perceptions. *Learning and Instruction, 45*, 61–71. https://doi.org/10.1016/j.learninstruc.2016.07.001.

Hsu, H. C. K., Wang, C. V., & Levesque-Bristol, C. (2019). Reexamining the impact of self-determination theory on learning outcomes in the online learning environment. *Education and Information Technologies, 24*(3), 2159–2174. https://doi.org/10.1007/s10639-019-09863-w.

Jansen, R. S., van Leeuwen, A., Janssen, J., Conijn, R., & Kester, L. (2020). Supporting learners' self-regulated learning in Massive Open Online Courses. *Computers and Education, 146*, 103771. https://doi.org/10.1016/j.compedu.2019.103771.

Jansen, R. S., van Leeuwen, A., Janssen, J., Jak, S., & Kester, L. (2019). Self-regulated learning partially mediates the effect of self- regulated learning interventions on achievement in higher education : A meta-analysis. *Educational Research Review, 28*. https://doi.org/10.1016/j.edurev.2019.100292.

Jansen, R. S., van Leeuwen, A., Janssen, J., & Kester, L. (2018). Validation of the revised self-regulated online learning questionnaire. In V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachsler, R. Elferink, & M. Scheffel (Eds.), *Lifelong technology-enhanced learning* (pp. 116–121). Cham: Springer International Publishing.

Jansen, R. S., van Leeuwen, A., Janssen, J., Kester, L., & Kalz, M. (2017). Validation of the self-regulated online learning questionnaire. *Journal of Computing in Higher Education, 29*(1), 6–27. https://doi.org/10.1007/s12528-016-9125-x.

Jo, B. (2002). Statistical power in randomized intervention studies with noncompliance. *Psychological Methods, 7*(2), 178–193. https://doi.org/10.1037/1082-989X.7.2.178.

Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *Internet and Higher Education, 33*, 74–85. https://doi.org/10.1016/j.iheduc.2017.02.001.

Kauffman, D. F. (2004). Self-regulated learning in web-based environments: Instructional tools designed to facilitate cognitive strategy use, metacognitive processing, and motivational beliefs. *Journal of Educational Computing Research, 30*(1–2), 139–161. https://doi.org/10.2190/AX2D-Y9VM-V7PX-0TAD.

Kistner, S., Rakoczy, K., Otto, B., Dignath-van Ewijk, C., Büttner, G., & Klieme, E. (2010). Promotion of self-regulated learning in classrooms: Investigating frequency, quality, and consequences for student performance. *Metacognition and Learning, 5*(2), 157–171. https://doi.org/10.1007/s11409-010-9055-3.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education, 104*, 18–33. https://doi.org/10.1016/j.compedu.2016.10.001.

Låg, T., & Sæle, R. G. (2019). Does the flipped classroom improve student learning and satisfaction? A systematic review and meta-analysis. *AERA Open, 5*(3). https://doi.org/10.1177/2332858419870489, 233285841987048.

Lai, C.-L., & Hwang, G.-J. (2016). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers & Education, 100*, 126–140. https://doi.org/10.1016/j.compedu.2016.05.006.

Lape, N. K., Levy, R., Yong, D. H., Haushalter, K. A., Eddy, R., & Hankel, N. (2014). Probing the inverted classroom: A controlled study of teaching and learning outcomes in undergraduate engineering and mathematics (9475. In *121st ASEE annual conference & exposition*. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-84905192244&partnerID=40&md5=f30d884fb1086d6e5dfcb8d18f9ef67e.

Lee, J., & Choi, H. (2019). Rethinking the flipped learning pre-class: Its influence on the success of flipped learning and related factors. *British Journal of Educational Technology, 50*(2), 934–945. https://doi.org/10.1111/bjet.12618.

Lee, H. W., Lim, K. Y., & Grabowski, B. L. (2010). Improving self-regulation, learning strategy use, and achievement with metacognitive feedback. *Educational Technology Research & Development, 58*(6), 629–648. https://doi.org/10.1007/s11423-010-9153-6.

Lee, S. W. Y., & Tsai, C. C. (2011). Students' perceptions of collaboration, self-regulated learning, and information seeking in the context of Internet-based learning and traditional learning. *Computers in Human Behavior, 27*(2), 905–914. https://doi.org/10.1016/j.chb.2010.11.016.

Lim, D. H., & Morris, M. L. (2009). Learner and instructional factors influencing learning outcomes within a blended learning environment. *Educational Technology & Society, 12*(4), 282–293. https://doi.org/10.2307/jeductechsoci.12.4.282.

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Munoz-Gama, J. (2018). Mining theory-based patterns from big data: Identifying self-regulated learning strategies in massive open online courses. *Computers in Human Behavior, 80*, 179–196. https://doi.org/10.1016/j.chb.2017.11.011.

Moos, D. C., & Bonde, C. (2016). Flipping the classroom: Embedding self-regulated learning prompts in videos. *Technology, Knowledge and Learning*, 1–18. https://doi.org/10.1007/s10758-015-9269-1.

Müller, N. M., & Seufert, T. (2018). Effects of self-regulation prompts in hypermedia learning on learning performance and self-efficacy. *Learning and Instruction, 58* (April), 1–11. https://doi.org/10.1016/j.learninstruc.2018.04.011.

Nye, B., Spyros, K., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237–257. https://doi.org/10.3102/01623737026003237.

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*, 1–28. https://doi.org/10.3389/fpsyg.2017.00422.

Panadero, E., Klug, J., & Järvelä, S. (2016). Third wave of measurement in the self-regulated learning field: When measurement and intervention come hand in hand. *Scandinavian Journal of Educational Research, 60*(6), 723–735. https://doi.org/10.1080/00313831.2015.1066436.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom Academic performance. *Journal of Educational Psychology, 82*(1), 33–40.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. https://doi.org/10.5901/mjss.2015.v6n1p156. Ann Arbor, MI.

Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: A review. *Scandinavian Journal of Educational Research, 45*(3), 269–286. https://doi.org/10.1080/00313830120074206.

Ramdass, D., & Zimmerman, B. J. (2011). Developing self-regulation skills: The important role of homework. *Journal of Advanced Academics, 22*(2), 194–218. https://doi.org/10.1177/1932202X1102200202.

Rovers, S. F. E., Clarebout, G., Savelberg, H. H. C. M., de Bruin, A. B. H., & van Merriënboer, J. J. G. (2019). Granularity matters: Comparing different ways of measuring self-regulated learning. *Metacognition and Learning, 14*(1), 1–19. https://doi.org/10.1007/s11409-019-09188-6.

Shibukawa, S., & Taguchi, M. (2019). Exploring the difficulty on students' preparation and the effective instruction in the flipped classroom. *Journal of Computing in Higher Education, 31*(2), 311–339. https://doi.org/10.1007/s12528-019-09220-3.

Shyr, W. J., & Chen, C. H. (2018). Designing a technology-enhanced flipped learning system to facilitate students' self-regulation and performance. *Journal of Computer Assisted Learning, 34*(1), 53–62. https://doi.org/10.1111/jcal.12213.

Sitzmann, T., Bell, B. S., Kraiger, K., & Kanar, A. M. (2009). A multilevel analysis of the effect of prompting self-regulation in technology-delivered instruction. *Personnel Psychology, 62*, 697–734.

Sitzmann, T., & Ely, K. (2010). Sometimes you need a reminder: The effects of prompting self-regulation on regulatory processes, learning, and attrition. *Journal of Applied Psychology, 95*(1), 132–144. https://doi.org/10.1037/a0018080.

Sletten, S. R. (2017). Investigating flipped learning: Student self-regulated learning, perceptions, and achievement in an introductory biology course. *Journal of Science Education and Technology, 26*(3), 347–358. https://doi.org/10.1007/s10956-016-9683-8.

Sun, J. C., Wu, Y., & Lee, W. (2016). The effect of the flipped classroom approach to OpenCourseWare instruction on students' self-regulation. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.12444.

Tan, C., Yue, W.-G., & Fu, Y. (2017). Effectiveness of flipped classrooms in nursing education: Systematic review and meta-analysis. *Chinese Nursing Research, 4*(4), 192–200. https://doi.org/10.1016/J.CNRE.2017.10.006.

Taub, M., Azevedo, R., Bouchet, F., & Khosravifar, B. (2014). Can the use of cognitive and metacognitive self-regulated learning strategies be predicted by learners' levels of prior knowledge in hypermedia-learning environments? *Computers in Human Behavior, 39*, 356–367. https://doi.org/10.1016/j.chb.2014.07.018.

Vukman, K. B., & Licardo, M. (2010). How cognitive, metacognitive, motivational and emotional self-regulation influence school performance in adolescence and early adulthood. *Educational Studies, 36*(3), 259–268. https://doi.org/10.1080/03055690903180376.

Wigfield, A., Klauda, S. L., & Cambria, J. (2011). Influences on the development of academic self-regulatory processes. In D. H. S. Barry, & J. Zimmerman (Eds.), *Handbook of self-regulation of learning and performance* (pp. 33–48). New York: Routledge. https://doi.org/10.4324/9780203839010.ch3.

Winne, P. H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition and Learning, 9*(2), 229–237. https://doi.org/10.1007/s11409-014-9113-3.

Wolters, C. A., Pintrich, P. R., & Karabenick, S. A. (2005). Assessing academic self-regulated learning. In K. A. Moore, & L. H. Lippman (Eds.), *What do children need to flourish: Conceptualizing and measuring indica- tors of positive development*. Berlin, New York, NY: Springer Science Business Media. https://doi.org/10.1007/0-387-23823-9.

Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G. J., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. *International Journal of Human-Computer Interaction, 35*(4–5), 356–373. https://doi.org/10.1080/10447318.2018.1543084.

Yeh, Y. F., Chen, M. C., Hung, P. H., & Hwang, G. J. (2010). Optimal self-explanation prompt design in dynamic multi-representational learning environments. *Computers & Education, 54*(4), 1089–1100. https://doi.org/10.1016/j.compedu.2009.10.013.

Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review, 17*(2), 187–202. https://doi.org/10.1007/s12564-016-9426-9.

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*(1), 51–59. https://doi.org/10.1037/0022-0663.82.1.51.

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In A. C. Hacker D.J., Dunlosky J., Graesser (Ed.), Handbook of metacognition in education (pp. 299–315). New York, NY: Routledge..