frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 28 August 2024
DOI 10.3389/feduc.2024.1412683

OPEN ACCESS

EDITED BY
David Rodriguez-Gomez,
Universitat Autònoma de Barcelona, Spain

REVIEWED BY
Mojtaba Rajabi,
Gonbad Kavous University, Iran
Wahyu Widada,
University of Bengkulu, Indonesia
Nuno Ricardo Oliveira,
Institute of Education and Sciences (ISEC Lisboa), Portugal

*CORRESPONDENCE:
Gesa Bintz
✉ gesa.bintz@uni-muenster.de

RECEIVED 05 April 2024
ACCEPTED 07 August 2024
PUBLISHED 28 August 2024

CITATION
Bintz G, Barenberg J and Dutke S (2024) Components of the flipped classroom in higher education: disentangling flipping and enrichment.
Front. Educ. 9:1412683.
doi: 10.3389/feduc.2024.1412683

Components of the flipped classroom in higher education: disentangling flipping and enrichment

Gesa Bintz*, Jonathan Barenberg and Stephan Dutke

Institute for Psychology in Education, University of Münster, Münster, Germany

Introduction: The flipped classroom (FC) model shifts the initial presentation of new content from in-class to out-of-class, while in-class time focuses on elaborating previously presented content. Although FC's benefits on learning outcomes are well-documented, the mediating processes remain unclear due to the simultaneous enrichment of learning activities in many studies. This study investigates whether merely flipping the initial presentation of new content, without additional enrichment, enhances learning outcomes.

Method: In two studies, psychology students (Study 1: $N = 306$; Study 2: $N = 413$) participated in either a non-flipped lecture series (new content presented in class) or a flipped lecture series (new content presented out-of-class via lecture recordings). Learning outcomes were assessed in terms of knowledge acquisition, knowledge application, and metacognitive monitoring.

Results: Results showed that students in the flipped classroom condition had significantly higher learning outcomes compared to those in the non-flipped condition. Specifically, the flipped approach improved knowledge acquisition, knowledge application, and metacognitive monitoring.

Discussion: These findings indicate that simply shifting the initial presentation of new content to an out-of-class setting is sufficient to enhance learning outcomes in a flipped classroom. This study highlights the effectiveness of the flipped classroom model in fostering student learning without the need for additional enrichment activities, suggesting a structured approach to analyze the components that contribute to learning benefit from flipping the classroom.

KEYWORDS

flipped classroom, lecture recording, metacognitive monitoring, higher education, learning outcome

1 Introduction

In recent years, flipped classroom is frequently discussed as an innovative method to foster instruction. This approach suggests reconsidering which learning processes are prompted in class and which out of class. In traditional higher education, courses usually follow this sequence: (a) Instructors synchronously present learning content in a classroom setting and (b) learners organize and elaborate the presented learning content out of class in self-study. Flipped classroom approaches invert this pattern (Lage et al., 2000). What was traditionally presented in class, is now acquired out of class; what was traditionally elaborated individually out of class is now subject of in-class processes. With accumulating research, it became evident that, in most implementations of flipped classroom, activities were not simply switched between in and out of class (e.g., Bishop and Verleger, 2013; Kapur et al., 2022), but flipped classroom represents "an expansion of the curriculum, rather than a mere re-arrangement of activities" (Bishop and Verleger, 2013, p. 4). Flipping a traditional teaching concept may be a chance for practitioners to create new learning opportunities and thereby to enhance learning. From a research perspective, however, this approach complicates investigating the extent that beneficial effects of flipped classroom can be attributed to flipping the setting of learning or to enriching learning activities. The present work aims at contributing to attempts of disentangling the impact of these two factors by examining traditional and flipped university lectures.

2 Components of flipped classroom

In the following, we denominate the learners' first contact with new learning material as primary processing of learning content, in traditional university lectures often a teacher-centered presentation. In addition to the presentation, primary processing may also include first questions to instructors or peers in order to initially understand new learning content. Secondary processing includes more active processing and elaboration of learning content by learners, in traditional lectures usually realized by individual or group work after and out of class. This simplified description may also apply to flipped classroom, with the difference that primary processing takes place before and out of class, whereas secondary processing occurs in class. If the key to flipped classroom approaches laid entirely in the "flipping" of where these components take place (in or out of class), experimental and control groups in flipped classroom research should work on the same tasks. Otherwise, it is impossible to differentiate the extent that increased learning outcomes are due to flipping the setting of primary and secondary processing (in or out of class) or due to qualitative changes in primary and secondary processing. This ambiguity is aggravated as researchers and instructors in secondary and tertiary education have contributed to a growing but heterogenous pool of studies on flipped classroom, and thereby generated a large pool of additional learning methods and opportunities suitable to enrich primary and secondary processing.

2.1 Challenge for research: heterogeneity and enrichment

Flipped classroom research has grown exponentially within the last decade (Talbert, 2022), resulting in a great body of articles, as well as reviews and meta-analyses with various foci (O'Flaherty and Phillips, 2015; Zainuddin and Halili, 2016; Låg and Sæle, 2019; Zheng et al., 2020). Only few studies, however, were dedicated to explore the reasons why and how flipped classroom affects learning (Abeysekera and Dawson, 2014; Kapur et al., 2022). There are several interesting, yet vague ideas why flipped classroom might enhance learning (Abeysekera and Dawson, 2014; DeLozier and Rhodes, 2017) although they do not provide conclusive explanations of flipped classroom effects on learning success (Låg and Sæle, 2019). Partially, this situation might be due to the fact that most reported implementations of flipped classroom change the setting of primary and secondary processing *and* add various activities to primary and secondary processing. Thus, many implementations of flipped classroom entail enrichment of primary and secondary processing.

McLaughlin et al. (2014), for example, flipped a first-year pharmaceutics course. Primary processing in the traditional lecture consisted of assigned reading (out of class) and instructor-centered lecture (in class). The lecture included occasional quizzes, interactive tasks as well as opportunities to ask questions. In the flipped condition, primary processing included the same reading assignments plus self-paced interactive learning modules. Secondary processing comprised various tasks (regularly answering clicker questions, interactive tasks, quizzes and micro-lectures). Additionally, students gave presentations, held discussions in class and worked on course projects throughout the semester. This is a typical example of flipped classroom implementations with various additional active and interactive learning opportunities in the flipped condition (Kapur et al., 2022). Primary processing was enriched through interactive learning modules and secondary processing via more frequent and additional elaboration tasks such as application of learning content in a course project. Although, in this case, the flipped condition led to a significant increase in exam grades compared to the traditional condition (McLaughlin et al., 2014), the reason for this increase remains unclear.

Giving a rather different example, Burgoyne and Eaton (2018) implemented flipped classroom in a course on research methods in social sciences. In the non-flipped condition, the instructor presented learning content with slides for primary processing in class. For secondary processing, students were assigned critical thinking tasks as homework. Students in the flipped condition received recordings of the same slides presented by the same instructor via their online learning platform. Secondary processing included the same critical thinking tasks. The flipped condition additionally included working in groups and sharing their results in class. Remarkably, this flipped classroom implementation contained enrichments only in secondary processing but not in primary processes. Nevertheless, Burgoyne and Eaton (2018) reported a small but significant difference in quiz results between traditional and flipped condition.

These examples illustrate how differently flipped classroom can be implemented in higher education. In both cases, students achieved higher performance in the flipped conditions. Yet, we are unable to conclude from the heterogeneous body of (quasi-) experimental manipulations which components induce beneficial effects because flipped classroom conditions often offer more learning opportunities than the non-flipped conditions by enriching primary and secondary processing of learning content. Consequently, several researchers pointed out that more homogenous standards in research on flipped classroom could facilitate the identification of underlying mechanisms (Jensen et al., 2015; Chen et al., 2018). Kapur et al. (2022) demand that in research on flipped classroom, the interventions should be better controlled in order to gain more solid insights.

Table 1 schematically summarizes the research options which result from the distinction of primary and secondary processing (columns 2 and 3), the idea of shifting primary processing to out of class and secondary processing into class (rows 2–5), and the opportunity to enrich primary or secondary processing or both (rows 3–5). We suggest this schema as an orientation framework for designing and evaluating research on flipped classroom in a more differentiated way.

2.2 Framework of the present studies

We recommend investigating flipped classroom by varying single, distinctly defined components of flipped classroom step by step instead of investigating arrangements varying many components of

TABLE 1 Illustration of traditional and flipped classroom approaches.

Flipped classroom implementation		Primary processing of learning contents	Secondary processing of learning contents
1.	Traditional classroom/non-flipped classroom	In class (supervised learning)	Out of class (self-study)
2.	Merely flipped classroom	Out of class (self-study)	In class (supervised learning)
3.	Flipped classroom (primary processing enriched)	Out of class (self-study with enrichment)	In class (supervised learning)
4.	Flipped classroom (secondary processing enriched)	Out of class (self-study)	In class (supervised learning with enrichment)
5.	Flipped classroom (primary and secondary processing enriched)	Out of class (self-study with enrichment)	In class (supervised learning with enrichment)

Each row describes a slightly different implementation of flipped classrooms, the variation from traditional classroom increasing with the descending rows.

flipped classroom at once. Following this approach, research may answer the question whether (a) flipping primary processing from in to out of class, (b) flipping secondary processing from out of into class, (c) enrichments to primary processing, (d) enrichments to secondary processing and/or (e) the interplay between the aforementioned variations produce essential effects on learning.

In the present studies, we focused on option (a) to observe if there are any effects on learning if flipping entails changes exclusively to primary processing. We therefore implemented a minimalistic concept of flipping primary processing by using the video recordings of a traditional lecture to present learning content out of class in the flipped condition. Secondary processing, however, was operationalized highly similar in the non-flipped and flipped condition of the lecture. This design ensures that the students were presented with the same contents in the flipped and non-flipped condition. This type of comparison between flipped and non-flipped conditions was conducted in two thematically different psychology lectures (studies 1 and 2).

3 Study 1

We investigated the effect of merely flipping primary processing (see Table 1, row 2) on learning performance in psychology lectures at university. Participants were teacher students of three psychology courses (lecture series), each spanning a complete semester. One lecture series implemented primary processing in class (non-flipped condition), two lecture series implemented primary processing out of class (flipped condition) using the video recordings of the non-flipped lecture series for presenting the learning content. Secondary processing was highly similar in all lecture series. To assess and compare the students' learning performance, we evaluated the quality of written products which students had to submit to pass the course. If the mere flipping of learning activities from in to out of class is effective, the quality of students' written products in the flipped condition should be higher compared to the non-flipped condition.

3.1 Method

3.1.1 Design and participants

In Winter Semester 2019/20, $N = 61$ students participated in a traditionally organized lecture series (non-flipped condition) consisting of 12 weekly sessions of 90 min each. Primary processing included nine sessions, in which the lecturer used Power Point to present the learning topics orally and answered to students' questions in class. Each 90-min session during the semester was completely recorded on video and the recordings could be accessed on the university's learning platform one or two days later. Each video showed the lecturer's presentation in full length including, in a separate window, the Power Point slides. All recordings were available for the students until the end of the semester. For secondary processing, students received study questions with the lecture recordings and three writing assignments. These tasks were accomplished during three self-study phases (out of class).

In the following Summer Semester 2020 ($N = 138$) and Winter Semester 2020/21 ($N = 107$), students participated in a flipped version of the same lecture series (flipped condition). In these semesters, the lecture recordings (recorded during the Winter Semester 2019/20) were made available to students on the university's learning platform. The videos were published once per week in the same temporal pattern as in the non-flipped condition. All recordings were available to the students until the end of the semester. The students were asked to study the lecture recordings self-paced out of class. Questions concerning the content of specific sessions could be uploaded to a question forum on the learning platform within the same week of presentation to be answered by the lecturer in written form. Beyond, the students had the opportunity to discuss their questions orally with the lecturer in three (summer semester 2020) or five (winter Semester 2020/21) discussion sessions distributed across the semester. These discussion sessions lasted up to 45 min and were conducted – due to Covid-19 regulations – in a synchronous video conference format. The three self-study phases were conducted in the same way as in the non-flipped condition. The flipped and the non-flipped lectures were given by the same lecturer.

Flipped and non-flipped condition were highly similar. In primary processing, the same learning content was presented by the same lecturer with the same methods and in the same style in both conditions. However, in the non-flipped condition, the lecturer presented the learning contents in class, whereas in the flipped condition, the students perceived the lecturer via recordings of exactly these lectures. In both conditions, the students had opportunities to ask questions. In the non-flipped condition, students could ask during the live lecture in presence, whereas in the flipped condition, students could raise questions via the learning platform in a question forum or the discussion sessions after having worked through the lecture recording. In both conditions, the lecture recordings were available for recapitulating the lecture's content. Secondary processing included the same tasks for both conditions. Under the flipped condition no additional learning activities were requested compared to the non-flipped condition. The implementation of flipped classroom in Study 1 can therefore be classified as merely flipped classroom, as categorized in Table 1 (row 2).

3.1.2 Learning materials

The lecture series is part of a psychology module for teacher students with a Bachelor's (of Education) degree focussing the interplay of psychological research and application at school. Students are required to have completed successfully a prior basic psychology module. The lecture series addressed research and application of evidence-based learning and teaching methods, such as spaced learning, retrieval practice, multimedia learning, and metacognition.

In presenting the learning content, the lecturer informed about research and theoretical accounts of the evidence-based learning and teaching methods using Power Point presentations. The presentations were uploaded in portable data format to the university's learning platform one day before each week's lecture (non-flipped condition) or together with the current week's lecture recording (flipped condition). The lecture recordings presented the lecturer, his voice and in split screen the current page of the Power Point presentation. The students could switch to a full-screen mode showing only the current page of the presentation but not the lecturer. Clicking on the side bar showing the miniaturized pages of the presentation allowed for jumping back and forth in the video. The recordings were available from the date of their publication until the students passed their oral exam at the end of the module. Within this time frame, there was no limitation in watching the recordings.

Furthermore, the students in both conditions received a list of references mentioned during the lectures and a list of self-study questions to help students reflecting their understanding of topics and to guide their self-regulated learning activities. Answering the study questions, however, was optional and not part of the study requirements.

In the self-study phases, students were prompted to apply the previously presented learning contents by designing applications of evidence-based learning and teaching methods in school, as concrete as possible and in written form.

3.1.3 Learning performance and dependent variables

In the flipped and the non-flipped condition, the same type of learning performance was requested. The students submitted three written products to pass the lecture series and to be allowed to register for the final exam of the module. In the three self-study phases across the semester, the students designed an application of an evidence-based learning and teaching method in written form. The students were instructed to include the following aspects: (a) short definition of the evidence-based learning and teaching method, (b) elaborated description of the application in a concrete learning situation with a concrete learning content at school, (c) reflection on the theoretically assumed mechanisms connecting learning method and performance in the given situation, and (d) reflection on potential practical problems within this context (including solution approaches). The written products were not graded but checked by the lecturer and a summarizing feedback across all written products was given to the students including suggestions to review and revise their applications. Reviewing and revising their applications, however, was optional and not part of the study requirements.

For research purposes, we assessed the quality of the students' written products. To this end, we selected four indicators: (a) the extent of the written product, (b) the concreteness of the application description, (c) the reflection on theoretical mechanisms in the described situation, and (d) the reflection on practical problems.

The first indicator reflects the student's effort and was operationalized by the number of words in the written product. The other three indicators represent the task goals of the assignment. Trained and blinded evaluators checked the written products and rated the extent the task goals were achieved. On a three-point scale, the evaluators rated whether the student addressed and achieved each goal "appropriately" (2 points), "in part appropriately" (1 point), or "not at all "(0). We decided to realize this evaluation process only with the students' first submission within each of the three semesters to avoid confounding with feedback effects. At this time, students could select one of two topics (spaced learning or testing effect). Each written product was rated by two of three evaluators. The evaluators were trained with ten written products of a prior semester. The assignment of two evaluators to one written product was counterbalanced across topics and semesters.

3.1.4 Procedure

The lecture series of all three semesters were facultative and the students signed on deliberately. However, participation was only approved if the student submitted three written products at definite dates during the semester. Two weeks after submitting, a summarizing feedback across all submissions was given and made available on the learning platform of the university.

3.1.5 Statistical analyses

The ratings of the quality of the students' written products were checked for their inter-rater reliability by computing Spearman correlations between the two independent ratings of each quality indicator. In case of mismatching ratings, the mean of the two ratings was computed for further analyses.

To assess the effect of flipping the setting of presenting contents (in class or on video recordings out of class), a univariate ANOVA was conducted with the number of words of each written product. For the remaining quality indicators (concreteness of the description, reflection on of theoretical mechanisms, and reflection on practical problems), due to the scale level, Kruskal-Wallis-tests were computed. In each analysis, semester was the between-participants factor (Winter Semester 2019/20 vs. Summer Semester 2020 and Winter Semester 2020/21). An alpha of 0.05 was used as the level of statistical significance and effect sizes are reported as partial eta square (η_p^2) for indicator (a) and eta square (η_H^2) for indicators (b–d). In case of significant effects of the semester, pairwise comparisons were executed (including Bonferroni corrections).

3.2 Results

3.2.1 Preparatory analyses

The interrater reliability of the quality indicators across all three criteria was $\rho = 0.68$ ($p < 0.001$). Furthermore, we computed the specific interrater reliability for each criterion: concreteness of the application description ($\rho = 0.28$, $p = 0.031$), reflection on theoretical mechanisms in the described situation ($\rho = 0.68$, $p < 0.001$), reflection on practical problems ($\rho = 0.80$, $p < 0.001$). Boxplot analyses identified two extreme cases in the Winter Semester 2019/20 (non-flipped condition) regarding the extent of the written products. Consequently, these two cases were excluded from further analyses.

3.2.2 Hypotheses testing

Means and standard deviations of the dependent variables are reported in Table 2. Comparing the number of words of the written products revealed a significant effect of the semester: $F(2, 301) = 25.58$, $p < 0.001$, $\eta_p^2 = 0.14$. Pairwise comparisons further revealed that written products in the non-flipped condition of the Winter Semester 2019/20 consisted of less words than written products in the flipped condition of the Summer Semester 2020 ($p < 0.001$) and in the flipped condition of the Winter Semester 2020/21 ($p < 0.001$).

Comparing the ratings of the concreteness in the application description (see Table 2) revealed a significant effect of the semester: $H(2) = 7.24$, $p = 0.03$, $\eta_H^2 = 0.02$. Pairwise comparisons via Mann-Whithey-U Tests further revealed that the concreteness of the description was significantly lower in the non-flipped condition than in the flipped condition of the Summer Semester 2020 ($p = 0.009$), and numerically lower compared to the flipped condition of the Winter Semester 2020/21 ($p = 0.07$).

Comparing the ratings concerning the reflection on theoretical mechanisms in the described situation (see Table 2) revealed no significant effect of the semester: $H(2) = 1.52$, $p = 0.47$.

Comparing the ratings concerning the reflection on practical problems (see Table 2) revealed a significant effect of the semester: $H(2) = 17,04$, $p < 0.001$, $\eta_p^2 = 0.05$. Pairwise comparisons via Mann-Whithey-U Tests further revealed significantly lower ratings of reflection in the non-flipped condition compared to the flipped condition, in the Summer Semester 2020 ($p < 0.001$) and Winter Semester 2020/21 ($p = 0.02$).

3.3 Discussion

In all four indicators of learning performance mean performance values were descriptively lower in the non-flipped condition than in both flipped conditions. This led to a significant overall effect in three of the four quality indicators (extent of the written products, concreteness of the description, and reflection on practical problems). At the level of Bonferroni-corrected pairwise comparisons, the written products in the non-flipped condition were significantly shorter und presented significantly less profound practical reflections than in both flipped conditions. Differences in performance reached small to medium effect sizes (Cohen, 1992, p. 157; Cohen, 1988, pp. 283–286). The concreteness of the description was significantly lower in the non-flipped condition than in the flipped condition of the Summer Semester 2020 but the difference between the non-flipped and the flipped condition in the Winter Semester 2020/21 narrowly missed the level of significance. The reason might the unsatisfactory interrater-reliability of the concreteness variable.

This pattern of results is compatible with the hypothesis that enrichment of the teaching activities is no necessary condition for improving learning in flipped classroom. No doubt, enriching the learning setting may gain additional benefits. However, even flipping primary processing to out of class without any enrichment seems to be sufficient for a learning-supportive effect. Moreover, Study 1 demonstrated this effect not only once. The comparison of both flipped versions of the lecture series with the traditional, non-flipped lecture series showed similar differences. Thus, the findings demonstrate that mere flipping of primary processing can improve learning performance.

4 Study 2

In the second study, we pursued the same research question as in Study 1. Does the mere flipping of primary processing (see Table 1) suffice to enhance learning performance in a university lecture? We investigated teacher students in two psychology lecture series, each spanning a complete semester. One lecture series implemented primary processing in class (non-flipped condition); in the other lecture series, primary processing was flipped to out of class (flipped condition) using the video recordings of the non-flipped lecture series as input. As in the first study, the conditions for secondary processing were highly similar in the flipped and the non-flipped condition.

The first study demonstrated the stability of the effect of merely flipping primary processing from in to out of class: The flipped versions of both semesters showed similar differences to the non-flipped lecture. The second study aimed at broadening the data base by examining a thematically different lecture given by a different teacher and measuring different dependent variables. Whereas, in the first study, we focused on the quality of midterm writing assignments, we measured performance in a knowledge test at the end of the semester in the second study. The dependent measures not only included learning outcome at the cognitive level but also at the metacognitive level. Given that (a) mere flipping is a sufficient condition for enhancing learning outcome and (b) enrichment of the teaching activities is no necessary condition, we expected that performance in a knowledge test at the end of the semester should be higher in the flipped than in the non-flipped condition.

TABLE 2 Measures of learning outcomes in study 1 (means, SD in parentheses).

Dependent variables (learning outcomes)	Non-flipped $N = 61$	Flipped 20 $N = 138$	Flipped 21 $N = 107$
a) Number of words in written product	319 (137)	482 (154)	426 (150)
b) Concreteness of description	1.03 (0.66)	1.31 (0.63)	1.24 (0.55)
c) Reflection on theoretical mechanisms	0.59 (0.76)	0.73 (0.85)	0.77 (0.83)
d) Reflection on practical problems	0.53 (0.63)	1.02 (0.72)	0.83 (0.71)

Flipped 20 abbreviates summer semester 2020, flipped 21 winter semester 2020/21. Quality of learning outcome regarding indicators b-d were rated on a three-point scale of 0–2 points regarding each answer's appropriateness.

4.1 Method

4.1.1 Design and participants

The hypothesis was tested by comparing students' learning outcomes between two semester-spanning lecture series on educational psychology for teacher students in their master's phase. In Winter Semester 2019/20, $N = 187$ students ($n = 139$ female, $n = 48$ male) participated in a traditionally organized lecture series (non-flipped condition) consisting of 14 weekly sessions of 90 min each. The lecturer presented the learning contents orally supported by Power Point presentations and answered the students' questions while they were present in the lecture hall. As in Study 1, each 90-min session of the semester was completely recorded on video, which was accessible one day later on the university's learning platform and remained available to the students until the final test in the end of the semester.

In the following Winter Semester (2020/21), $N = 226$ students ($n = 137$ female, $n = 89$ male) participated in the flipped implementation of the same lecture series. In the flipped condition, the recordings of the 14 lectures of the non-flipped condition (recorded during the previous winter semester) served to present the learning content and were made available to the participants on the university's learning platform. Each video showed the lecturer's presentation in full length including, in a separate window, the Power Point slides. The participants watched the recordings instead of attending a live lecture in class and could use it later for recapitalizing the lecture's content. The videos were published one per week in the same temporal pattern as the corresponding live lectures in the non-flipped condition. All recordings were available for the students until the final test in the end of the semester. The students were asked to study the lecture recordings self-paced and to prepare for sessions with the lecturer in which their questions would be discussed. These discussion sessions ($n = 6$) took place every two weeks (with one exception due to winter holidays), lasted for 90 min and were conducted – due to Covid-19 regulations – in a synchronous video conference format. The flipped and the non-flipped lectures were given by the same lecturer.

Summarized, due to using the lecture recordings for primary processing, in the flipped and the non-flipped condition, exactly the same learning content was presented by the same lecturer with the same methods and in the same style. However, in the non-flipped condition, the lecturer presented the learning contents live in class, whereas in the flipped condition the students saw and heard the lecturer in the recordings of exactly these lectures. In both conditions, the students had opportunities to ask questions: in the non-flipped condition, students could ask during the lecture, in the flipped condition, students could ask in separate discussion sessions after working through the lecture recording. In both conditions, the lecture recordings were available for recapitulating the lecture's content. Secondary processing included working on the self-study questions, which were the same for both conditions. Under the flipped condition no additional learning activities were requested compared to the non-flipped condition. Therefore, as in Study 1, the main difference between non-flipped and flipped condition consisted in the setting of presenting learning content during primary processing. The implementation of flipped classroom in Study 2 can therefore be classified as merely flipped classroom, as categorized in Table 1.

4.1.2 Learning materials

The lectures informed about psychological research relevant to learning and teaching at school, such as cognitive learning, memory, motivation, development during adolescence, social processes in and between groups. The topics were prepared for teacher students with a Bachelor's (of Education) degree with only minor prior knowledge of psychology.

The Power Point presentations visualized information that complemented the lecturer's talk and were thought to support comprehension, such as pre-and reviews, citations from the research literature, figures showing research designs and results. The presentations were uploaded in portable data format to the university's learning platform one day before each week's lecture (non-flipped condition) or before the current week's lecture recording was published (flipped condition).

Furthermore, the students in both conditions received a list of references mentioned during the lectures, a commented list of textbooks, and a list of self-study questions. Reflecting on these questions was intended to help students to guide their self-regulated learning activities during secondary processing. Answering the study questions was not obligatory and not part of the study requirements.

4.1.3 Test materials and dependent variables

The students in both conditions completed the same test at the end of the semester. The test consisted of 40 confidence-weighted true-false items in the form of one-sentence statements (Dutke and Barenberg, 2015). The statements tested facts presented in the lectures or propositions that could be inferred from the knowledge acquired in the lectures.

For each statement in the test, the students judged whether it was true or false and concurrently indicated their confidence in the correctness of each decision. The true-false decision and the confidence judgment were integrated on a single scale resulting in four response options: (1) I am *sure* the statement is *correct*, (2) I think the statement is *correct*, but I am *unsure*, (3) I think the statement is *incorrect*, but I am *unsure*, or (4) I am *sure* the statement is *incorrect*. Thus, students had to give only one answer to each item.

From these data, different measures were derived assessing the students' cognitive and metacognitive learning performance. As a measure of correctness (cognitive performance), the relative number of correct true-false decisions was computed, irrespective of the reported confidence level. As a measure of confidence (metacognitive performance), the relative number of confident judgments was computed, irrespective of the answers' correctness. To assess the quality of metacognitive monitoring, measures of correctness and confidence were combined (see Appendix A for formulas). The following indices reflect different facets of metacognitive monitoring accuracy typically assessed in the field of metacognition (see Schraw, 2009; Schraw et al., 2013; Barenberg and Dutke, 2019). The *bias* of the confidence judgments (BS) was calculated by subtracting the proportion of correct answers from the proportion of confident answers. Positive or negative values reflect general over-or underconfidence across all items. As the BS score does not reveal whether correct and incorrect answers are biased in the same way, two conditional probabilities were calculated. The *confident-correct probability* (CCP) indicates the proportion of confident answers given the answer is correct, and the *confident-incorrect probability* (CIP) indicates the proportion of confident answers given the answer is incorrect.

4.1.4 Procedure

Both lectures were facultative and students signed on deliberately. However, participation was only approved if the student passed the end-term test. The test was written at the end of the semester, one week after the final session. The students were familiar with the test format due to presenting sample items throughout the semester. The test was written in the lecture hall and the participants had 40 min to complete it.

4.1.5 Statistical analyses

T-tests were used to test the hypothesis that the dependent variables differed between the flipped and the non-flipped condition, showing a higher learning outcome in the flipped condition. The sizes of the effects were measured with Cohen's *d*.

4.2 Results

Comparing the relative number of correct answers demonstrated the expected difference between the experimental conditions. Means and standard deviations of the dependent variables are reported in Table 3. In the flipped condition, more items were answered correctly than in the, non-flipped condition. Though the effect is small ($d = 0.22$), it is significant, $t(411) = 2.22$, $p = 0.027$. The relative number of answers given with confidence did not differ between the flipped and the non-flipped condition, $t(411) = 1.28$, $p = 0.200$.

The metacognitive variables indicate that participants of the flipped condition underestimated the correctness of their answers more than participants of the non-flipped condition, $t(411) = 2.98$, $p = 0.003$, $d = 0.29$. However, confidence was not generally reduced in the flipped condition. Particularly, the probability of answering with confidence given the answer was incorrect (CIP) was lower in the flipped condition than in the non-flipped condition, $t(410) = 3.27$, $p < 0.001$, $d = 0.32$. In contrast, the probability of answering with confidence given the answer was correct (CCP) did not differ significantly between the flipped and the non-flipped condition.

4.3 Discussion

Regarding the cognitive level of learning success, Study 2 shows evidence in support of our expectation: Students in the flipped condition performed better in the knowledge test at the end of the semester than students in the non-flipped condition. Evidence at the metacognitive level of learning success, at the first glance, seemed to point in the opposite direction, as participants in the flipped condition underestimated their performance more than students in the non-flipped condition. Further analysis, however, revealed that students in the flipped condition, compared to the non-flipped condition, judged the correctness of their answers particularly critically when their answers were actually wrong. This effect indicates that students in the flipped condition were more sensitive to uncertain knowledge, which is a desirable learning outcome. Differentiating adequately between correct and incorrect answers before external feedback is provided, is essential to successfully regulate one's learning process. Thus, these results suggest benefits at the cognitive as well as the metacognitive level of learning as a function of flipping primary processing.

TABLE 3 Measures of learning outcomes in study 2 (means, SD in parentheses).

Dependent variables (learning outcomes)	Non-flipped $N = 187$	Flipped $N = 226$
Correct answers	0.80 (0.10)	0.82 (0.10)
Confident answers	0.73 (0.19)	0.70 (0.18)
Bias of the confidence judgements (BS)	−0.07 (0.16)	−0.12 (0.15)
Confident-correct probability (CCP)	0.78 (0.17)	0.76 (0.17)
Confident-incorrect Probability (CIP)	0.50 (0.28)	0.41 (0.28)

Values regarding correct answers and confident answers represent the according percentage, values range from 0 to 1. BS can reach values between 1 and-1, indicating over-or underconfidence in answers when deviating from 0. The remaining variables indicate proportion of confident answers, given that the answer is correct (CCP) or incorrect (CIP), on a range 0 to 1.

5 General discussion

We proposed, for investigating how flipping a classroom fosters learning, to vary distinctly defined components of flipped classroom step by step rather than comparing complex and variously enriched learning arrangements. In the two studies reported above, we demonstrated learning benefits for flipped over non-flipped conditions while merely flipping primary processing. Thereby, we neglected possible enrichments and took a defined first step towards understanding mechanisms underlying flipped classroom learning. Most measures indicated a learning benefit on the cognitive level of learning in both studies. Study 2 additionally investigated metacognitive measures of learning, which indicated a more accurate differentiation between correct and incorrect answers within the flipped condition.

5.1 Learning performance

Recent meta-analyses (Chen et al., 2018; Låg and Sæle, 2019; Zheng et al., 2020) demonstrated statistically significant small effects for flipped classroom conditions on cognitive learning performance. These results are in line with the results found in our studies. Study 1 showed that learning performance in flipped classroom surpassed cognitive learning performance in the non-flipped comparison group regarding the extent and quality of written products with small to medium effect sizes ($\eta_p^2 = 0.14$ and $\eta_H^2 = 0.02$–0.05). Study 2 extended these results, finding higher performance in a knowledge test at the end of the semester in the flipped than in the non-flipped condition with a small effect size ($d = 0.22$).

Metacognitive monitoring has been established to enhance learning (Swanson, 1990). This especially applies to situations in which learners have to regulate their own learning process (Boekaerts, 1999; Pintrich, 2002), such as the self-study part (primary processing) of flipped classroom. While research on metacognition in the context of flipped classroom is scarce, few studies indicate learning benefit on the metacognitive level (Hsu and Hsieh, 2014; van Vliet et al., 2015; Bredow et al., 2021). The results from our second study add to this research by demonstrating that students in the flipped condition monitored their learning more critically as they were less confident in incorrect answers. As students in the flipped condition also answered more often correctly than the participants of the non-flipped lecture series, our second study supported the idea that increased quality of monitoring contributed to enhanced learning performance on a cognitive level.

5.2 Enrichment as a facultative condition

A variety of researchers (e.g., McLaughlin et al., 2014; Altaii et al., 2017; Gilette et al., 2018; Cannon, 2021) report extensive use of resources to implement flipped classroom approaches, especially consuming time as well as personnel in the preparation of comprehensive material for primary processing, specifically videos. Noetel et al. (2021) suggest that extensive revision and employing design principles may increase efficacy in teaching via video. This would point to enrichment of primary processing materials resulting in higher learning outcomes for learners in flipped classrooms. The present studies, however, show that enrichment in primary processing is not necessary in order to enhance learning. Merely utilizing lecture recordings held the design of information input in primary processing constant between flipped and non-flipped conditions. Therefore, observed differences in learning performance cannot be explained by enrichment which indicates that flipping primary processing from in to out of class alone fostered learning performance. According to these results, enrichment (regarding learning material as well as elaboration tasks) is not required in order to successfully flip a classroom although it may contribute to the further optimization of learning arrangements.

5.3 Differences in primary processing

What, if not enrichment, caused the advantage of the flipped lecture condition? As the design of the learning activities with regard to secondary processing did not differ substantially between flipped and non-flipped condition, the reason could only lie in the flipping of primary processing from in class to out of class. However, neither the amount of information delivered nor the way how this information was conveyed differed, as the recordings of non-flipped lectures served as input for flipped lectures. If, at all, an advantage may be seen for the non-flipped condition, as students in this condition were presented with the live lecture in classroom *and* lecture recordings, whereas the students in the flipped condition only had the recordings at their disposal. Thus, the observed learning benefits in flipped conditions might seem counterintuitive (cf. Missildine et al., 2013). Apparently, even when lecture recordings are available in non-flipped conditions, they may not be used or be used in an ineffective way when seen as additional material to in-class lectures. Learning analytics suggest that additional availability of lecture recordings in traditional lectures show only marginal improvement in student academic performance (Owston et al., 2011). It seems likely that students in non-flipped conditions used lecture-recordings mostly for exam preparations (elaboration in secondary processing). Support for this thought can be seen in levels of effect sizes, which were small to medium in Study1, where data was gathered at the beginning of the semester, before exam preparation efforts can be expected, while in Study 2, effect sizes were small and data was collected at the end of semester, after exam preparations.

A second idea why primary processing in flipped conditions was more effective than in non-flipped conditions, refers to learners' increased requirements and opportunities of self-regulation in the flipped lectures. Especially when learners prepare learning material out of class (primary processing) for in class secondary processing, the role of metacognition in successful learning may be emphasized. In flipped classrooms, learners are aware that after out of class primary processing, the quality of their acquired knowledge co-determines how much they can benefit from supervised secondary processing in class. This may result in more frequent monitoring of their own learning process. In contrast, during secondary processing in non-flipped higher education courses, no immediate consequences arise due to learners unsuccessful self-regulated primary processing. Typically, greater efforts in secondary processing occur shortly before achievement tests or exams at the end of the semester (e.g., Algayres and Triantafyllou, 2020), which may decrease the sense of urgency to monitor learning during in class primary processing.

Increased or more effective metacognitive monitoring in flipped lectures may also affect the way how students controlled information input via the lecture recording. For example, Abeysekera and Dawson (2014) as well as Rankl (2016) suggest that cognitive load during primary processing can be managed more easily by means of controlling primary processing in flipped classroom learning (such as pausing, rewinding, and repeating the lecture video). Thereby, the pace of presentation can be aligned with the learner's pace of processing learning content allowing more time to utilize learning strategies that benefit primary processing of learning material.

In traditional lectures, students can interact with learning material by listening, taking notes, highlighting, or asking questions which represent passive to active engagement, according to Chi and Wylie (2014). With a shift of primary processing from in-class lecture to out of class, the variety of learning strategies and methods available in primary processing increases for students. Increased variety of strategies may enhance opportunities to engage actively or constructively (Chi and Wylie, 2014), thereby enable further elaboration. Considering disuse theory (Bjork and Bjork, 1992), we can assume that differences in primary processing likely affect storage strength of new learning content. The option to individually adapt the pace of presentation, insert pauses to utilize learning strategies, or rewind and re-watch parts of a lecture, seem appropriate to make a difference in information processing.

To date, few studies focused on learners' interaction with out of class learning material, as a gateway to assess primary processing in flipped classrooms. In a recent study with aspiring math-teachers, Dusanek and Kollar (2023) reported increasing extraneous cognitive load along with decreasing opportunities to interact with learning videos. Compared to traditional lectures, Karaca and Ocak (2017) as well as Müller and Wulf (2023) demonstrated how primary processing in flipped classroom can lead to reduced cognitive load in university students. Both results imply evidence in support of Abeysekera and Dawson's (2014) rational regarding cognitive load as potential bridge between primary processing and learning performance in flipped classroom. Altierie et al. (2018) documented increased learning success in engineering students when pausing while learning with a video compared to watching the same video without pausing. While the gap between these yet isolated findings and a comprehensive understanding of flipped classroom mechanisms remains, such findings support the perspective that learners' interaction with learning material in primary processing affects their capacity and success in primary processing and thereby their subsequent learning outcomes.

5.4 Limitations and research perspectives

Data collection regarding the flipped conditions was conducted in 2020 and 2021 during distance learning within the Covid-19 pandemic. It is possible, that aspects of observed learning benefit can be attributed to this unique setting. However, our results demonstrated some robustness. We obtained measures of learning outcome within the course of the semester (Study 1) as well as at the end of semester (Study 2), offering a variety of supplementary indicators of student's learning performance before and after a period of exam preparations including data from open-and closed-answer items. Moreover, in Study 1, the positive effects of flipping primary processing from in to out of class were not only observed in Summer 2020, but also in the following winter semester. Thus, it is improbable that these findings are simply artefacts of this specific time period. Nevertheless, results need to be replicated under conditions beyond the specific pandemic situation.

In Study 2, it becomes evident that students' cognitive performance was at the higher end of the performance spectrum although no indications of a ceiling effect were found. Nevertheless, observed effects may be even larger when replicated in a field with more variance within the dependent measures.

The mere flipping approach, as implemented in the present studies, brings primary processing into focus, thereby neglecting potential additional effects of enriching group interaction, which is often implemented in secondary processing of flipped classrooms. It could be argued that our research approach does not allow a full picture of potential effects of flipping. Yet, this is the core idea of this approach: In carefully controlling implementation with precisely defined changes from non-flipped to flipped classroom, we can gather empirical evidence and clarify which of these components make a difference in learning performance. Therefore, this research needs to be extended by replicating the mere flipping effect as well as gradually enrich flipped classroom settings as illustrated in Table 1. In the present study, we opted for investigating our research question in a genuine teaching context and avoided an artificial learning situation. If, in future research, the effects of possible enrichments of the learning context are focused, it might be necessary to add follow-up tests in order to assess long-term effects, commonly fostered by evidence-based teaching methods (Lyle et al., 2020; Kim and Webb, 2022) and switch to even better controlled laboratory learning setting.

In the present design, students had many degrees of freedom how they organized their learning activities. For example, accessing lecture recordings via learning platform, working with the study question (Study 1 and 2) or revising their written products (Study 1) was optional. One may speculate whether making these activities mandatory would have diminished inter-participants variance and, therefore, augment the flipped classroom effects. However, in compliance with our study program regulations, arrangements that introduce additional study requirements, such as additional mandatory activities could not be realized. For similar reasons, it was not possible assign students randomly to the flipped or non-flipped condition. Whereas these features may be evaluated as a limitation, it is also a strength of our studies that they were conducted in the context of the ordinary teaching program, which enhances external validity.

Measuring cognitive load, time spent with learning material as well as the use of means to control the pace of primary processing (such as pausing, rewinding, repeating the lecture recordings), might help to explore the mechanisms underlying the effect of merely flipping primary processing. We suggest to invest research efforts into identifying how learners interact with the learning content in primary processing out of class compared to in class, for our research shows that merely flipping does result in different learning outcomes. Our results also indicate clear benefits on the cognitive and the metacognitive level of learning performance. This is why we suggest zooming into this aspect of flipped classroom and exploring the opportunities and incentives to monitor one's own learning under flipped and non-flipped conditions.

In any case, future research would be well advised to differentiate between effects that could be attributed to the flipping of learning activities itself and effects of additional learning activities that become possible because learning activities were flipped from in to out of class. We anticipate that such an approach will enhance our understanding of *how* flipped classroom fosters learning – beyond measuring that the effect is existent.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Based on the declaration in the ethical guidelines of the Deutsche Gesellschaft für Psychologie (2022) as well as the American Psychological Association (2017), psychologists may dispense with informed consent where research would not reasonably be assumed to create distress or harm and involves the study of normal educational practices, curricula, or classroom management methods conducted in educational settings. Since this applies to these studies, with data derived from lectures held within the regular teacher education curricula at the University of Münster, individual written consent was not requested. This application of the ethical guidelines was conducted in accordance with the ethics checklist of the department of Psychology/Sport and Exercise Science at the University of Münster. Thus, ethical review and approval was nor required in this study on human participants in accordance with legislation and institutional requirements.

Author contributions

GB: Conceptualization, Formal analysis, Project administration, Visualization, Writing – original draft, Writing – review & editing. JB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. SD: Conceptualization, Data curation, Formal analysis, Investigation, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abeysekera, L., and Dawson, P. (2014). Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research. *Higher Educ. Res. Dev.* 34, 1–14. doi: 10.1080/07294360.2014.934336

Algayres, M., and Triantafyllou, E. (2020). Learning analytics in flipped classrooms: a scoping review. *Electr. J. E-Learn.* 18:3. doi: 10.34190/JEL.18.5.003

Altaii, K., Reagle, C. J., and Handley, M. K. (2017), (June 24–28), Flipping an engineering thermodynamics course to improve student self-efficacy (Paper presentation). ASEE Annual Conference & Exposition, Columbus, Ohio.

Altierie, M., Köster, A., Friese, N., and Paluch, D. (2018). Größerer Lernerfolg durch Pausen in Lernvideos? Eine Untersuchung zu segmentierten Lernvideos in der Ingenieurmathematik. [higher learning success through pauses during learning videos? A study on segmented learning videos in mathematics for engineers] [paper presentation]. 52. Tagung für Didaktik der Mathematik-Jahrestagung der Gesellschaft für Didaktik der Mathematik, Paderborn.

American Psychological Association (2017). Ethical Principles of Psychologists and Code of Conduct (2002, Amended Effective June 1, 2010, and January 1, 2017). Available at: https://www.apa.org/ethics/code/ethics-code-2017.pdf (Accessed July 8, 2024).

Barenberg, J., and Dutke, S. (2013). Metacognitive monitoring in university classes: anticipating a graded vs. a pass-fail test affects monitoring accuracy. *Metacogn. Learn.* 8, 121–143. doi: 10.1007/s11409-013-9098-3

Barenberg, J., and Dutke, S. (2019). Testing and metacognition: retrieval practise effects on metacognitive monitoring in learning from text. *Memory* 27, 269–279. doi: 10.1080/09658211.2018.1506481

Bishop, J. L., and Verleger, M. A. (2013, June 23–26). The flipped classroom: a survey of the research. [Paper presentation]. ASEE Annual Conference & Exposition, Atlanta, GA.

Bjork, R. A., and Bjork, E. L. (1992). "A new theory of disuse and an old theory of stimulus fluctuation" in From learning processes to cognitive processes: Essays in honor of William K. Estes. eds. A. Healy, S. Kosslyn and R. Shiffrin. *2nd* ed (Hillsdale, NJ: Erlbaum), 35–67.

Boekaerts, M. (1999). Self-regulated learning: where we are today. *Int. J. Educ. Res.* 31, 445–457. doi: 10.1016/S0883-0355(99)00014-2

Bredow, C. A., Roehling, P. V., Knorp, A. J., and Sweet, A. M. (2021). To flip or not to flip? A meta-analysis of the efficacy of flipped learning in higher education. *Rev. Educ. Res.* 91, 878–918. doi: 10.3102/00346543211019122

Burgoyne, S., and Eaton, J. (2018). The partially flipped classroom. *Teach. Psychol.* 45, 154–157. doi: 10.1177/0098628318762894

Cannon, J. (2021). A first experience with video-based flipped-classroom teaching. Kyushu University. Available at: http://www.jamescannon.net/ja/a-first-experience-with-video-based-flipped-classroom-teaching/ (Accessed March 25, 2024).

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. *Med. Educ. Rev.* 52, 910–924. doi: 10.1111/medu.13616

Chi, M. T. H., and Wylie, R. (2014). The ICAP framework: linking cognitive engagement to active learning outcomes. *Educ. Psychol.* 49, 219–243. doi: 10.1080/00461520.2014.965823

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *2nd* Edn. New York: Routledge.

Cohen, J. (1992). Quantitative methods in psychology: a power primer. *Psychol. Bull.* 112, 155–159. doi: 10.1037/0033-2909.112.1.155

DeLozier, S. J., and Rhodes, M. G. (2017). Flipped classrooms: a review of key ideas and recommendations for practice. *Educ. Psychol. Rev.* 29, 141–151. doi: 10.1007/s10648-015-9356-9

Deutsche Gesellschaft für Psychologie (2022). Berufsethische Richtlinien [Professional Ethical Guidelines]. Available at: https://www.dgps.de/die-dgps/aufgaben-und-ziele/berufsethische-richtlinien/ (Accessed July 8, 2024).

Dusanek, V., and Kollar, I. (2023), (September 20). Interaktive gamifizierte Online-Lernvideos als Mittel zur Förderung des Wissenserwerbs bei Studierenden – Effekte unterschiedlicher Gestaltungsmerkmale [Interactive gamified online learning videos as a means to enhance knowledge acquisition in university students – effects of different design attributes]. In Y. Omarchevska (Chair) Lernen mit Vidoes 2 [learning with videos 2] [Symposium]. 19. Fachgruppentagung Pädagogische Psychologie (PAEPS) der Deutschen Gesellschaft für Psychologie (DGPs), Kiel. Available at: https://www.conftool.pro/paeps2023/index.php?page=browseSessions&form_session=11

Dutke, S., and Barenberg, J. (2015). Easy and informative: using confidence-weighted true–false items for knowledge tests in psychology courses. *Psychol. Learn. Teach.* 14, 250–259. doi: 10.1177/1475725715605627

Gilette, C., Rudolph, M., Kimble, C., Rockich-Winston, N., Smith, L., and Broedel-Zaugg, K. (2018). A meta-analysis of outcomes comparing flipped classroom and lecture. *Am. J. Pharm. Educ.* 82, 6898–6440. doi: 10.5688/ajpe6898

Hsu, L.-L., and Hsieh, S.-I. (2014). Factors affecting metacognition of undergraduate nursing students in a blended learning environment. *Int. J. Nurs. Pract.* 20, 233–241. doi: 10.1111/ijn.12131

Jensen, J. L., Kummer, T. A., and Godoy, P. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE-Life Sci. Educ.* 14, 1–12. doi: 10.1187/cbe.14-08-0129

Kapur, M., Hattie, J., Grossman, I., and Sinha, T. (2022). Fail, flip, fix, and feed – rethinking flipped learning: a review of meta-analyses and a subsequent meta-analysis. *Front. Educ.* 7:956416. doi: 10.3389/feduc.2022.956416

Karaca, C., and Ocak, M. A. (2017). Effect of flipped learning on cognitive load: a higher education research. *J. Learn. Teach. Digit. Age* 2, 20–27. Available at: https://dergipark.org.tr/en/pub/joltida/issue/55466/760080

Kim, S. K., and Webb, S. (2022). The effects of spaced practice on second language learning: a Meta-analysis. *Lang. Learn.* 72, 269–319. doi: 10.1111/lang.12479

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

Lage, M. J., Platt, G. J., and Treglia, M. (2000). Inverting the classroom: a gateway to creating an inclusive learning environment. *J. Econ. Educ.* 31, 30–43. doi: 10.1080/00220480009596759

Lyle, K. B., Bego, C. R., Hopkins, R. F., Hieb, J. L., and Ralston, P. A. S. (2020). How the amount and spacing of tretrieval practice affect the short-and long-term retention of mathematics knowledge. *Educ. Psychol. Rev.* 32, 277–295. doi: 10.1007/s10648-019-09489-x

McLaughlin, J. E., Roth, M. T., Glatt, D. M., Gharkholonarehe, N., Davidson, C. A., Griffin, L. M., et al. (2014). The flipped classroom: a course redesign to foster learning and engagement in a health professions school. *Acad. Med.* 89, 236–243. doi: 10.1097/ACM.0000000000000086

Missildine, K., Fountain, R., Summers, L., and Gosselin, K. (2013). Flipping the classroom to improve student performance and satisfaction. *J. Nurs. Educ.* 52, 597–599. doi: 10.3928/01484834-20130919-03

Müller, F. A., and Wulf, T. (2023). Differences in learning effectiveness across management learning environments: a cognitive load theory perspective. *J. Manag. Educ.* 48, 802–828. doi: 10.1177/10525629231200206

Noetel, M., Griffith, S., Delaney, O., Sanders, T., Parker, P., Del Pozo Cruz, B., et al. (2021). Video improves learning in higher education: a systematic review. *Rev. Educ. Res.* 91, 204–236. doi: 10.3102/0034654321990713

O'Flaherty, J., and Phillips, C. (2015). The use of flipped classrooms in higher education: a scoping review. *Internet High. Educ.* 25, 85–95. doi: 10.1016/j.iheduc.2015.02.002

Owston, R., Lupshenyuk, D., and Wideman, H. (2011). Lecture capture in large undergraduate classes: What is the impact on the teaching and learning environment? [Paper presentation]. New Orleans, USA: Annual Meeting of the American Educational Research.

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory Pract.* 41, 219–225. doi: 10.1207/s15430421tip4104_3

Rankl, M. (2016). Kognitive Belastungen in Lernprozessen–mediendidaktische Überlegungen zum inverted classroom model. [cognitive load in learning processes – media educational considerations to the inverted classroom model] PHScript 10, 31–35. Available at: https://phsalzburg.at/wp-content/uploads/2023/01/phscript_16_online-1.pdf

Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacogn. Learn.* 4, 33–45. doi: 10.1007/s11409-008-9031-3

Schraw, G., Kuch, F., and Gutierrez, A. P. (2013). Measure for measure: calibrating ten commonly used calibration scores. *Learn. Instr.* 24, 48–57. doi: 10.1016/j.learninstruc.2012.08.007

Swanson, H. L. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *J. Educ. Psychol.* 82, 306–314. doi: 10.1037/0022-0663.82.2.306

Talbert, R. (2022). How much research has been done on flipped learning? The 2022 (and final?) update. Available at: https://rtalbert.org/how-much-research-has-been-done-on-flipped-learning-the-2022-and-final-update/ (Available at: July 7, 2022).

van Vliet, E. A., Winnips, J. C., and Brouwer, N. (2015). Flipped-class pedagogy enhances student metacognition and collaborative-learning strategies in higher education but effect does not persist. *CBE Life Sci. Educ.* 14:141. doi: 10.1187/cbe.14-09-0141

Zainuddin, Z., and Halili, S. H. (2016). Flipped classroom research and trends from different fields of study. *Int. Rev. Res. Open Distrib. Learn.* 17:2274. doi: 10.19173/irrodl.v17i3.2274

Zheng, L., Bhagat, K. K., Zhen, Y., and Zhang, X. (2020). The effectiveness of the flipped classroom on students' learning achievement and learning motivation: a meta-analysis. *Educ. Technol. Soc. 23, 1–15.* Available at: https://psycnet.apa.org/record/2020-39682-001

Appendix

APPENDIX A 2 × 2 (correctness × confidence) data array resulting from the response scale (adapted from Barenberg and Dutke, 2013).

	Correct	Incorrect	
Confident	a	b	a + b
Unconfident	c	d	c + d
	a + c	b + d	a + b + c + d

The indices of metacognitive monitoring accuracy were computed according to the following formulas:

1. Bias of the confidence judgments (BS):

$$BS = \frac{(a+b)-(a+c)}{a+b+c+d}$$

2. Confident-correct probability (CCP):

$$CCP = \frac{a}{a+c}$$

3. Confident-incorrect probability (CIP):

$$CIP = \frac{b}{b+d}$$

The attributes of the answer to each item (*correct* or *incorrect, confident* or *unconfident*) were coded as 1 (given) or 0 (not given).