

Investigating the Effectiveness of exam wrappers in Biology Education: Exploring learning Environment, Student Factors and Research Design

L.C. (Lotte) van der Krabben (6419763)

Utrecht University

Graduate School of Life Sciences

Master Writing Assignment

Environmental Biology (Plant track)

Daily Supervisor: J.B. (Jonno) van Vulpen MSc

1st Coordinator: Dr. M.C. (Margot) Koster

2nd Coordinator: Dr. ing. J. (Jelle) Zandveld

April 24th, 2024

Investigating the Effectiveness of exam wrappers in Biology Education: Exploring learning Environment, Student Factors and Research Design

Abstract

Academic success is often a result of acquired metacognitive and self-regulated learning skills. Despite their acknowledged significance, many students still encounter difficulties in developing effective metacognitive strategies and reflective thinking abilities. Post-exam wrappers are a method to cultivate metacognitive skills used across diverse academic disciplines, notably in biology education. These wrappers function as reflective exercises aimed at stimulating metacognition, encouraging students to critically evaluate their learning processes, and enhancing their academic performance. However, the effectiveness of exam wrappers remains a subject of debate, with literature presenting a spectrum of findings regarding their impact on metacognition and course outcomes. This review examines the effectiveness of post-exam wrappers in promoting metacognition and/or course performance in biology education, considering a range of factors, related to students' characteristics, curriculum design, and educational research design. Through an analysis of existing research, this review offers insights into optimizing the implementation of exam wrapper to maximize its outcomes. Overall, exam-wrappers could enhance metacognition and academic performances if implemented optimally with the appropriate strategies, considering factors like student characteristics and course environments.

Layman's Summary

In education, students are often struggling with effectively developing learning strategies. One approach to tackle this issue is through a method called "post-exam wrappers". These wrappers are reflective exercises that help students think about how they studied and how to improve their grades. But there is a lot of controversy about if they really work. This review looks at different studies to see when and how these wrappers actually help students in biology classes. It considers things like student differences, how the class is set up, and how often the wrappers are used. By looking at all these factors, we can figure out the best ways to use these wrappers to help students develop learning skills.

Introduction

Life-long learning, self-directed learning, and reflection are key elements within the field of biology education. However, a lot of students still do not have the capability for these cognitive processes, like reflective thinking and metacognition (Pate et al., 2019). Metacognition refers to the awareness and understanding of one's thought processes. Students can develop metacognitive skills through the use of educational resources and reflection instruments (Ratnayake et al., 2023). Research by Martín et al. (2000) has shown that most biology students exhibit what the researchers describe as "weak" restructuring. This term refers to a gradual and incremental process of modifying existing knowledge, rather than making significant and abrupt changes. In essence, many students struggle with restructuring their knowledge. Students who were able to apply these significant abrupt conceptual changes were students with great metacognitive skills. They showed high self-knowledge and a strong capacity to monitor their self-learning and control it, which is the backbone of self-regulated learning. In addition, students who were able to reflect on their study strategy and indicate plans can shift their study-related behaviours from passive to active techniques (Ratnayake et al., 2023). As reflective thinking and metacognition are important for self-regulated learning and not many students are capable of it, a challenge is faced (Pate et al., 2019)

To be efficient in self-regulated learning, learners have to actively manage their thoughts, feelings, and actions that are focused on the accomplishment of their objectives. Through setting personal goals, students create self-oriented feedback loops through which they can evaluate their effectiveness and progress and adjust their strategies accordingly for better results. It's important to note that self-regulation doesn't solely rely on individual efforts; it also incorporates social learning, such as seeking assistance from peers, mentors, and instructors (Schunk & Zimmerman, 2011).

Yet, only a small fraction of students seek advice from instructors or advisors after an unsatisfactory performance (Rieger et al., 2023). A way in which teachers can help students obtain skills for self-regulated learning is to facilitate metacognitive practices that can aid in analysing unsatisfactory performances. In research by Medina et al. (2017), a lot of strategies in which teachers can facilitate their students to enhance metacognition are described. One of these strategies is by reviewing the exam. By combining the exam review with an exam wrapper (a tool commonly used to enhance metacognition) students are supported in their reflection of their skills. Exam reviews and exam wrappers are both educational tools utilized to support student learning and comprehension, particularly concerning exams. An exam review session typically occurs after the completion and grading of an exam. Led by the instructor or teaching assistant, students engage in a collective review of exam questions, answers, and explanations. The primary aim of an exam review is to offer clarity on concepts, reinforce learning, and provide feedback to students regarding their performance. By dissecting the exam, students gain insights into their errors and areas for improvement, facilitating their understanding of the material. While exam reviews are valuable for comprehension, they predominantly focus on furnishing feedback rather than nurturing metacognitive skills (Silver et al., 2023).

In contrast, by combining the exam review with an exam wrapper, students are supported in the reflection of their skills. An exam wrapper is a reflective exercise completed after an exam. They are designed to stimulate the metacognition and self-regulation of students (Carpenter et al., 2020). It prompts students to introspect on their exam preparation, performance, and study techniques (Silver et al., 2023). Typical exam wrappers ask students how they studied, what they missed and why, and how they would change their study behaviours to improve on the next exam (Carpenter et al., 2020). Although the format of an exam wrapper can vary (Pate et al., 2019), the fundamental goal of exam wrappers is to cultivate metacognition by fostering self-awareness, critical evaluation, and adaptive learning strategies. By encouraging students to assess their learning processes and make adjustments, exam wrappers aim to enhance their overall understanding and performance in subsequent assessments (Silver et al., 2023).

In essence, exam wrappers serve as a bridge between assessment and learning, encouraging students to transition from focusing on performing well on exams to actively engaging in the process of learning and self-improvement. In short: **Exam wrappers stimulate the students to learn from the exam instead of learning for the exam.** A lot of disciplines in higher education already use them in a variety of forms (Pate et al., 2019). For example, teachers can automatically provide additional grades that inform students of their performance in four basic question categories that are related to levels in Bloom's taxonomy (Bloom, 1956; Rieger et al., 2023). Combining this with specific recommendations for how to improve in each of these categories, students can reflect on their learning strategies. It also gives the teacher insight into what the student(s) might be doing poorly in their studies (Rieger et al., 2023).

Even though the implementation of exam wrappers (in biology or other STEM courses) has been described a lot in literature, there is still a lot of variability in their impact on metacognition and academic performance. While some studies found no significant impact on exam scores or final grades, other studies showed that students who utilized exam wrappers showed improvement in their exam scores (Ridall, 2018). There has been little to no description of why these studies differ so much in effectiveness. That's why this review focuses on the question. **What factors may influence the effectiveness of post-exam wrappers for enhancing metacognition or course performance within biology education?**

To tackle this question, it is necessary to first describe the discrepancies in outcomes found within the literature. Secondly, to address the variations observed in the outcomes of previous studies, it's crucial to explore several factors:

Student characteristics, such as achievement levels, motivation, approaches, and gender are of relevance. These factors may influence students' receptiveness to reflective practices and their ability to utilize post-exam wrappers effectively.

Additionally, understanding the impact of the course itself, including the course environment and course layout is of the same importance. These factors can interact with the implementation of post-exam wrappers, shaping their effectiveness in different contexts.

By exploring these factors, this research tries to gain an understanding of the effectiveness of post-exam wrappers in different contexts within biology education.

Methods

A review of existing literature was conducted to explore the effectiveness of post-exam wrappers in enhancing metacognition and course performance within biology education. Relevant studies were identified through searches of the databases PubMed and Google Scholar. Keywords used were: 'post-exam wrappers', 'exam wrapper', 'metacognition', 'biology education' or 'STEM education' and 'student achievement levels'. In Google Scholar, the search gave 17.300 hits, while PubMed returned 8 hits. Data collection involved gathering information from peer-reviewed articles, scholarly journals, and publications focused on post-exam wrappers and biology education. As a criteria article published within the last 10 years were used for data extraction of factors within the student and course population and outcomes related to metacognition or course performance. Only articles and academic publications specifically focused on STEM courses, with a preference for biology courses, were selected for inclusion. This selection criterion was applied to frame the context within the field of biology education.

For this review the AI Petal was used to analyse some of the research papers on exam wrapper implementation, generating a table to visualize positive or negative effects of implementing exam wrappers. The results created by AI Petal were critically evaluated to ensure validation, as AI comes with ethical implications and biases. The AI-generated content was acknowledged in compliance with academic integrity standards (GSLs, 2024). The AI was only used to highlight factors among the papers and was not used to generate text.

Results

1 Variations in the effect of exam wrappers

The effectiveness of exam wrappers is a topic of considerable interest in STEM education. Some studies suggest positive outcomes, such as enhanced grades and metacognition (Yoder & Chenoweth, 2020; Sethares & Asselin, 2021), while other studies challenge their intuitive appeal (Stephenson et al., 2017; Vemu et al., 2022). Additionally, within biology education, the implementation of exam wrappers may lead to unexpected consequences, as seen in a study by Swalve et al. (2020) where grades declined after their introduction. A more detailed investigation into the different studies is important to understand the inconsistent results.

1.1 Positive effects

Studies have shown that exam wrappers are related to positive effects on course and exam grades, as well as metacognition, and noted changes in future study habits within STEM courses (Yoder & Chenoweth, 2020; Sethares & Asselin, 2021). A study by Gezer-Templeton et al. (2017) indicated that the majority of students reported that they believed exam wrappers positively impacted their study habits and exam scores. They also expressed intentions to continue using the exam wrapper process in future classes. These findings suggest that exam wrappers can enhance students' self-reported study habits. Exam wrappers were also linked to higher rates of gathered tests and improved scores on a course evaluation question that assessed the fairness of the evaluation methods (Stephenson et al., 2017).

1.2 Neutral/No effect

Exam wrappers have intuitive appeal as reflective exercises to improve student performance on assessments. However, some studies suggest that this intuition may not be correct (Stephenson et al., 2017; Vemu et al., 2022); they show no effect. A study by Soicher & Gurung (2016) found no improvements in final grades or metacognitive ability as measured with the Metacognitive Awareness Inventory (MAI) by implementing exam wrappers. All students showed an increase in their metacognitive ability over the semester, regardless of the use of exam wrappers (Swalve, et al., 2020).

1.3 Negative effects

Besides studies that show positive or neutral effects of exam wrappers, some studies show negative results of exam wrapper implementation. Swalve et al. (2020) showed that grades in an introductory biology course were notably lower in the year when exam wrappers were implemented compared to previous years (Figure 1). Initially, students reported similar attitudes towards factors such as the perceived difficulty of the course and their level of preparedness. However, perceptions shifted by the end of the course, with the course *Introduction to Biology* being perceived as more challenging when the exam wrappers were used. In contrast, *Chemical Analysis* showed minimal change in perception and grades throughout the term. This shift in perception was also reflected in other aspects of the *Introduction to Biology* course. By the end of the semester, there was a significant reduction in the percentage of students who believed they could earn an A or B grade, indicating a decrease in confidence regarding their academic performance. Specifically, the percentage of students responding positively to this question dropped to 37.4% in Biology, after the introduction of the exam wrappers, marking a 48.8% reduction compared to previous years without the wrappers.

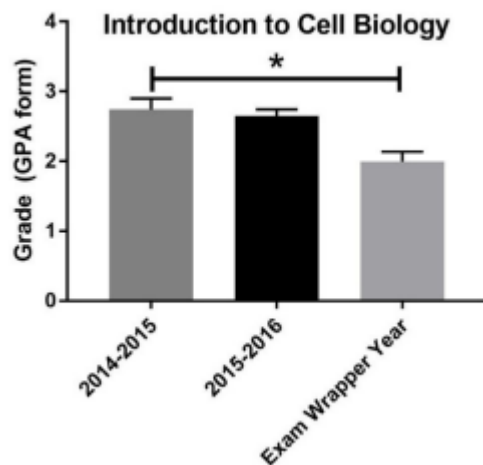


Figure 1: Average grade in GPA (grade point average) form in the exam wrapper year and previous years in an introductory Cell Biology Course. Stars indicate a $p < 0.0001$ Tukey's post hoc comparisons show that the exam wrapper year was significantly different from both previous years' data ($p \approx 0.0013$ and $p \approx 0.0002$ respectively) while the previous two years' data did not differ significantly from one another ($p > 0.05$). Figure obtained from Swalve et al. (2020).

2. Student-related characteristics

Appropriate metacognitive strategies are a significant predictor of academic achievement. However, this prediction is influenced by differences among student-related characteristics, like gender or their socioeconomic status (Callan et al., 2016). Post-exam wrappers could also differentially affect performances depending on differences in students entering courses and their perceptions of themselves or the course (Swalve, et al., 2020). Since student characteristics are important for academic performance and metacognition, these factors may influence the effect of exam wrappers.

2.1 Achievement levels

One of the student-related characteristics that could influence the impact of exam wrappers is the achievement levels. Research indicates that student achievement levels have a substantial influence on post-test performance, with a lot of students, particularly those with lower achievement levels, demonstrating difficulty in effectively acquiring content knowledge within academic activities (Simons & Klein, 2006). Research by Yoder & Chenoweth (2020) shows that in terms when an exam wrapper was present, there was a better performance on the final exam compared to the first exam than in terms without the exam wrapper. This relative improvement was the greatest for low-scoring students. Exam wrappers could potentially offer significant benefits for low-achieving students. Additionally, another study by Hodges et al. (2020) demonstrates that first-year students with higher scores on the first two exams were more likely to use the first two exam wrappers than those with lower exam scores.

So while low-achieving students benefit the most from post-exam wrappers, high-achieving students were the ones that were most likely to engage with the post-exam wrappers (Simons & Klein, 2006). It could be possible that high-achieving students feel more successful than low-ability students, which may lead to a greater likelihood of wanting to work on this type of study material again and reflect on their study performance. In the study of Simons & Klein (2006), this hypothesis seems to be consistent with the responses of lower-achieving students on an open-ended attitude survey. The students showed fewer responses for specific content learning as something they liked about the project, compared to high-achieving students. So while high-achieving students are the ones who show more engagement with post-exam wrappers, low-achieving students could benefit more from the wrappers in terms of improving their academic performance.

2.1.1 Gender

Gender has a significant effect on knowledge construction (Martín et al., 2000). Female students, both first-years and continuing students, show higher GPAs in general (Hodges, et al. 2020) and they are also better at visualizing interactions and concepts (Martín et al., 2000). In the study of Hodges, et al. (2020), there was a significant positive relation between exam wrapper use and grade point averages. This was both significant for male and female students. Yet, female students were more likely to make use of the offered exam wrappers. Female students also showed higher MAI scores, a self-report questionnaire designed to measure students' awareness and understanding of their metacognitive processes. This indicates that female students show higher levels of metacognition and that could potentially explain the higher grades (Hodges, et al. 2020).

2.2 Motivation

Motivation serves as a powerful force in education, as it is intertwined with student engagement and academic success. Yet, when faced with challenges or low grades, students may find themselves disheartened, and their motivation dwindling (Simons & Klein, 2006). This lack of motivation can affect their willingness to actively use exam wrappers. For self-regulated learning, proactive engagement is the key. Students must take the initiative to set goals and actively participate in the self-regulatory cycle. However, this proactive stance is influenced by motivational beliefs. If motivated, students are more likely to persevere through challenges, maintain a positive outlook, and actively seek growth opportunities (Schunk & Zimmerman, 2011).

The research conducted by Grandoit et al. (2020) revealed that numerous of their students demonstrated the capacity to assess their study habits and academic performance critically in the exam wrappers; Yet, these efforts did not yield positive effects on their academic performance, particularly among academically weaker students. Unsatisfactorily performing students, on the whole, did not tend to overestimate their abilities. This aligns with the Kruger-Dunning effect, which suggests that individuals with lower abilities may have more accurate self-assessments. In contrast, satisfactorily performing students were found to underestimate their abilities, indicating a lack of confidence or awareness in their capabilities. Additionally, "risk" students who eventually withdrew from the course were somewhat likely to overestimate their abilities. This highlights the discrepancy between students' ability to assess their study

habits and performance using exam wrappers and the actual improvement in academic outcomes, particularly among weaker students.

Another problem is that some students may lack the motivation to use an exam wrapper (Ridall et al., 2018). In the study of Grandoit et al. (2020), all students displayed a lack of motivation when reviewing their answers and deriving insights from their mistakes, irrespective of achievement levels. This lack of motivation was measured by the infrequency or absence of reported plans to adopt specific behaviours known to enhance learning outcomes. These behaviours included actions such as joining a study group, attending office hours with instructors, reviewing lecture notes outside of class, checking one's answers against a key after completing assignments or exams, and actively learning from mistakes made in coursework. The rarity or non-existence of these plans among students of all categories suggests a general lack of initiative-taking engagement with strategies that could potentially benefit their academic progress. However, another study showed that students who were struggling academically showed effective outcomes when implementing exam wrappers (Ridall, 2018).

2.3 Other student characteristics

The study of Butzler, (2016) showed that mathematics level and high school class rank emerged as significant predictors of overall grade in a college chemistry course, irrespective of the learning environment. Moreover, the effectiveness of exam wrappers appears to be influenced by the instructional approach employed in the classroom (see also section on course-related factors below). For instance, the study by Butzler (2016) compared the effectiveness of implementing notetaking and exam wrappers in both flipped and non-flipped classroom settings. They found that, in a flipped classroom learning environment, the addition of notetaking and exam wrappers did not lead to significant differences in individual assessment outcomes. However, a closer examination reveals that students graduating in the top, middle, and bottom third of their high school class witnessed increases of 7%, 3%, and 6%, respectively, in their overall course grades when using exam wrappers (Butzler, 2016). This suggests that while the overall impact of exam wrappers may not be uniform across all students, those with varying academic backgrounds and learning needs may benefit to differing extents from the implementation of reflective strategies.

3 Course related factors

Besides the fact that student-related characteristics could influence the effectiveness of exam wrappers on student performances and metacognition, the course itself is a crucial factor as well. There is the suggestion that post-exam learning effects, the learning from the exam rather than from the exam, could be different depending on the features of the courses themselves. This could be explained by differences over the years and changing perceptions by the students of the course, among several other explanations (Swalve, et al., 2020). Researchers and instructors determine for their classroom whether and how the metacognitive wrappers will be applied in their classrooms. Consider that metacognitive reflection depends on the instructor and student, they will not work the same for everyone, as the results may vary with the type of class (e.g. biology), semester (at the start or the end of the year), and grade level

(first year or continuing students) (Chambers, 2020). In addition, beyond variations in student characteristics and course dynamics, there is also an element of frequency that influences how effective exam wrappers are in improving both performance and metacognition (Ridall, 2018). In other words, the effects of exam wrappers cannot solely be attributed to student characteristics, but also to how and when they are applied and the extent to which they're used.

3.1 Frequency of Exam Wrappers

Students report that one of the motivational influences behind completing the exam wrappers was because they believed that the previous exam wrapper helped them improve their score for the next exam (Gezer-Templeton et al., 2017), suggesting that doing more exam wrappers may increase their effectiveness. As previously discussed, female students were more likely to make use of exam wrappers. Yet, there is some indication that increasing the frequency of offered exam wrappers is particularly helpful to male students (Hodges, et al. 2020). In their literature review, Chambers (2020) investigated whether increasing the number of and varying the type of metacognitive wrappers can augment academic performance and metacognitive awareness. They did not show any significant difference in the improvement of academic performance and metacognitive skills in relation to an increase in frequency.

Yet the repeated application of exam wrappers did assist students in perceiving this self-reflection tool as a method to enhance not just their study routines, but also their performance on exams (Gezer-Templeton et al., 2017). Contrasting to the research of Chambers, the research of Hodge et al. (2020) discussed that at least first-year students did benefit from increasing the frequency of the exam wrappers. When offering more exam wrappers throughout the year, the students also completed a higher percentage of the exam wrappers. Higher completion of the exam wrappers also was related to the student's cumulative grades and higher MAI scores. These outcomes also suggest that implementing more exam wrappers in (more than) one course could impact a student's performance in more than only the course they have been offered (Hodge et al., 2020).

3.2 Course Environment

Education types also play a role in shaping the effectiveness of implementing exam wrappers, as shown by recent research findings. Exam wrappers have demonstrated improvements in students' assessment performances, particularly in smaller classes (Gezer-Templeton et al., 2017; Soicher & Gurung, 2017). In smaller classrooms, the instructors were able to offer one-on-one support, but this was not the case for larger introductory classes (Ridall, 2018). In larger courses, the effectiveness of exam wrappers depends on offering additional motivational support for the students to complete the wrappers and to develop the metacognitive awareness to appreciate their value (Carpenter et al., 2020).

3.3 Timing of the course

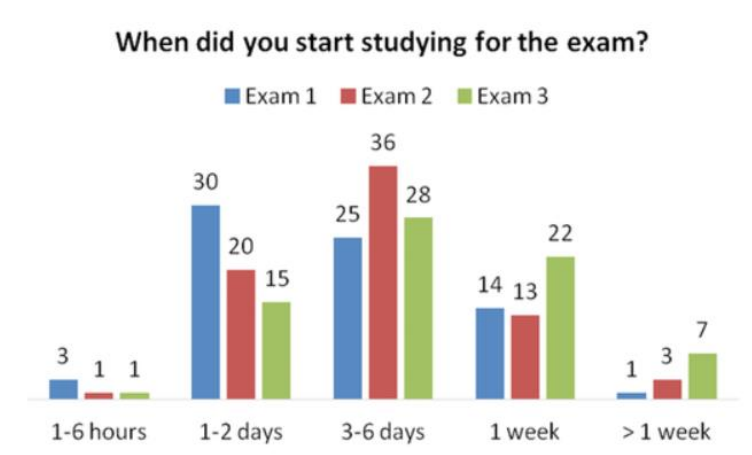


Figure 2: Frequency of students who started preparation at different time intervals for exams 1,2 and 3 during the semester. Numbers on the bars indicate the number of students who started studying for the exam in the given time interval. Exams 1, 2, and 3 were given at different time points during the semester, respectively. Figure obtained from Gezer-Templeton et al. (2017)

Not only do the number of exam wrappers implemented and the course environment affect the effectiveness of implemented exam wrappers on the metacognition of the students or promoting their grade, but also the placement of the course and the exam wrapper throughout the year. We have seen exam wrappers were more promising to be effective in enhancing metacognition among first-year students (Hodges, et al. 2020). But the same thing goes for restructuring knowledge by the students. For instance, it has been shown that the first six weeks of a biology course had the most radical restructuring of biology knowledge (Martín et al., 2000). Therefore, strategically placing exam wrappers during this critical period of knowledge restructuring, may maximize their effectiveness by facilitating the reorganization of students' understanding. When examined on an individual basis, most students reported spending more hours on exams that pose the greatest challenges through the course or year. Students may find themselves contending with a lot of competing stress factors, such as assignments and exams for other classes, which may limit the time available for studying for the exams (Figure 2) (Gezer-Templeton et al., 2017). Consequently, despite the implementation of exam wrappers, students may struggle to achieve significant improvements in their grades and metacognitive skills if they are unable to dedicate sufficient time and attention to the reflective process encouraged by the wrappers. As the learning experiences in the initial stages of education often set the tone for subsequent periods, it becomes clear that exam wrappers have to be strategically implemented at optimal times, while ensuring continuity in their implementation throughout the academic term.

3.4 Using incentives

In some studies, students were able to choose whether or not they would engage with the exam wrappers. Yet, this self-selection often leads to low completion rates of exam wrappers (Soicher & Gurung, 2017; Carpenter et al., 2020). Offering incentives, like offering extra credits or requiring exam wrappers for completion of the course leads to stronger effects on completion rates (Hodges, et al. 2020). Overall, biology classes contain quite a large number of students (Universiteit Utrecht, z.d.). In large courses, the success of exam wrappers depends on providing incentives for students to obtain the motivation to complete it and should provide additional support for students to develop and appreciate the value of metacognitive awareness (Carpenter et al., 2020). Students, especially first-years, also needed extra enticement to complete the exam wrappers when their performances did not meet expectations (Hodges et al. 2020). Thus, for some studies, these incentives were offered as extra credits (Gezer-Templeton et al., 2017). In these courses, most students noted extra credits as their motivational factor for completing the exam wrapper. For example, in the study of Gezer-Templeton et al., (2017), only half of the students stated that they finished the wrapper because they thought it might be helpful for future exam scores or preparation for future exams. As it seems that offering incentives provides more motivation for exam wrapper completion than the usefulness of the exam wrapper itself it is offered as a key to gain the motivation of the students.

Discussion

In education, one of the goals for students to gain academic success is obtaining metacognition and self-regulated learning skills. However, despite the recognized importance of these cognitive processes, many students still struggle to develop effective metacognitive strategies and reflective thinking skills. A method to gain metacognitive skills gaining attention in the literature is the use of post-exam wrappers in various disciplines, including biology education. These wrappers serve as reflective exercises designed to stimulate metacognition, encourage students to critically evaluate their learning processes, and improve their grades. While the implementation of exam wrappers holds promise, literature presents a diverse range of findings regarding their effectiveness in enhancing metacognition and course performance. This review aims to explore when post-exam wrappers are effective in enhancing metacognition and/or course performance in biology education, considering several factors influencing outcomes.

Overview of the findings

The effectiveness of exam wrappers in improving student performance and fostering metacognition varies across studies, with some demonstrating positive effects, others showing neutral or no effects, and some indicating negative outcomes. Positive effects include improved course and exam grades, enhanced metacognition, and positive changes in future study habits, particularly for students with average exam grades. However, conflicting evidence challenges the intuitive appeal of exam wrappers, with some studies showing no

improvement in final grades or metacognitive ability. In some cases, the implementation of exam wrappers may even lead to a decline in student grades, suggesting potential drawbacks.

Factors such as student-related characteristics (e.g., achievement levels, gender, motivation), course-related factors (e.g., course placement, course environment), and the frequency of exam wrapper implementation have been shown to influence the effectiveness of exam wrappers. Figure 3 displays the varied factors influencing the effective outcome of exam wrappers (e.g. improving metacognition and performance) (Table 1).

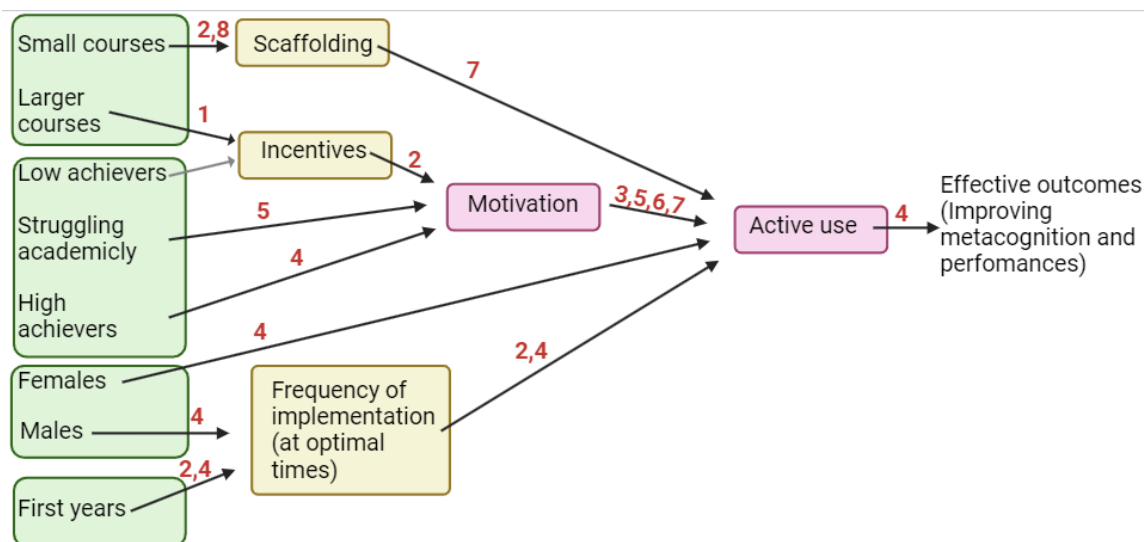


Figure 3: Different factors affecting the effectiveness of the implementation of exam wrappers Interactions are depicted with black arrows. Red numbers on top of the interactions indicate the corresponding research that has illustrated these interactions (Table 1). The gray arrow represents a hypothesized interaction. Various components are represented by green, yellow, and pink boxes symbolizing distinct categories such as course/student factors, strategies, and outcomes, respectively. The figure was made using BioRender ©.

Table 1: Corresponding research to the numbers in Figure 3

Number	Research
1	(Carpenter et al., 2020)
2	(Gezer-Templeton et al., 2017)
3	(Grandoit et al., 2020)
4	(Hodges et al., 2020)
5	(Ridall, 2018)
6	(Schunk & Zimmerman, 2011)
7	(Simons & Klein, 2006)
8	(Soicher & Gurung, 2016)

Exam wrappers appear to be most effective in small courses, where instructors can provide effective scaffolding to students (Gezer-Templeton et al., 2017; Soicher & Gurung, 2016). Instructors can offer one-on-one support, facilitating the development of students' metacognitive awareness and their appreciation of its value. This is less feasible in larger (introductory) classes due to the limited possibility of one-on-one interaction. By offering an incentive (such as extra credit or requiring the exam wrapper as part of the course), the course provides an external motivational factor for students (Carpenter et al., 2020; Gezer-Templeton et al., 2017). Both the scaffolding technique in small courses and the motivation through incentives in larger courses result in more active usage and a more effective outcome of exam wrappers.

In exploring the effectiveness of exam wrappers, it is notable that while they often prove more effective for low achievers, it is the high achievers who frequently utilize them more effectively, leading to improved metacognition and higher grades. Despite the demonstrated efficacy of exam wrappers for low-achieving students, they are not the primary beneficiaries of this resource. High-achieving students tend to exhibit more proficient utilization of exam wrappers, resulting in enhanced metacognitive abilities and academic performance (Hodges et al., 2020). Conversely, low-achieving students often display diminished motivation. Research indicates that lower grades often contribute to a more negative perception of the course and a lack of effective engagement with exam wrappers. Thus, a lack of motivation becomes apparent. One potential solution to this issue could involve the addition of incentives (Gezer-Templeton et al., 2017). Such incentives could, in turn, foster more effective outcomes in the implementation of exam wrappers. Lastly, there are students who are struggling academically. These students are mostly aware of their mistakes and aware that their current study techniques are ineffective. These students are motivated to alter their strategies to effectively employ metacognitive strategies (Ridall, 2018).

Student characteristics also play a significant role in the effectiveness of exam wrappers. Initially, female students appear to benefit the most from this intervention. Female students generally exhibit higher MAI scores and GPAs compared to their male counterparts (Hodges et al., 2020). Moreover, they tend to engage more actively with exam wrappers, positioning them akin to high achievers. Conversely, male students exhibit lower engagement with exam wrappers (Hodges et al., 2020). However, similar to low achievers, they seem to derive greater benefits when they do engage with them. Increasing the frequency of exam wrapper implementation appears to assist male students in promoting active use, thereby yielding the desired effects. This increase in frequency also appears to have a positive impact on first-year students. Additionally, first-year students are highly influenced by the timing of exam wrapper implementation. Optimizing both the frequency and timing of implementation maximizes the effectiveness of exam wrappers for first-year students (Gezer-Templeton et al., 2017; Hodges et al., 2020).

Limitations of the research

In addition to student and course-related factors, the effectiveness of exam wrapper applications within biology education faces a significant challenge posed by potential bias in the literature, primarily stemming from sample selection. Researchers often select participants based on convenience or availability, which can result in a non-representative sample. This bias may skew findings, as the selected participants may not accurately reflect the broader population of the classroom in which the exam wrappers are intended to be applied (Collins, 2017). Consequently, the generalizability and validity of study results may be compromised, emphasizing the importance of careful consideration (and transparency) in sample selection procedures within educational research.

Consent for participation

As a non-biased sample would be a random group of students completing the wrapper, obtaining this unbiased group would be of the biggest concern. As students have to give their consent for a study the sample selection process will be influenced (Gezer-Templeton et al., 2017). For example, in the study of Gezer-Templeton et al. (2017), only 83 out of 100 students gave their consent for the research. Students who actively participate in class discussions, engage with course materials, and seek opportunities for academic involvement may also be more likely to consent to participate in research studies. This can lead to a sample that overrepresents students who are more academically motivated or invested in their studies (Schunk & Zimmerman, 2011). Science has to respect confidentiality, anonymity, privacy, and non-traceability as they are ethical concerns (Cohen et al., 2007). On the other hand, students who have concerns about privacy or data security may be less likely to provide consent for the use of their data in research studies involving exam wrappers. This can result in a sample that underrepresents individuals who are more cautious about sharing personal information for research purposes. Students' consent to use their data in exam wrapper research can impact sample selection by influencing the composition of the participant pool based on factors such as engagement, motivation, and privacy concerns (Schunk & Zimmerman, 2011). The right of consent contributes to the already discussed effect of student-related characteristics on exam wrapper effectiveness. This problem is exacerbated when the data are archived for future use, as informed consent should also include the right not to participate or to withdraw at any time (Cohen et al., 2007). Researchers should be mindful of these potential biases and take steps to ensure that consent processes are transparent, respectful, and inclusive to promote the ethical conduct of research.

Self-selection

Not only consent but also participation in education students offers a big challenge for investigating the effectiveness of wrappers within biology education. Like most educational research, there is a selection bias, where students volunteer to participate, and thus represent a biased sample (Halmo et al., 2024). For example, Pate et al. (2019) had the exam wrapper and non-wrapper group self-selected. Here only 53 students participated in at least one wrapper and 35 never utilized a wrapper. As motivation comes in yet again as a factor, this self-selection contributes to the sample bias. Because the number of students who did not complete all exam wrappers was significantly smaller than those who did, comparisons between these different populations were not meaningful (Gezer-Templeton et al., 2017).

Randomly assignment and a placebo group

To achieve determination of the exact contribution of exam wrappers a controlled study is needed without sample biases, however, this would require a more intensive IRB approval process (a systematic review conducted by academic institutions or research organizations to ensure that research involving human subjects adheres to ethical standards and regulatory requirements) (Gezer-Templeton et al., 2017). Yet, to minimize this bias in group selection, students could also be randomly assigned an anonymized number at the beginning of the semester that divided them into control and experimental groups, while making the exam wrapper a qualification for completion of the course (Swalve, et al., 2020). Students in the placebo group (could) could be asked questions about their perceptions of the course and biology in general; and not the lost-marks, typical exam wrapper questions (Stephenson et al., 2017). Randomly assigning students to either the placebo or the exam wrapper group could potentially provide less biased results in determining the effectiveness of exam wrappers on metacognition and enhancing student performances. However, it is important to acknowledge that student consent remains a factor in such studies, thereby potentially remaining a group selection bias.

Implications of the findings

The findings of this study contribute to our understanding of the complex interplay between educational interventions like exam wrappers and student outcomes (e.g. enhanced metacognition and promoting academic performances), particularly in biology education. However, the effectiveness of exam wrappers is affected by various factors, including student characteristics and course environment. By providing incentives, increasing the frequency of the implementation of the exam wrappers, and scaffolding support, educators can enhance student engagement and maximize the benefits of exam wrappers, especially for low-achieving and first-year students. Additionally, understanding how gender differences can affect engagement of targeted interventions can help to promote active use among all students. Educators should alter their approaches to maximize learning outcomes for all of their students, considering their achievement levels, motivation, and gender differences. Moreover, the frequency and timing of exam wrapper implementation should be carefully considered to ensure optimal engagement and effectiveness. Incentives can be useful in

encouraging participation. Overall, while exam wrappers hold promise as reflective tools to enhance student learning and metacognition, their effectiveness depends on a range of factors. Further research is needed to explore these factors in more depth and to develop evidence-based strategies for maximizing the benefits of exam wrappers in educational settings.

Further Research

With the numerous factors discussed in this study, the factors influencing the effectiveness of exam wrappers have not been fully explored. Beyond examinations, there are other components on which course grades depend, including assignments and presentations. These activities, whether within the course curriculum or extracurricular, may require additional attention to metacognitive instructions (Hodges et al., 2020). Additionally, research has indicated that transfer or non-matriculated students were identified as those who often experienced academic failure or had higher risks of dropping the course (Grandoit et al., 2020). This group of students and these activities have not been discussed in this paper, and future studies should take these into account.

In addition to the extra factors that may influence effectiveness, there are also significant differences in education systems around the world. The studies used for this review come from various continents. These diverse educational backgrounds may also affect outcomes, particularly when comparing studies. For instance, American education standards differ significantly from those in the Netherlands, resulting in additional variation. Subsequent research should focus on education in a specific region to improve quality, but the quantity of educational research on exam wrappers within STEM subjects still hinders this. This limitation also applies to students with different socioeconomic backgrounds (Callan et al., 2016). Future research on exam wrappers within STEM subjects should consider and address these differences in education systems across geographical regions and socioeconomic backgrounds, to enhance the quality and applicability of findings.

Furthermore, in Figure 3, a grey arrow is indicated, representing the connection between low achievers and incentives. The study suggests that incentives could serve as motivation for low achievers to engage more actively with exam wrappers. However, this aspect has not yet been described or investigated in the literature. Further research on motivation among low-achieving students, with and without incentives, could provide insight into this matter. As motivation will always remain a crucial factor influencing the effectiveness of exam wrapper implementation, it is important to investigate further possibilities to gather and retain student engagement and participation.

Conclusion

In summary, the effectiveness of post-exam wrappers in biology education varies based on multiple factors such as student characteristics, course environment, and optimized implementation. High achievers tend to benefit more, while incentives may enhance engagement for low achievers. Optimizing implementation frequency and timing can maximize benefits, particularly for first-year or male students. Motivation will always remain a crucial factor influencing the effectiveness of exam wrapper implementation, underscoring the need for approaches to gain student engagement and participation. Despite promising outcomes, biases in sample selection and the need for further research remain challenges. Overall, while exam wrappers show potential for improving metacognition and academic performance, careful attention to implementation factors is necessary for their success.

Literature

1. Bloom, B.S. (1956) *Taxonomy of Educational Objectives, Handbook: The Cognitive Domain*. David McKay, New York.
2. 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>
3. Callan, G. L., Marchant, G. J., Finch, W. H., & German, R. L. (2016, October). Metacognition, Strategies, Achievement, and Demographics: Relationships across Countries. <https://eric.ed.gov/?id=EJ1115048>
4. Carpenter, T. S., Beall, L. C., & Hodges, L. C. (2020). Using the LMS for Exam Wrapper Feedback to Prompt Metacognitive Awareness in Large Courses. *Journal Of Teaching And Learning With Technology*, 9(1). <https://doi.org/10.14434/jotlt.v9i1.29156>
5. Chambers, J. M. (2020). Amp up the wrappers: Multiple Metacognitive Wrappers do not improve Student Academic Performance nor Metacognition in a Single-Course Intervention, Appalachian State University In School of Graduate Studies https://libres.uncg.edu/ir/asu/f/Chambers_Jessica_December%202020_Thesis.pdf
6. Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (8th edition). Routledge. <https://doi.org/10.4324/9780203029053>
7. Gezer-Templeton, P. G., Mayhew, E. J., Korte, D. S., & Schmidt, S. J. (2017). Use of Exam Wrappers to Enhance Students' Metacognitive Skills in a Large Introductory Food Science and Human Nutrition Course. *Journal Of Food Science Education*, 16(1), 28–36. <https://doi.org/10.1111/1541-4329.12103>
8. Grandoit, E., Bergdoll, R., Rosales, E., Turbeville, D., Mayer, S., & Horowitz, G. (2020). Exploring Organic Chemistry I Students' Responses to an Exam Wrappers Intervention. <https://eric.ed.gov/?id=EJ1254350>
9. GSLS. (2024). Guidelines for Using GenAI in Research Projects within the GSLS Framework. <https://students.uu.nl/sites/default/files/GSLS%20GenAI%20Research%20Guidelines%20for%20Students%20and%20SupervisorsGSLS%20GenAI%20Research%20Guidelines.pdf>
10. Halmo, S. M., Yamini, K. A., & Stanton, J. D. (2024). Metacognition and Self-Efficacy in Action: How First-Year Students Monitor and Use Self-Coaching to Move Past Metacognitive Discomfort During Problem-Solving. *CBE- Life Sciences Education*, 23(2). <https://doi.org/10.1187/cbe.23-08-0158>

11. Hodges, L. C., Beall, L. C., Anderson, E. C., Carpenter, T. S., Cui, L., Feeser, E., Gierasch, T., Nanes, K. M., Perks, H. M., & Wagner, C. (2020). Effect of exam wrappers on student achievement in multiple, large STEM courses. (z.d.). *Journal Of College Science Teaching*, 50(1), 69–79., 50(1), 69–79. <https://www.jstor.org/stable/27119232>
12. Martín, B. L., Mintzes, J. J., & Clavijo, I. E. (2000). Restructuring knowledge in Biology: cognitive processes and metacognitive reflections. *International Journal Of Science Education*, 22(3), 303–323. <https://doi.org/10.1080/095006900289895>
13. Medina, M. S., Castleberry, A. N., & Persky, A. M. (2017). Strategies for Improving Learner Metacognition in Health Professional Education. *American Journal Of Pharmaceutical Education*, 81(4), 78. <https://doi.org/10.5688/ajpe81478>
14. OpenAI. (2022). AI Petal [Artificial intelligence tool]. Retrieved from <https://www.petal.org/reference-manager>
15. Pate, A. N., Lafitte, E. M., Ramachandran, S., & Caldwell, D. J. (2019). The use of exam wrappers to promote metacognition. *Currents in Pharmacy Teaching And Learning*, 11(5), 492–498. <https://doi.org/10.1016/j.cptl.2019.02.008>
16. Ratnayake, A. W., Bansal, A., Wong, N., Saseetharan, T., Prompiengchai, S., Jenne, A., Thiagavel, J., & Ashok, A. (2023). All “wrapped” up in reflection: supporting metacognitive awareness to promote students’ self-regulated learning. *Journal Of Microbiology & Biology Education*. <https://doi.org/10.1128/jmbe.00103-23>
17. Ridall, A. M. (2018, 8 December). Rethinking Metacognitive Intervention: a scaffolded exam wrapper strategy. <https://soar.suny.edu/handle/20.500.12648/5236>
18. Rieger, G. W., Mclver, J., Mazabel, S., Cooper, S., & Lichtenberg, G. (2023). Getting More Out of Midterm Assessments. *The Physics Teacher*, 61(3), 207–209. <https://doi.org/10.1119/5.0075793>
19. Ruban, L. M., & Reis, S. M. (2006). Patterns of self-regulatory strategy use among low-achieving and high-achieving university students. *Roeper Review*, 28(3), 148–156. <https://doi.org/10.1080/02783190609554354>
20. Schunk, D. H., & Zimmerman, B. (2011). *Handbook of Self-Regulation of Learning and Performance*. Taylor & Francis.
21. Sethares, K. A., & Asselin, M. E. (2021). Use of Exam Wrapper Metacognitive Strategy to Promote Student Self-assessment of Learning. *Nurse Educator*, 47(1), 37–41. <https://doi.org/10.1097/nne.0000000000001026>
22. Silver, N., Kaplan, M., LaVaque-Manty, D., & Meizlish, D. (2023). *Using Reflection and Metacognition to Improve Student Learning: Across the Disciplines, Across the Academy*. Taylor & Francis.

23. Simons, K. D., & Klein, J. D. (2006). The Impact of Scaffolding and Student Achievement Levels in a Problem-based Learning Environment. *Instructional Science*, 35(1), 41–72. <https://doi.org/10.1007/s11251-006-9002-5>
24. Soicher, R. N., & Gurung, R. A. R. (2016). Do Exam Wrappers Increase Metacognition and Performance? A Single Course Intervention. *Psychology Learning & Teaching*, 16(1), 64–73. <https://doi.org/10.1177/1475725716661872>
25. Stephenson, B., Craig, M., Zingaro, D., Horton, D., Heap, D., & Huynh, E. (2017). Exam Wrappers. SIGCSE '17: Proceedings Of The 2017 ACM SIGCSE Technical Symposium On Computer Science Education. <https://doi.org/10.1145/3017680.3017701>
26. Swalve, N., Harwood, A. D., & Calhoun, E. S. (2020). Use of Exam Wrappers and Measures of Anxiety on Class Performance in Six Gateway STEM Courses at a Small Liberal Arts College. *College Teaching*, 69(3), 138–149. <https://doi.org/10.1080/87567555.2020.1833826>
27. Universiteit Utrecht. (z.d.). Universiteit Utrecht. Geraadpleegd op 15 april 2024, van <https://www.uu.nl/organisatie/departement-biologie/onderwijs#:~:text=De%20opleiding%20Biologie%20aan%20de,studenten%20de%20grootste%20keuzemogelijkheden%20hebben>.
28. Vemu, S., Denaro, K., Sato, B. K., & Williams, A. E. (2022). Moving the Needle: Evidence of an Effective Study Strategy Intervention in a Community College Biology Course. *CBE- Life Sciences Education*, 21(2). <https://doi.org/10.1187/cbe.21-08-0216>
29. Yoder, J., & Chenoweth, S. F. (2020). Reflection for Recovery: Exam Wrappers in an Object-Oriented Software Development Course Help Struggling Students Improve Future Exam Scores. 32nd IEEE Intl. Conference On Software Engineering Education & Training. <https://doi.org/10.1109/cseet49119.2020.9206181>