Gibbons, R.E., Hokien, D., Hutson, B.L., Maness, H.T.D. (2025). The potential of exam wrappers in higher education assessment practice: Fostering self-efficacy through guided self-reflection. *Intersection: A Journal at the Intersection of Assessment and Learning, 6*(1). Early View.

The Potential of Exam Wrappers in Higher Education Assessment Practice: Fostering Self-Efficacy through Guided Self-Reflection

Rebecca E. Gibbons, Ph.D., Deborah Hokien, Ph.D., Bryant L. Hutson, Ph.D., Heather T.D. Maness, Ph.D.

Author Note

Rebecca E. Gibbons, https://orcid.org/0000-0002-8345-9938
Deborah Hokien, https://orcid.org/0000-0001-5452-692X
Bryant L. Hutson, https://orcid.org/0000-0003-3180-563X
Heather T.D. Maness, https://orcid.org/0000-0003-0373-3333

We have no conflicts of interest to disclose.

Intersection: A Journal at the Intersection of Assessment and Learning
Vol 6, Issue 1, 2025

Abstract: This literature review (n = 30) examines the potential of exam wrappers to rapidly improve student learning in higher education. Exam wrappers, introduced as pre- and post-test reflection prompts (though widely varied), encourage students to engage in self-assessment and self-regulated learning (SRL). The vast majority of studies exclusively involved Science, Technology, Engineering, and Mathematics undergraduate courses. Both quantitative and qualitative methodologies were implemented, but only five were randomized control trials. Findings indicate that exam wrappers can lead to improvements in metacognition, SRL, and academic achievement, particularly when coupled with SRL training. While the magnitude of these positive effects can vary depending on the academic context and student characteristics, there is evidence exam wrappers hold significant promise for enhancing teaching and learning. Future research should focus on refining exam wrapper design for varied contexts, documenting fidelity, and conducting rigorous studies across diverse disciplines to better understand their impact and optimize effectiveness.

Keywords: *Exam wrapper, self-efficacy, assessment, exam autopsy, meta-learning tasks, metacognitive wrapper*

Introduction

Self-efficacy, defined as a context-specific evaluation of one's belief in their ability to execute behaviors needed to produce desired outcomes, strongly influences student learning and achievement (Bandura, 1977). Students with robust self-efficacy approach difficult academic tasks as challenges to master rather than threats to avoid, sustain motivation despite setbacks, and utilize more effective self-regulated learning strategies (Bandura, 1993; Schunk, 1995). Students do not enter higher education with a high ability for the use of self-regulated learning strategies, so strategies to increase this ability via self-efficacy can provide avenues for improving learning in postsecondary contexts (Sebesta & Bray

Speth, 2017). Moreover, self-efficacy is contextually variable, and an individual student's self-efficacy could be strong in some course environments, but weak in others. It should not be assumed as strong in every course, even with past academic success. Given the impact of self-efficacy, a body of research in postsecondary education has explored the use of techniques to cultivate students' confidence in their capabilities (Honicke & Broadbent, 2016). This work has found that intentionally working to facilitate an increase in student self-efficacy can help to increase disciplinary achievement (Honicke & Broadbent, 2016). Given the significant impact of self-efficacy, instructors should intentionally employ evidence-based techniques to cultivate adaptive confidence in students' capabilities.

Exam wrappers represent one promising instructional intervention to strengthen postsecondary students' self-efficacy. Exam wrappers are structured self-assessment activities focused on metacognitive monitoring and contribute to self-regulation. Exam wrappers have been structured in many ways but are frequently reflection prompt assignments given pre- and post-testing to encourage more effective preparation methods. It is ideal to incorporate self-assessment into the formative assessment feedback loop because engaging students in assessment emphasizes, and norms, the importance of using feedback in deliberate practice to improve learning (Niedwiecki, 2012). This structured literature review explores exam wrappers as a mechanism to enhance student self-efficacy and self-regulated learning. First, the conceptual background connecting self-efficacy, metacognition, and self-regulated learning is established. Next, the implementation of exam wrappers is described along with variations. Key studies on the impacts of exam wrappers are analyzed, focusing on self-efficacy and achievement outcomes. Finally, implications for practice and future research are discussed. This synthesis contends that by deliberately prompting self-reflection, calibration, and strategy refinement, exam wrappers allow responsive bolstering of student self-efficacy and self-regulated learning skills, leading to increased disciplinary learning.

As part of the **Grand Challenges in Assessment in Higher Education (GCA)** working group on the rapid improvement of pedagogy, and therefore student learning (Singer-Freeman & Robinson, 2020a), self-efficacy has become a recurring area of focus for the authors. This general review of the literature posits that within a single learning context (i.e., course semester, quarter, mini-mester, etc.), instructors can create an environment in which self-assessment is structured through exam wrappers (and their analogues), building students' self-efficacy and use of self-regulated learning strategies, therefore enhancing students' development of disciplinary knowledge (Singer-Freeman & Robinson, 2020a, 2020b). Calls for increased understanding of how exam wrappers influence self-efficacy and learning have already been articulated in nursing education and are extended to other disciplines herein (Poorman & Mastorovich, 2016; Williams, 2021). Equipped with more assessment insights, instructors can rapidly deploy exam wrappers to address disengagement and build students' beliefs in their capabilities.

Conceptual Framework: Self-Efficacy, Metacognition, and Self-Regulated Learning

Self-efficacy

Bandura's (1977) social cognitive theory emphasizes the role of self-referential thinking in human motivation and behavior. Self-efficacy refers specifically to individuals' beliefs that they can

successfully carry out courses of action required to achieve goals. Self-efficacy influences the challenges people undertake, the effort they expend, as well as their perseverance, resilience, stress levels, and performance attainments by shaping cognitive, motivational, affective, and decision-making processes (Bandura, 1977).

Self-efficacy is notably malleable and context-specific, often differing for a student in each course experience. Self-efficacy is developed through four sources: mastery experiences of success, social modeling of competencies, verbal persuasions of capability, and inferences from physiological states signaling competence or anxiety (Bandura, 1977). Mastery experiences of success include both positive and negative experiences, with relative impacts on self-efficacy. For example, past experiences in similar courses with high scores will subsequently positively influence efficacy when entering a course, but a low score on a lower-stakes assessment would have a relatively impactful negative influence (Bandura, 1997). Mastery experiences can be built into course structures through creating frequent opportunities for demonstrating mastery (Schunk & Ertmer, 1999). Social modeling of competence, also known as vicarious experience, comes from the observation of peers successfully completing other tasks, which can allow for efficacy development. This factor is most influential when students see similar-seeming peers' success (Adams, 2004). Verbal persuasions of capability, especially from perceived experts like faculty members, generates an environment in which efficacy can develop from trust. This goes beyond mere praise, which may damage self-efficacy, and should specifically focus on providing frequent feedback on their sustained efforts, to facilitate the efficacy-development process (Mueller & Dweck, 1998). Physiological inference is a straightforward developer of efficacy; this state is most frequently referenced in relation to test anxiety in the assessment context, where high heart rate feeds into perceptions of lower efficacy, and can be overcome with training in managing physiological states (Cioffi, 1991).

In academic settings, self-efficacy refers to students' appraisals of their abilities to fulfill course requirements, master content, and meet learning objectives. Robust self-efficacy promotes active learning strategies, intrinsic motivation, and deep engagement that produce achievement gains (Cleary & Zimmerman, 2006; Multon et al., 1991; Richardson et al., 2012). Conversely, fragile academic self-efficacy engenders task avoidance, disengagement, negative affect, and poor regulatory skills that inhibit success. Accordingly, cultivating adaptive confidence represents a pivotal instructional goal given self-efficacy's profound impact on students' discipline-specific learning and outcomes. Although self-efficacy is contextually variable, and an individual students' self-efficacy could be strong in some course environments, but weak in others, interventions, especially those informed by social cognitive theory, can improve self-efficacy during the course learning experience (van Dinther et al., 2011).

Metacognition

One strategy for cultivating self-efficacy in academic settings is increasing opportunities for student evaluation of mastery experiences of success, such as classroom activities that require metacognition. Metacognition describes the acts of monitoring and controlling one's cognition (Flavell, 1979). Key

metacognitive processes include connecting new information to prior knowledge, planning how to approach tasks, employing comprehension strategies, self-testing understanding, correcting errors, analyzing the effectiveness of techniques, and adjusting approaches (Ibabe & Jauregizar, 2010; Young & Fry, 2008). These reflective acts optimize learning by enhancing encoding, deepening understanding, identifying gaps, and improving recall. Students with robust metacognitive skills actively regulate engagement to maximize achievement (Broadbent & Poon, 2015; Coutinho, 2007).

Metacognitive monitoring provides the necessary feedback for informing self-efficacy beliefs (Bandura, 1977). Judgments of self-efficacy rely on appraisals of current performance capabilities and difficulties. Accurate monitoring information ensures appropriate calibration of efficacy beliefs. However, research reveals that students struggle with monitoring accuracy. Glenberg, Sanocki, Epstein, and Morris (1982) demonstrated the "illusion of knowing" phenomenon where students believe they understand the material better than they do due to metacognitive disconnects. Similarly, Hacker et al. (2000) found that lower-performing students often significantly overestimate their comprehension and exam readiness due to the inability to distinguish known from not-yet-learned content. Schneider et al (2014) explored self-evaluation capability in pharmacy students and discovered a similar trend, in which higher-performing students were more accurate in their prediction of which items on a high-stakes examination were answered correctly. This disconnect stems from failing to recognize the limits of their abilities. Kruger and Dunning (1999) extended these findings, showing that poor performers are typically least accurate in appraising their abilities. They found that building monitoring skills improved self-evaluations and achievement among struggling students. However, studies show that metacognitive monitoring, and subsequent academic achievement, can be improved through interventions (Nietfeld et al., 2006; Wagener, 2016).

Exam wrappers aim to strengthen metacognitive monitoring accuracy, planning, strategy use, perceived control, attributional beliefs, and self-efficacy—key drivers of achievement. Research demonstrates that guided reflection and self-explanation improve metacognitive monitoring precision (Callender et al., 2016; Dang et al., 2018; Knight et al. 2022). Metacognitive monitoring represents a central emphasis given its role as a gateway skill enabling effective self-regulation and calibration of self-efficacy. Flavell (1979) first highlighted the critical role of metacognitive knowledge and experiences in learning. Accurate monitoring allows learners to gauge understanding, identify knowledge gaps, and implement strategies to address weaknesses (Nietfeld et al., 2005).

Self-regulated learning

Self-regulated learning (SRL) involves proactively managing cognition, behavior, and motivation to successfully accomplish learning tasks (Pintrich, 2000; Zimmerman, 1990). Self-regulated students strategically set goals, select effective tactics, monitor progress, adjust approaches based on feedback, and reflect on performance. These processes underscore SRL's alignment with metacognition. Students with strong self-regulation display perseverance and resilience that yield positive achievement patterns. In the higher education context, the use of self-regulated learning strategies is associated

with higher achievement (Colthorpe et al., 2017). However, many students do not enter the higher education classroom with a high degree of self-regulation (Sebesta & Bray Speth, 2017), so, increasing students' self-regulation through classroom interventions can be anticipated to increase not only performance in the single course context, but across the postsecondary learning experience (Ergen & Kanadli, 2017).

The reflective steps encouraged by exam wrappers help students to identify learning gaps and highlight potential regulatory processes needing adjustment. In this way, exam wrappers encourage strategic planning and goal-setting for upcoming assessments, key components of SRL. Studies show that implementing study intentions and plans enhances self-regulation and performance (Gollwitzer & Sheeran, 2006; Schwinger & Otterpohl, 2017). Exam wrappers typically include forward-looking prompts to set specific preparation goals and develop implementation intentions centered on improved strategies (Colthorpe et al., 2017, 2019). This reflects Zimmerman's (1990) feedback loop model where students adjust tactics and efforts based on self-monitoring.

Exam Wrappers Implementation Features

Bandura (1993) contends that efficacy beliefs shape whether individuals can productively deploy regulatory processes. Conversely, employing self-regulatory strategies successfully strengthens efficacy beliefs. Given these interconnections, instructional interventions designed to increase self-efficacy are likely to work best when they incorporate metacognitive and self-regulatory skill-building. Exam wrappers embed prompts to apply metacognitive monitoring and SRL strategies through structured reflective activities. Lovett (2013) formally presented the exam wrapper concept, building on foundational work on reflection and SRL like Zimmerman's (1990) cyclical feedback model. Lovett (2013) outlines three reflective steps in exam wrappers: (1) Reviewing preparation approaches and estimations of performance; (2) Analyzing successes, errors, strategy effectiveness, and knowledge gaps; and (3) Setting specific goals and plans for the next assessment.

These three steps are designed to have the greatest impact on learning, and evidence supports the use of these steps. When guided by faculty in individual remediation sessions, for example, pharmacy students were found to have a significant increase in future content examination scores (Wang et al., 2018). Similar results have been found in the nursing context, where instructor-guided pre- and post-test reviews result in increased student confidence and use of self-regulated learning strategies (Poorman & Mastorovich, 2008; Tinnon, 2018). However, it is not always practical to engage in individualized feedback sessions with students. Exam wrappers incorporate the approach of encouraging students to reflect on their learning strategies through prompts guiding analysis of preparation, performance, errors, and strategy effectiveness, allowing them to be deployed at a larger scale. One of the relevant recommendations for the use of exam wrappers is their ability to achieve the kind of impacts that metacognitive self-regulation can provide in a relatively easy-to-use and scalable context.

As originally designed, exam wrappers are short activities, approximately 10 minutes, administered on paper via a series of metacognitive guiding questions that students complete after receiving the graded result of their first learning assessment and then as a preparation and reflection activity surrounding each assessment thereafter. For instance, Lovett's (2013) wrapper includes:

- How did you prepare for this exam? How much time did you spend?
- Were your exam preparation strategies effective? Why or why not?
- How will you change your study strategies for the next exam? Be specific.

These open-ended but directed questions are designed to stimulate independent, productive analysis and planning. **Appendix I** contains other examples of published exam wrapper prompts that can be adapted. After the first exam wrapper is completed, it is collected (to review for insights into potential adjustments to the course teaching plan) and redistributed at the beginning of the study period for the second exam. Another reflection prompt assignment is included to review the previous wrapper when they create their study preparation plan for the upcoming exam. Lovett (2013) notes that it is desirable to continue this reflection activity cycle throughout the course.

Literature Review Methodology

Identification of Articles

This review was inspired by a set of articles used as references during previous work by our GCA working group, which were identified through preliminary searches and expert recommendations. These 20 core articles (**Appendix II**) served as "seeds" for a chain-referral sampling approach (Figure 1), also known as snowball sampling (Biernacki & Waldorf, 1981; Goodman, 1961). This non-probability sampling technique involves using a small pool of initial sources to nominate other potential data sources that can be used in the research (Dusek et al., 2015). The reference lists of these initial articles were then scanned for additional relevant studies on the topic. This process was repeated iteratively, with each new set of identified articles serving as a source for further references, until a point of saturation was reached where no new significant studies were being discovered (Heckathorn & Cameron, 2017). Articles did not need to include the term "exam wrapper," as other descriptors were often used in our seeds, but it was the most common (29 uses when also including "examination wrapper"). Furthermore, not all studies used a concise term but provided a description that seemed relevant. Other unique terms observed included enhanced answer keys and reflection questions, exam autopsy, exam review assignments, Metacognitive Exam Preparation Assignments, meta-learning assessment tasks (2, same author), Post exam Review Activity, posttest reflection exercises, post-quiz reflection, self-reflection forms, Student Self-evaluation After Nursing Examinations, as well as these terms from later excluded studies: examination review, goal setting and activity report forms, independent post midterm question analysis exercise, metacognitive judgments, metacognitive learning strategies intervention, metacognitive monitoring with test-item confidence ratings, monitoring worksheet exercises, post-examination one-on-one remediation, Post examination review, post-test analysis, Pretest and Posttest Review, reflective test review, self-assessment exercises with

feedback, self-evaluation assignments, Strategic Resource Use intervention with pre- and post exam surveys, strategy microanalytic assessment, student-driven examination review process, and written exam corrections and peer-reviewed writing assignments including a metacognitive component.

While chain-referral sampling can be an efficient way to locate studies on a specific and emerging topic like exam wrappers, it is important to acknowledge potential limitations such as sampling bias and reduced generalizability (Etikan, 2016). To mitigate these concerns, efforts were made to include a diverse range of initial seed articles and to critically assess the representativeness of the final sample of studies included in the review. Our collection included articles sourced from ERIC, EBSCOhost, JSTOR, PsycINFO, SAGE, ScienceDirect, Springer, Taylor & Francis, and Wiley. Furthermore, for the studies with key findings identified (**Tables 1 & 2**), we additionally used Connected Papers (Tarnavsky Eitan et al., n.d.) to search for more articles, which did uncover additional studies of interest. In total, the collection of qualifying studies contained 58 unique article records, after duplicates were removed.

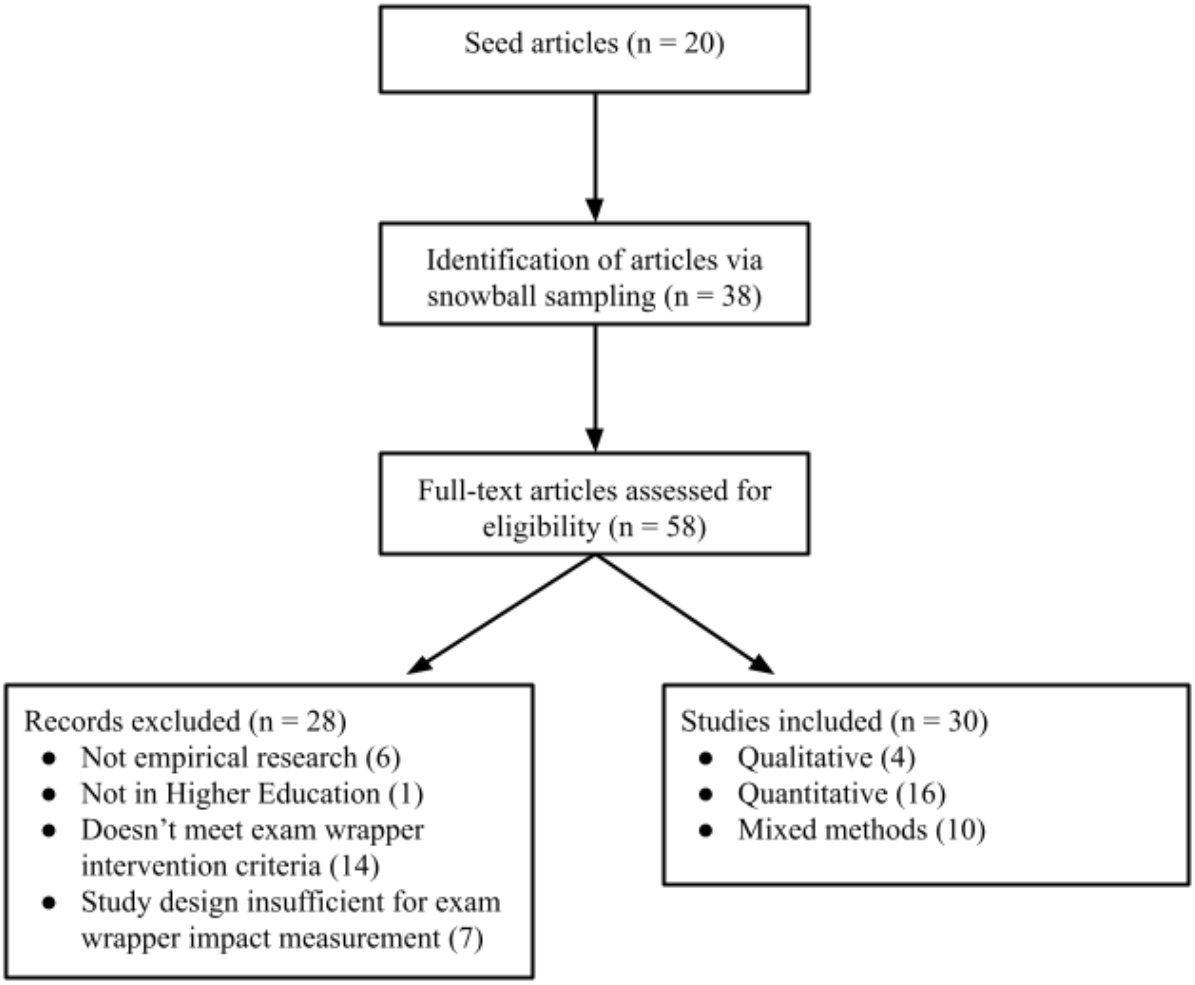
Eligibility Review

Screening based on title or abstract was impossible because of the variety in terms for this type of activity and the multitude of practitioner adaptations (see Variations below for more details). Instead, all 58 studies were evaluated based on exclusion criteria. During this phase, 28 records were excluded for one or more of the following reasons (ultimately categorized as excluded by the first appropriate reason on the list, for distinct counts in **Figure 1**, further detailed in **Appendix III**):

- Articles lacked empirical research (e.g., reviews).
- Studies conducted outside of higher education contexts.
- Lack of focus on exam wrappers. Interventions had to explicitly satisfy at least two basic elements of the protocol criteria established by Lovett (2013) for prompting independent self-reflection on assessment performance. These key elements are 1) self-reflecting on assessment preparation or goal-setting (Lovett's first and third questions) and 2) self-evaluating the effectiveness and weaknesses/errors on assessment performance (Lovett's second question). See **Table 1** in Variations and **Appendix IV** for details on how studies aligned.
- Study design was insufficient to directly measure exam wrapper impact to course-level test scores/grade performance or learning outcomes achievement (e.g., research focus was tangential to exam wrappers, behavior change investigated without correlation to performance metrics).

Figure 1

Article Review Process



Note: The remaining 30 studies represented a range of methodologies (**Appendix V**). The majority were quantitative (n = 16), with five being randomized control trials. Another 10 were mixed methods, combining quantitative and qualitative approaches for greater insights, and the remaining 4 were qualitative. The studies spanned multiple disciplines, with an overwhelming representation from Science, Technology, Engineering, and Mathematics (STEM) fields (n = 26).

Variations

The pioneering works, with the terminology of "exam wrapper" by Lovett (2013) and earlier works such as Achacoso (2004), provide guide rails for designing and implementing exam wrappers in the classroom to have a high degree of alignment with theoretical grounding. However, the literature examples of exam wrapper implementations vary in the administration of wrappers. The variations in implementation found in this review of the literature are summarized in **Appendix IV**. All but three of these 30 studies engage in the first three of four components in the exam wrapper protocol established by Lovett. Six studies followed all four Lovett principles. Most frequently, the implementation strategy of the instructor collecting and subsequently returning exam wrappers to students was not implemented, rather, students completed wrappers and kept them. Wrappers are most often implemented only after exams, with little reference to additional reflection activities using previous post-exam reflections before the next exam (n = 8). The number of wrappers and questions was highly varied, and concerningly, sometimes not specified. While exam wrappers were designed to be associated with major examinations, applications in literature associate wrappers with various kinds of assessments, from in-class quizzes to papers.

Other examples of important features of variations are the association of wrappers with credit, metacognitive training, and peer engagement. Some instructors distribute wrappers non-compulsorily for extra credit to encourage buy-in and avoid student resistance. In our review, grading varied from being required as a small proportion of the course grade to extra credit or for no credit. In total, eight studies had a graded approach, while nine provided extra credit (1 study used a mixed approach of these strategies), three were not graded, and nine had unclear grade relationships (grading impact was not addressed or was generalized without comparison to the impact on the final grade). Regarding metacognitive training or discussion to introduce the concept to students, 13 studies included this element. Some researchers combined explicit SRL training with wrappers to strengthen strategy implementation, such as individual sessions providing insights into metacognitive techniques or regularly embedded direct instruction on how to successfully regulate one's own learning. Studies also implemented wrappers as individual or cooperative activities during class or as independent graded out-of-class assignments; individual work was the most common approach. Some works noted the "exam autopsy" model which is similar to exam wrappers yet often specifically includes peer feedback to create a holistic post-exam environment promoting metacognition (Havis, 2019). However, only four studies addressed peer involvement in the wrapper activities. Additionally, in one study reviewed, but deemed not inclusive of sufficient wrapper criteria, additional peer support processes of the calibrated peer review (CPR) rubric were discussed (Mynlieff et al., 2014).

Table 1

Characteristics of Exam Wrapper Implementation in Key Impact Studies (n = 7).

Study	Measurement of Wrapper Impact	Similarity to Lovett (2013) Protocol [a]	Implementation: # of wrappers; # of questions; Pre-/Post-Exam (or both) use	SRL Training Inclusion	Relevancy to Grade
Angell et al. (2024)	Exam scores, prediction accuracy	1,2,3	3; 10-15; Pre	None	Required, points (equivalent to 3-5% of each exam grade) were awarded "as long as they responded thoroughly"
Chen et al. (2017)	Final grade, Self-Reflection on Learning (MSLQ[b]), negative affect[c], resource use, and perceived effectiveness of resource used [d]	1,2,3	4; varied, approximately 8; Both	Strategic Resource Use exercise	Voluntary, extra credit
Knight et al. (2022)	Quiz grades, prediction accuracy, Metacognitive Reflection Score	1,2 (but did not claim exam wrapper methodology)	6; 1-4; Both (quizzes)	None	Extra credit, 1 point per post-survey, 35% of students fully participated in post activities

LaCaille et al. (2019)	Quiz and exam scores, Knowledge of Cognition, Regulation of Cognition (MAI[e]), Competence (Perceived competence scale for learning[f])	1,2,3	15-16; 13; Post (quizzes and 1st exam)	Video series and instruction on deep learning & study strategies	Extra credit awarded based on completeness
Mutambuki et al. (2020)	Midterm and final exam scores, Final grade	1,2,3	2; 2; Post	Training in metacognition provided at beginning of course and reminders throughout	Unclear
Rosales et al. (2019)	Final exam score, course grades, study strategies (categorized via coding)	1,2,3	4; 3-5; Post (quizzes)	A syllabus handout on study tips was discussed in class	In-class, credit unclear
Stephenson et al. (2017)	Final exam score	1,2,3,4	1; 6-8; Post-Midterm Exam	Brief discussion when wrappers were returned	Voluntary; increase of 1/3-letter grade on midterm exam if "successfully" completed

Note:

[a]Alignment to the protocol criteria described in Lovett (2013) of prompting assessment performance self-reflection on 1) preparation methods, 2) error evaluation for effectiveness and weakness, 3) future planning and goal-setting, as well as 4) returning the exam wrapper to the student just prior to the next exam and encouraging review of their previous work and adherence to their preparation plan.

[b]Motivated Strategies for Learning Questionnaire (Pintrich et al., 1991)

[c]Adapted from Smith & Ellsworth (1987)

[d]Adapted from the Resource Questionnaire (Brown et al., 1996)

[e]Metacognitive Awareness Inventory (Schraw & Dennison, 1994)

[f]Perceived competence scale for learning (Williams & Deci, 1996)

Impact Studies

The empirical literature reveals modest yet promising effects of exam wrappers on metacognition, self-regulated learning strategy use, self-efficacy, and achievement, with minimal risks. Sethares and Asselin (2022) conducted a systematic review of exam wrapper studies across disciplines. They found sparse overall impacts, with some studies showing positive effects on test scores and metacognitive variables. However, the review noted a lack of rigorous controlled studies, which we saw as well.

Table 2 provides a summary of key findings from empirical studies investigating the impact of exam wrappers on student outcomes in higher education. A summary of the study design and key findings for all qualifying articles reviewed (n = 30) is available in **Appendix V**. Controlled studies on exam wrappers reveal some metacognitive, SRL, and performance advantages but inconsistent effects depending on contextual factors. This review of published studies reveals mixed implementation factors (see **Table 1**). These factors can influence the impact found. For example, Pate et al. (2019) found voluntary exam wrapper participation declined over a semester in a pathophysiology course. These findings suggest participation and proper scaffolding remain challenges in exam wrapper implementation. While much work reinforces the themes of at least modest positive impacts on grades, metacognition, and self-efficacy, the general body of literature on exam wrappers in higher education illustrates mixed results. Because of the range of participation rates and implementation factors, it is difficult to identify a single trend in the impact of exam wrappers on either academic achievement or affective characteristics.

Table 2

Summary of Key Findings from Exam Wrapper Impact Studies

Study	Context	Method	Results Summary
Angell et al. (2024)	Biology (N = 233)	Quantitative - Quasi-experimental	⇀ Pre-exam assignment wrappers provided moderate exam score increases for students with lower incoming standardized test scores ⟳ Exam score prediction accuracy was associated with incoming standardized test scores, not wrappers
Chen et al. (2017)	Statistics (Study 1: N = 178, Study 2: N = 203)	Quantitative - Randomized control trials, separate and pooled analyses (No significant differences in groups on pre-intervention measures of motivation, importance, & confidence)	↑ Final grades (4% increase overall and 8-10% increase after 2 wrappers; no effect differences in sex, race, class standing, performance cohort, GPA[a]) ↑ Self-reflection on learning score increased (Study 2 only, adapted 8-item metacognitive self-regulation subscale of MSLQ[b]) ↑ Perceived effectiveness of resource use (adapted measure) ↑ Affect toward the exams (adapted measure) ⟳ Variety use of resources (Intervention group used fewer resources; Actual use of planned resources was needed for impact, but self-reports of plan quality and adherence was not correlated.)
Knight et al. (2022)	Genetics (N = 496)	Mixed Methods - Descriptive Convergent Parallel Design	⟳ Frequency of metacognitive responses over time was not predictive of grade increase ⇀ Those who shift from overpredicting to matching or underpredicting performance had improved performance.

LaCaille et al. (2019)	Psychology (N = 244, intervention = 123)	Quantitative - Quasi-experimental	↑ Moderately higher quiz and exam scores for students reporting higher use of wrappers ⇀ Metacognition knowledge was higher than control, but not regulation (MAI[c]) ↑ Competence (Perceived competence scale for learning[d]) ↑ Students reported enjoying learning the material more strongly when using wrappers.
Mutambuki et al. (2020)	General Chemistry (N = 427, intervention = 239)	Quantitative - Quasi-experimental	↑ Exam 3 and final scores, and lower withdrawal rates
Rosales et al. (2019)	Organic Chemistry (N = 154, intervention = 71)	Quantitative - Quasi-experimental, interrupted time series design	↑ Final exam score ↑ 1 specific item % correct on final exam ↑ Course grades ↑ Withdrawal (vs failure) rates ⟳ Study strategies (active v passive)
Stephenson et al. (2017)	Computer Science (Study 1 N = 289, Study 2 N = 752)	Quantitative - Study 1 = Quasi-experimental, interrupted time series design Study 2 = Randomized control trial, plus open-ended survey items	⟳ Final exam score ↑ Students appreciated the wrappers and they reported increased use of active study strategies.

Note: Students are undergraduates unless otherwise noted. Symbol Key: Improvement: ↑, Improvement for only Some: ⇀, No improvement: ⟳

[a]Grade Point Average

[b]Motivated Strategies for Learning Questionnaire (Pintrich et al., 1991)

[c]Adapted from Smith & Ellsworth (1987)

[d]Adapted from the Resource Questionnaire (Brown et al., 1996)

[e]Metacognitive Awareness Inventory (Schraw & Dennison, 1994)

[f]Perceived competence scale for learning (Williams & Deci, 1996)

Impact on Non-Academic Constructs

Metacognition

A cornerstone aim of exam wrappers is strengthening metacognitive monitoring to enable effective self-regulation. Studies demonstrate predominantly positive impacts on reflection and awareness using wrappers, but this is not a universal finding. Thompson (2012) observed increased metacognitive self-regulation among first-year Spanish students receiving an exam wrapper intervention. Trogden & Royal (2019) similarly saw an increase in metacognitive awareness for the group of students participating in wrappers, including courses in both mathematics and chemistry. In a review of research studies across disciplines, including two in nursing education, Williams (2021) found evidence that exam wrappers which prompt students to reflect on test performance and study behaviors before reviewing exams can enhance calibration of knowledge. These findings suggest that exam wrappers may be particularly beneficial for improving metacognitive monitoring skills.

In many cases, training in metacognitive practice combined with wrappers, provided benefits to students. LaCaille et al. (2019) demonstrated modest positive impacts on metacognitive skills in undergraduate psychology classes using wrappers with corrective feedback. Nietfeld et al. (2006) found that students receiving monthly monitoring accuracy training and reflective wrappers were better able to calibrate exam score predictions and post-dictions than those using wrappers alone. Callender et al. (2016) demonstrated that a wrapper intervention with monitoring exercises yielded significant gains in calibration and achievement. In the implementation of “strategic resource use” guidance, undergraduate statistics students reported an increase in their ability to self-monitor their learning compared to a control group in the study by Chen et al. (2017).

Not all studies found a positive relationship between metacognition and exam wrapper completion. Hodges et al. (2020) observed limited metacognitive gains from multiple course wrappers. Soicher & Gurung (2017) found no differences in metacognition between control and wrapper conditions. In Nietfeld et al.’s (2005) study, students receiving in-exam confidence monitoring questions, rather than full wrapper assignments, were not able to increase their ability to engage in metacognitive monitoring. In Knight’s (2022) study, students who were better able to predict their exam scores over time had better performance; however, this was not directly correlated to the completion of metacognitive questions in a wrapper-style assignment. Furthermore Angell et al. (2024) found exam score prediction accuracy was associated with incoming standardized test score, not wrappers.

Self-Regulated Learning

Given the aim of improving students’ learning tactics, studies of exam wrapper implementation also explore the impact of wrappers on SRL strategy use. Results demonstrate increased planning and strategy refinement after implementing wrappers. Lovett (2013) documented STEM students’ open-ended reports of modified techniques like spacing study sessions after using exam wrappers. In another study, engaging students in planning study strategies before examinations via “meta-learning” activities akin to exam wrappers was demonstrated to help them to use more effective strategies

(Colthorpe et al., 2017). LaCaille et al. (2019) demonstrated increased strategic planning for tests among psychology students using the Learning and Study Strategies Inventory. Chew et al. (2016) found that students qualitatively reported increased planning and strategizing.

In the literature, there are mixed findings regarding the impact exam wrappers have on the enhancement of students' demonstration of SRL behaviors development. For example, Grandoit et al. (2020) found that only about half of the students who had committed to increasing study efforts in wrappers at the beginning of the semester reported actually engaging in such activities in later wrappers. Chen et al. (2017) had similar findings, with students in a treatment wrapper condition using no more resources than those in a control, and a negative relationship between the number of planned study strategies to those actually used, despite a high perception of resource use effectiveness in treatment students. LaCaille et al. (2019) documented higher metacognitive knowledge, but not regulation. Students in the nursing context reported via focus groups that they found direct feedback from the instructor to be more helpful in determining future study strategy planning rather than self-reflection via exam wrappers as well as finding more value in their pre-existing SRL strategies, indicating that training in metacognition might be an important component in exam wrapper efficacy (Schuler & Chung, 2019). In the undergraduate physiology context, many students reported not changing their strategies, instead relying on previous strategies that were easier and had worked for them in the past, despite being guided through metacognitive practices (Colthorpe et al., 2019). These findings indicate wrappers can prompt productive regulation improvements, but that additional support for students might be required to show value in new strategies and impact sustained behavior changes needed to gain maximum SRL benefits.

Impact on Academic Performance

The central aim of strengthening metacognition, regulation, and self-efficacy through exam wrappers is enhancing academic achievement. Naturally, this is the area in which most studies of exam wrapper efficacy are focused. While effects vary across disciplines (and most studies occur in STEM fields), overall patterns indicate potential learning benefits from wrappers. Chen et al. (2017) implemented wrappers in a randomized control trial, finding grade increases mediated by enhanced self-regulatory behaviors. Similarly, Edlund's (2020) RCT findings noted students with high critical thinking skills benefitted the most in exam score improvement after wrapper use. In chemistry courses, reflective wrappers related to improved final exam performance, amongst other benefits (Mutambuki et al., 2020; Rosales et al., 2019; Trogden & Royal, 2019). Hodges et al. (2020) observed modest grade improvements from multiple course wrappers. Butzlaff et al. (2018) found improvements in exam scores when using wrappers in undergraduate nursing. In a review of research studies across multiple disciplines, Williams (2021) identified evidence that exam wrappers can improve grades in nursing courses specifically. This finding, along with those from Butzlaff et al. (2018), suggests that the benefits of exam wrappers for academic performance extend to nursing education. Pate et al. (2019) found a low level of participation with an increase (although non-significant) in exam performance.

While exam wrappers as an addition to a course can provide positive impacts on academic performance, the further enhancement of metacognitive training shows increases as well. Zimmerman et al. (2011) redesigned math courses and trained instructors to complete an extensive checklist of SRL activities throughout the semester with improved exam performance results as well as on a standardized test. In general chemistry, when exam wrappers were one part of a substantial instruction on metacognitive engagement, Mutambuki et al. (2020) found significant gains in student performance on the final exam. LaCaille et al. (2019) and Chen et al. (2017) also saw positive impacts and included SRL training in their rigorous study designs. Furthermore, with the enhancement of peer feedback discussions and individual guidance sessions about how to regulate learning, Havis (2019) found the exam autopsy model (an exam wrapper adaptation) to be affiliated with significantly higher exam scores over the course of a semester.

In some cases, any positive effects on academic achievement were found to be the case for only certain groups of students. Butzler (2016) combined wrappers with other SRL encouragement and found top and bottom high school performers showed the greatest achievement gains, suggesting particular benefits for those with very high or low prior preparation. Colthorpe et al. (2017) found that, while the entire group participating in a wrapper activity had higher achievement scores, undergraduate students performing in the lowest levels had the greatest gains when adapting their study strategies as a result of the wrapper activity. Angell et al. (2024) also found improvements in performance for students with the lower incoming ACT scores. In the case of undergraduate business, Hartling (2022) found that those students scoring D or F, or, those students failing the course, had the greatest increase in scores associated with exam wrappers (although statistical significance was not tested). LaCaille et al. (2019) noted moderately higher quiz and exam scores for students reporting higher use of wrappers.

In contrast, Root Kustritz & Clarkson (2017) found no achievement differences in a veterinary course where there was minimal student participation in optional wrappers. Swalve et al. (2021) demonstrated mixed achievement results across STEM courses, with wrappers related to biology score decreases in one course but found wrappers boosted biology grades in others. Soicher & Gurung (2017) found no differences in achievement between control and wrapper conditions. Stephenson et al. (2017) found similar results in a random assignment condition as well as wrapper use when compared over past years without wrappers. Chew et al. (2016) found no quantitative performance improvements from engineering wrappers. Knight et al. (2022) could not correlate increased frequency of metacognitive responses with grade improvements. Furthermore, Andaya et al. (2017) reported a third of responses were low quality reflections with poor articulation of issues and a lack of appropriate, detailed plans for improvement, indicating a need for greater metacognitive support to achieve higher exam performance over time. In summary, multiple studies reveal tangible learning improvements, but effects are not universal. Based on the differences in wrapper implementation described in **Table 1** and treatment fidelity issues observed on our review, we posit that the efficacy of wrappers for improving self-regulated learning behaviors, and thus learning performance, directly depends on factors related to quality of implementation by instructors (e.g., gaining buy-in, prompts

used, integration into multiple course points throughout the term) and students (e.g., critical reflection, plan adherence).

Perceptions of Exam Wrapper Impact

Quantitative results showcase wrappers' impacts on measurable characteristics like metacognitive awareness and academic performance. In addition, qualitative data from impact studies provides insights into student and instructor perceptions of the effects of exam wrappers. In some cases, these perceptions are positive. For example, surveys consistently reveal that students view wrappers as beneficial for identifying gaps, planning improvements, and building confidence (Butzlaff et al., 2018; Gezer-Templeton et al., 2017). Students learn by using wrappers that their current study strategies may not be working and provide them with insights on how to improve in future semesters, potentially developing a growth mindset (Knight et al., 2022; Sabel et al., 2017). Wrappers can also be a venue for students to recognize the impact of their successful SRL activities (Chew et al., 2016). Evidence also indicates that students are better able to identify effective learning strategies after engaging in wrapper activities (Chen et al., 2017; Colthorpe et al., 2017; Craig et al., 2016; Trogden & Royal, 2019). In addition, students in one study reported liking exam wrappers and feeling as if they helped, even if the academic performance data did not support this claim (Stephenson et al., 2017). LaCaille et al. (2019) even noted students reporting they enjoyed learning the material more when using wrappers. Yet, El Bojairami & Driscoll (2019) point out that students with higher initial exam scores, seemingly of a self-satisfactory performance, often put forth little effort in their reflections and are less interested in the potential benefits of wrappers.

Discussion

The published literature demonstrates that exam wrappers, as a tool for student self-assessment and incorporated into traditional student learning outcomes assessment techniques like examinations and projects, can improve metacognition, SRL, and academic achievement for many students, especially in cases for those with traditionally lower performance. The extent of these effects varies with factors like discipline, instructional practices, and learner characteristics that require further investigation. Based on this analysis of the literature, we posit that exam wrappers hold promise to improve the teaching and learning environment of higher education and potentially bridge achievement gaps for students with varying incoming academic credentials. Wrappers and their analogues hold meaningful potential to strengthen student learning and success. However, this promise remains relatively poorly examined, with little standardization of exam wrapper implementation to allow for wide-scale empirical investigation.

Exam wrappers have been investigated in the higher education context under various names (i.e., exam autopsy, metacognitive wrapper) and with varying implementation contexts (e.g., quantity of wrapper assignments in the course, differences in prompts). It appears that instructors and researchers implement wrappers differently, concurrent with contextual circumstances. For example, in computer science, exam wrappers were implemented under an "after-action review" model that is seen in many firm's project implementations to suit career development (Davis, 2021). In another computer science

example, faculty devised customized exam wrappers to align with students' most-missed questions, potentially increasing the potential for positive impacts (Stephenson et al., 2017).

Overall, for broad claims of the efficacy of exam wrappers, there is a need to systematically investigate and refine the design, supports, and incentives needed to achieve consistent benefits. Mixed findings highlight areas for improvement through pedagogical innovation and further research. For example, there was a distribution of the implementation of exam wrappers with and without specific instruction on how to engage in metacognitive monitoring and other SRL strategies. Past research suggests that college students do not enter the higher education environment with a high degree of self-awareness sufficient to engage in SRL strategies (Sebesta & Bray Speth, 2017), so this kind of training might be needed to maximize the potential of exam wrappers. Positive results in several studies indicate that exam wrappers assisted some groups of students in recognizing the need for a new study preparation approach, but the extent to which wrappers are scaffolded with other training is not fully investigated in the literature.

Implications for Instruction

This analysis suggests that, while some criteria should be used to maximize the impact of exam wrappers based on their design (Lovett, 2013), instructors should adapt wrappers to their disciplinary and institutional context. Based on our reading of the evidence for exam wrapper efficacy in higher education and the theoretical background supporting by other empirical research, we propose several practical techniques instructors could apply in the classroom context to meaningfully impact student learning:

- Incorporate peer-peer interaction: Instructors can implement wrappers as individual or cooperative activities during class or as independent graded out-of-class assignments. Individual work was the most common approach (see **Table 1**), but it would seem more beneficial to have students work in small groups for the first step before the exam to talk through their post-analysis and form accountability partners for their future action commitments. For example, in informal observations in the nursing context, peer-affiliated exam review provided a friendly environment for students to process exam results (Ierardi, 2014; Masters, 2007).
- Teach students how to "think about thinking": Explicitly teach metacognitive and self-regulation skills consistently, not just through wrappers. Wrappers provide opportunities to apply acquired tactics. Metacognitive instruction can promote increased learning, so it stands to reason that incorporating explicit instruction can provide the boost exam wrappers need to significantly increase learning (Cook et al., 2013). This can be enhanced through adequate modeling, coaching, and feedback, as this kind of support moves skill-building beyond rote reflection.
- Adapt prompts to course content: The instructional technology environment creates many possibilities for creating exam wrappers that are deeply embedded within the context of student learning. In some of the exam wrapper literature, customized exam wrappers helped

students to identify their key needs for improving for future exams (Davis, 2021; Stephenson et al., 2017).

- Reward good practice: Consider incentives like incorporating wrapper completion into course credit or extra credit to encourage effort and student buy-in. Voluntary use of exam wrappers often declines over time (Root Kustritz & Clarkson, 2017). Other possibilities include pairing wrappers with the assurance of exam re-takes (or incorporating some "earned-back" credit) and adding tutoring, or other academic support, to enhance learning while building metacognitive and SRL skills. This will prevent motivational harm if efficacy wanes.
- Engage in the assessment process: When implementing exam wrappers for the first time, it would be valuable to sustain engagement by moving in manageable steps, such as adding some reflective elements to early assessments and building to full wrappers. In this process, it is important to remember that not all changes are positively perceived by students, and that the skill of metacognition and SRL will require a shift in mindset toward learning and coaching on the quality of their reflections. Monitoring of students' perspectives on the perceived utility of exam wrappers and strategies they're utilizing can help inform decisions on adding additional time for SRL training and peer discussion. Minimally, it would be expected that metacognitive prompting would need to occur several times throughout a course to sustain desired behaviors.

Through ongoing refinement guided by context-specific needs and data, exam wrappers have the potential to powerfully enhance student success. It is important to note in this context that the strategies used to assess student learning in a course environment are related to the learning strategies that students are more likely to engage with, so it is important to consider courses holistically so SRL can be promoted along with highly functional assessment strategies (Struyven et al., 2005).

Implications for Assessment Scholarship

This manuscript demonstrates the relative dearth of high-quality studies on the impact of exam wrappers on student learning or metacognition. In studies of similar evidence-based instructional practices such as experiential learning (Schellehase, 2006), there has also been varied quality in study rigor. Ongoing research should continue refining wrapper design and identifying optimal supports to enhance metacognitive and motivational impacts.

- Rigorous design, field expansion: Explore effects in new disciplines using rigorous research designs. Mixed findings necessitate replication in diverse contexts with detailed documentation for future meta-analysis. Most significantly, very little exploration of exam wrappers has occurred in the humanities disciplines; this review demonstrated that most emerge from STEM.
- Implementation fidelity: A critical area for future research is the examination of implementation fidelity in exam wrapper usage. Implementation fidelity refers to the degree to which an intervention is delivered as intended (Carroll et al., 2007). In the context of exam wrappers, this involves assessing whether instructors are implementing the tool in a manner consistent with its theoretical foundations and best practices (O'Donnell, 2008). Future studies should develop clear guidelines for exam wrapper implementation, create validated measures

of fidelity, and investigate the relationship between implementation fidelity and student outcomes (Durlak & DuPre, 2008). This research should also explore how instructors adapt exam wrappers to their specific contexts and identify factors that support or hinder high-fidelity implementation (Century et al., 2010). To fully understand the impact of implementation fidelity, researchers should conduct longitudinal studies to examine how fidelity changes over time and its effect on sustained use of exam wrappers (Proctor et al., 2011). Additionally, comparing different models of implementation (e.g., instructor-led vs. peer-led, individual vs. group-based) could identify the most effective approaches (Dane & Schneider, 1998). By focusing on implementation fidelity, researchers can help bridge the gap between the theoretical promise of exam wrappers and their practical effectiveness in diverse educational settings (Lovett, 2013). This research can inform the development of more targeted training and support for instructors, ultimately leading to more consistent and impactful use of exam wrappers in higher education (Havis, 2019).

- Document nonadherence: Similar to the above-described issue, it's equally important to document unintentional and intentional nonadherence. Behavior change is a lofty goal, and it should be expected that some students may actively resist metacognitive reflection prompts or fail to continually implement new desired behaviors. In addition, they may struggle with developing metacognitive skills, and not achieve the quality of self-reflection needed to improve their SRL habits, and thus, performance. Therefore, monitoring and documenting nonadherence and metacognitive quality should provide greater insight into the reasons for the mixed results currently reported.
- Detail pedagogical context: Instructors make many pedagogical choices in their courses making context a key area to document for broader applicability implications. Quantity and length of wrappers is one element we synthesized (when possible), but the overall reflection time and grade relevance of wrappers to the totality of course experiences are challenging to discern. Some contexts may find the short wrappers as sufficient, while others may need to expand reflection prompting, add more training and increase incentives. Furthermore, most SRL research has focused on the face-to-face context, and evidence suggests that these impacts might be decreased online, but without more rigorous study and documentation of varied pedagogical contexts, like modality, it is impossible to make this distinction (Broadbent & Poon, 2015).
- Long-term behavior change investigation: We did not identify any studies investigating long-term metacognitive and SRL gains beyond the immediate course, and thus, impact to outcomes such as higher GPA, program retention rate and graduation rates. Minimal exploration has occurred (Hodges et al., 2020; Lovett, 2013; Swalve, 2021) on the benefit of students tasked with exam wrappers in multiple courses in the same term, but further examination and research on reinforcement over multiple terms is warranted. Furthermore, studies on transferability over time (to positive self-regulated behaviors in subsequent terms without exam wrapper prompts) and lasting learning benefits are still needed.

Conclusion

Self-efficacy profoundly shapes academic motivation, learning strategies, resilience, and achievement, and ultimately, academic success. Equipping students with robust efficacy requires evidence-based tools tailored to revealed needs. Exam wrappers allow responsive strengthening of self-efficacy and regulation by leveraging guided reflection. After identifying gaps through assessments, instructors can rapidly implement wrappers to address disengagement and prompt students to calibrate efficacy beliefs, analyze errors, adjust tactics, set implementation goals, and plan focused improvements. Exam wrappers model the virtuous cycles of metacognitive monitoring, tactical adaptation, and motivational bolstering that underlie meaningful learning. While demanding effort and promoting consistent refinement, wrappers reinforce students' capabilities to monitor and control their success. By extension, wrappers help students develop the self-efficacy and agency needed to tackle challenges in their educational journey. Exam wrappers serve as a means to address the Grand Challenges in Assessment in Higher Education, specifically, the improvement of pedagogy and student learning. Ongoing research should continue honing wrapper design and identifying optimal supports throughout the course design. However, thoughtfully applied exam wrappers provide an actionable technique for instructors to cultivate the metacognitive skills and self-beliefs that can shift student potential to immediate behavior and outcome change.

Acknowledgements

The authors extend their sincere gratitude to our Implementation Team chairs, Drs. Yao Hill and Jessica Taylor, and Grand Challenges in Assessment in Higher Education directors, Drs. Karen Singer-Freeman and Christine Robinson, for their leadership and support. We are especially appreciative of Dr. Taylor for her insightful feedback and contributions to this manuscript.

Appendix I: Exam Wrappers Question Examples

Compiled from LaCaille et al. (2019) and, with minor modifications, Mutambuki et al. (2020).

Preparation method

- Recall exactly how you prepared for the task. How much time was devoted? What study strategies did you use and how successful were they? Why were they helpful or not helpful?
- List the methods you used to study your notes and the readings before the quiz besides just reading them over.
- Based on the information from the videos, identify the ineffective or counterproductive beliefs and study strategies you used in preparing for the last exam.

Effectiveness and weaknesses

- Scrutinize the errors/mistakes you made on the task. Why do you think each occurred? (modified from Mutambuki et al., 2020)
- What questions do you have that you still need to seek answers for? What is your plan to address these? (modified from Mutambuki et al., 2020)
- What are the most important points you have learned thus far? (modified from Mutambuki et al., 2020)

Future planning and goal setting

- What, if anything, will you do differently to prepare for your next quiz?
- Describe your plan to prepare for the next exam: How many days before the exam will you complete reading and viewing all materials for the first time? How many days before the exam will you begin reviewing and studying the materials? About how many hours each day do you plan to study for the exam? How do you plan to minimize distractions while studying?
- What study strategies will you use to make sure that you are processing the information deeply? Clearly address how these strategies will involve the four components of deep processing: elaboration, distinctiveness, personalization, and appropriate retrieval and application.

Appendix II: Seed articles

Full citation only provided for those not cited in References.

1. Achacoso (2004)
2. Butzlaff et al. (2018)
3. Chen et al. (2017)
4. Chew et al. (2016)
5. Craig et al. (2016)
6. Edlund (2020)
7. El Bojairami & Driscoll (2019)
8. Fleming, V. M. (2002). Improving Students' Exam Performance by Introducing Study Strategies and Goal Setting. Teaching of Psychology, 29(2), 115–119. https://doi.org/10.1207/S15328023TOP2902_07
9. Gezer- Templeton et al. (2017)
10. Havis (2019)
11. Hodges et al. (2020)
12. Lovett (2013)
13. Olszewski, P. (2016). Teaching Millennials how to Study Under the 21st Century Sky. Pyrex Journal of Educational Research and Reviews, 2(1), 1–9. https://silo.tips/download/teaching-millennials-how-to-study-under-the-21-st-century-sky
14. Root Kustritz & Clarkson (2017)
15. Rosales et al. (2019)
16. Sethares and Asselin (2022)
17. Soicher & Gurung (2017)
18. Swalve et al. (2021)
19. Thompson (2012)
20. Williams (2021)

Appendix III: Reviewed But Excluded Literature

Full citation only provided for those not cited in References.

Exclusion rationale (primary reason listed but some may also have multiple attributes):

[a]Study design insufficient for exam wrapper impact measurement to test scores/grade performance or course-level learning outcomes achievement

[b]Doesn't explicitly meet exam wrapper intervention criteria

[c]Not in higher education

[d]Not empirical research

1. Achacoso (2004)[a]
2. Burgess, A., Bateman, K., & Schucker, J. (2020). Post examination review using a standardized examination review form. Teaching and Learning in Nursing, 15(1), 15–18. https://doi.org/10.1016/j.teln.2019.07.003[b]
3. Callender et al. (2016)[b]
4. Cleary & Zimmerman (2006)[c]
5. Colthorpe et al., 2019[a]
6. Cook et al. (2013)[a]
7. Fleming, V. M. (2002). Improving Students' Exam Performance by Introducing Study Strategies and Goal Setting. Teaching of Psychology, 29(2), 115–119. https://doi.org/10.1207/S15328023TOP2902_07[b]
8. Ibabe & Jauregizar (2010)[b]
9. Ierardi (2014)[b]
10. Lovett (2013)[a]
11. Masters (2007)[b]
12. Muñiz, M. N., Altinis-Kiraz, C., & Emenike, M. E. (2022). Extending Equity, Access, and Inclusion: An Evolving Multifaceted Approach to Transform a General Chemistry Course at a Large, Flagship, Research Institution. Journal of Chemical Education, 99(1), 227–238.[d]
13. Mynlieff et al. (2014)[b]
14. Nietfeld et al.'s (2005)[b]
15. Nietfeld et al.'s (2006)[b]
16. Olszewski, P. (2016). Teaching Millennials how to Study Under the 21st Century Sky. Pyrex Journal of Educational Research and Reviews, 2(1), 1–9. https://silo.tips/ download/teaching-millennials-how-to-study-under-the-21-st-century-sky[b]
17. Poorman & Mastorovich (2008)[d]
18. Poorman & Mastorovich (2016)[d]
19. Schuler & Chung (2019)[a]
20. Sethares and Asselin (2022)[d]
21. Sethares, K. A., Asselin, M. E., Mahoney, D., Nicotera, J., Chung, J., & Schuler, M. (2021). Description and comparison of exam wrapper learning strategy use in baccalaureate and associate degree nursing students: A descriptive study. *Nurse Education Today*, *103*, 104961. https://doi.org/10.1016/j.nedt.2021.104961[a]

22. Stanton, J. D., Neider, X. N., Gallegos, I. J., & Clark, N. C. (2015). Differences in metacognitive regulation in introductory biology students: when prompts are not enough. CBE—Life Sciences Education, 14(2), 1-12. https://doi.org/10.1187/cbe.14-08-0135[b]
23. Stephan, A. T., Stephan, E. A., & Miller, M. K. (2020, June 22). Extended Exam Wrappers: A Comparison of Approaches in a Learning Strategies Course. 2020 ASEE Virtual Annual Conference Content Access. https://peer.asee.org/extended-exam-wrappers-a-comparison-of-approaches-in-a-learning-strategies-course[a]
24. Tibbetts, T. 2011. Metacognition and examination wrappers in biology. Associated Colleges of the Midwest Project Summaries 2010. Retrieved from: https://serc.carleton.edu/acm_teagle/projects/tibbetts.html[b]
25. Tinnon (2018)[d]
26. Wang et al. (2018)[b]
27. Williams, A. E., Aguilar-Roca, N. M., Tsai, M., Wong, M., Beaupré, M. M., & O'Dowd, D. K. (2011). Assessment of Learning Gains Associated with Independent Exam Analysis in Introductory Biology. CBE—Life Sciences Education, 10(4), 346–356. https://doi.org/10.1187/cbe.11-03-0025[b]
28. Williams (2021)[d]

Appendix IV: Characteristics of Exam Wrapper Implementations in All 30 Qualifying Studies

Study	Lovett (2013) protocol similarity*	Implementation: # of wrappers; # of questions; Pre-/Post-Exam (or both) use	Peer sharing	SRL training inclusion	Relevancy to grade
Andaya et al. (2017)	1,2,3 (planning but not goal-setting)	3; 8-9; Post	None	None	Required, 10 points
Angell et al. (2024)	1,2,3	3; 10-15; Pre	None	None	Required, points (equivalent to 3-5% of each exam grade) were awarded "as long as they responded thoroughly"
Butzlaff et al. (2018)	1,2,3	4; varied, approximately 6; Post	Saw increased group studying	None	Required, 5/400 course points
Butzler (2016)	1,2,3,4	Not specified; Not specified; Post each unit exam	None	Note-taking instruction provided	Extra credit if completed "thoroughly and with reflection"
Chen et al. (2017)	1,2,3	4; varied, approximately 8; Both	None	Strategic Resource Use exercise	Voluntary, extra credit
Chew et al. (2016)	1,2,3,4 (once)	4; 3-30; Post 2 of 6 homework, Both for 1 of 3 exams	None	Unclear	Voluntary, in-class, credit unclear

Study	**Lovett (2013) protocol similarity***	**Implementation:** # of wrappers; # of questions; Pre-/Post-Exam (or both) use	**Peer sharing**	**SRL training inclusion**	**Relevancy to grade**
Colthorpe et al. (2017)	1,2,3	3-4; 6-8; Both	None	None	Non-optional, graded for completion
Craig et al. (2016)	1,2,3	2; 6; Post	None	None	Voluntary, 1% added to final grade for each
Dang et al. (2018)	1,2,3 (planning but not goal-setting)	3, 4-10, Both	Group quizzes and peer discussion occurred but it was unclear if wrapper prompts were covered	Optional videos on metacognition and deep vs surface learning	1-exam point (out of 100), plus extra credit for 3 assignments
Davis (2021)	1,2,3	3; 7-14; 1 Pre & 2 Post	None	None	Non-optional, 5-point assignment
Edlund (2020)	1,2,3,4	2; 5; Post	None	None	Voluntary, in-class, no grade
El Bojairami & Driscoll (2019)	2,3	2; 4; Post	None	Discussed introducing students to the concepts and value of metacognition and exam wrappers	Unclear, asynchronous electronic assignment

Study	**Lovett (2013) protocol similarity***	**Implementation:** # of wrappers; # of questions; Pre-/Post-Exam (or both) use	**Peer sharing**	**SRL training inclusion**	**Relevancy to grade**
Gezer-Templeton et al. (2017)	1,2,3	3; 3-5; Post	None	None	Extra credit
Grandoit et al. (2020)	1,2,3	4; 3; Both (Quizzes)	None	Extra handout with content on syllabus and reviewed in class	Non-optional, part of course grade
Hartling (2022)	1,2,3	1; 3; Post	None	None	Voluntary; no credit
Havis (2019)	1,2,3	2 studies: 2; 6; Post 1; 6; Post	2nd study only - "Autopsy model": 20-min peer conversation and worksheet (Instructor feedback on worksheet also received)	1st study: Study skills information provided and peer tutoring resource recommended (for note-taking and anxiety). 2nd study: Individual discussion sessions	Required; Credit for completion
Hodges et al. (2020)	1,2,3	1-4; 9; Not specified	Unclear	None	Varied, some required for points, some for extra credit

Study	**Lovett (2013) protocol similarity***	**Implementation:** # of wrappers; # of questions; Pre-/Post-Exam (or both) use	**Peer sharing**	**SRL training inclusion**	**Relevancy to grade**
Knight et al. (2022)	1,2 (but did not claim exam wrapper methodology)	6; 1-4; Both (quizzes)	None	None	Extra credit, 1 point per post-survey, 35% of students fully participated in post activities
LaCallie et al. (2019)	1,2,3	15-16; 13; Post (quizzes and 1st exam)	None	Video series and instruction on deep learning & study strategies	Extra credit awarded based on completeness
Mutambuki et al. (2020)	1,2,3	2; 2; Post	None	Training in metacognition provided at beginning of course and reminders throughout	Unclear
Pate et al. (2019)	1,2,3	4; 8; Post	Peer group conversations about missed items	30-min intro to metacognition presented to all students	Voluntary, structured, one-hour exam reviews outside of class, no credit
Root Kustritz & Clarkson (2017)	1,2,3,4	2; 9; Post 2 of 4 exams	None	None	Voluntary, in-class, credit unclear

Study	**Lovett (2013) protocol similarity***	**Implementation:** # of wrappers; # of questions; Pre-/Post-Exam (or both) use	**Peer sharing**	**SRL training inclusion**	**Relevancy to grade**
Rosales et al. (2019)	1,2,3	4; 3-5; Post (quizzes)	None	A syllabus handout on study tips was discussed in class	In-class, credit unclear
Sabel et al. (2017)	1,2,3	3, 8-12, Post	None	23% of students received 1:1 training	Voluntary, credit unclear
Soicher & Gurung (2017)	1,2,3	3; 3; Post	None	Unclear	Voluntary, in-class, credit unclear
Stephenson et al. (2017)	1,2,3,4	1; Post 1 of 2 exams	None	Brief discussion when wrappers were returned	Voluntary; extra credit awarded for low performers who completed the wrapper
Swalve et al. (2021)	1,2,3	3-7; 7; Both	None	Brief description of the goal at beginning of the course	Voluntary, credit unclear
Thompson (2012)	1,2,3,4 (but returned earlier)	2; 7; Post 2 of 7 exams	After the second wrapper, in-class discussion time	None	In-class, credit unclear

Study	**Lovett (2013) protocol similarity***	**Implementation:** # of wrappers; # of questions; Pre-/Post-Exam (or both) use	**Peer sharing**	**SRL training inclusion**	**Relevancy to grade**
Trogden & Royal (2019)	2,3,4	4; 2; Post	None	Not formal, faculty "used metacognitive language"	Each counted as a quiz score
Zimmerman et al. (2011)	1,2	4, 7-10, Post	Encouraged to discuss with peers	Extensive – teachers were pre-trained and modeled task, plus students engaged in SRL activities every 2 to 3 sessions	Voluntary, incentive points applied to quiz for corrected work

*Met the protocol criteria described in Lovett (2013) of prompting assessment performance self-reflection on 1) preparation methods, 2) error evaluation for effectiveness and weakness, 3) future planning and goal-setting, as well as 4) returning the exam wrapper to the student just prior to the next exam and encouraging review of their previous work and adherence to their preparation plan.

Appendix V: Summary of All 30 Qualifying Exam Wrapper Studies

To use this table effectively, consider the following steps:

- Familiarize yourself with the symbol key: The table uses symbols to indicate the direction and magnitude of the effects observed in each study. Improvement is denoted by ↑, improvement for only some groups by →, and no improvement by ⟳.
- Identify the context and method of each study: The table provides information about the subject area, sample size (when available), and research design employed in each study. This can help you assess the generalizability and reliability of the findings.
- Review the results summary: For each study, the table presents a brief overview of the main findings, focusing on the impact of exam wrappers on performance, metacognition, self-regulated learning, and other relevant outcomes.
- Compare findings across studies: Look for patterns and inconsistencies in the results of different studies. Consider factors such as the context, method, and implementation of exam wrappers that may contribute to the observed effects.
- Synthesize the information: Use the table to draw overall conclusions about the effectiveness of exam wrappers in higher education. Identify the strengths and limitations of the existing research and consider the implications for future practice and research.

Study	Context	Method	Results Summary
Andaya et al. (2017)	Biology (N = 44)	Mixed Methods – Parallel design with correlational quantitative analysis and descriptive qualitative elements	⟳ Those with high quality scores on the wrapper had higher grades on Exam 1 & 2, but 2 was not higher than 1 (so no evidence of the wrapper impacting performance)
Angell et al. (2024)	Biology (N = 233)	Quantitative - Quasi-experimental	⇀ Pre-exam assignment wrappers provided moderate exam score increases for students with lower incoming standardized test scores ⟳ Exam score prediction accuracy was associated with incoming standardized test scores, not wrappers
Butzlaff et al. (2018)	Nursing (Intervention = 120)	Qualitative – Descriptive	↑ Mean grade between the first and last exam (not statistically evaluated) ↑ Use of more effective study strategies
Butzler (2016)	Gen Chem (Semester 1: N = 45, Semester 2: N = 45, Semester 3: N = 32, Semester 4-5: 75)	Mixed Methods – Comparative, quantitative data included correlation analysis, while qualitative data was descriptive	⇀ Top/bottom HS performers benefited most ⟳ Overall achievement ↑ Perception of wrappers assisting was high
Chen et al. (2017)	Statistics (Study 1: N = 178, Study 2: N = 203)	Quantitative – Randomized control trials, separate and pooled analyses (No significant differences in groups on pre-intervention measures of motivation, importance, & confidence)	↑ Final grades (4% increase overall and 8-10% increase after 2 wrappers; no effect differences in sex, race, class standing, performance cohort, GPA[a]) ↑ Self-reflection on learning score increased (Study 2 only, adapted 8-item metacognitive self-regulation subscale of MSLQ[b]) ↑ Affect toward the exams[c] ↑ Perceived effectiveness of resource use[d]

Study	Context	Method	Results Summary
			⟳ Variety use of resources (Intervention group used fewer resources; Actual use of planned resources was needed for impact, but self-reports of plan quality and adherence was not correlated.)
Chew et al. (2016)	Engineering (N = 70)	Mixed Methods – Quantitative data was analyzed using descriptive statistics and correlation analysis, while qualitative data was analyzed thematically	⇀ Identification of a decrease in general mistakes on 1 assignment was weakly correlated to exam performance ↑ Planning, implementing strategies, confidence
Colthorpe et al. (2017)	Physiology	Mixed Methods – Convergent Parallel Design with inferential statistics for quantitative data as well as inductive and deductive thematic analyses	↑ Achievement ↑ Use of high-quality learning strategies
Craig et al. (2016)	Computer science (N = 259)	Quantitative – Randomized Control Trial	⟳ Exam performance when question corrections were combined with 3 wrapper types (placebo, metacognitive, conceptual)
Dang et al. (2018)	Biology (N = 171)	Mixed Methods - Parallel design with correlational quantitative analysis and descriptive qualitative elements	↑ Exam score prediction accuracy, with fewer over-predictors, especially for lower-performing students (overprediction correlated to lower exam scores) ⇀ Self-reported increased use in new study techniques by 85% of opt-in (n = 20) students and 45% reported grade improvement (others did not see change or did not mention it) ⟳ Metacognition (MAI^{e}) score did not correlate to exam performance)
Davis (2021)	Computer science and Software Engineering	Mixed Methods –	↑ Score on next exam, especially for low performing students

Study	Context	Method	Results Summary
	(N = 59, 49 completed intervention)	Descriptive	↑ Students self-reporting enacting their study enhancement plans
Edlund (2020)	Social psychology (Study 1: N = 84; Study 2: N = 63, 28 completed intervention)	Quantitative - Study 1: Quasi-experimental Study 2: Randomized Control Trial	↑ Score on next exam ↑ Students considered the wrappers useful ↑ Students who had high value scores on a critical thinking instrument benefited most from wrappers
El Bojairami & Driscoll (2019)	Engineering (N = 41)	Qualitative - Descriptive	↑ Quality score on their second reflection was associated with ↑ exam performance.
Gezer- Templeton et al. (2017)	Introductory Food Science (N = 100, 73 completed intervention)	Qualitative - Descriptive	⇀ Higher future exam scores for middling students.
Grandoit et al. (2020)	Organic chemistry (N = 176)	Mixed Methods - Descriptive quantitative and thematic analysis for qualitative	↻ Students' conceptual understandings and problem-solving ability ↻ By exam wrapper 4, 70% of students reported dissatisfaction with their course performance. Only 11% of students increased their study time, and only 25% increased their problem-solving time.
Hartling (2022)	Business (N = 67)	Qualitative - Descriptive	↑ Exam score improvement (modest) for all, D and F students gained the most
Havis (2019)	Criminology (N = 74, Post-Only Wrapper Intervention = 22, Exam Autopsy Intervention = 23)	Quantitative - Quasi-experimental, interrupted time series design	↑ Exam performance with exam autopsy model (which triangulates self, instructor, & peer feedback, not just post-exam self-reflection wrappers)

Study	Context	Method	Results Summary
Hodges et al. (2020)	STEM: Biology, Chemistry, Math [Physics for control] (N = 2,393, intervention = 1,613)	Quantitative - Quasi-experimental	↑ Modest achievement gains ⇀ Gains in overall GPA for men only ⇀ Metacognition (via MAI[e]) increase only seen for non-first-year students in the wrapper courses
Knight et al. (2022)	Genetics (N = 496)	Mixed Methods - Descriptive Convergent Parallel Design	⟳ Frequency of metacognitive responses over time was not predictive of grade increase ⇀ Those who shift from overpredicting to matching or underpredicting performance had improved performance.
LaCaille et al. (2019)	Psychology (N = 244, intervention = 123)	Quantitative - Quasi-experimental	↑ Moderately higher quiz and exam scores for students reporting higher use of wrappers ⇀ Metacognition knowledge (MAI[e]) was higher than control, but not regulation ↑ Competence (Perceived competence scale for learning[f]) ↑ Students reported enjoying learning the material more strongly when using wrappers.
Mutambuki et al. (2020)	General Chemistry (N = 427, intervention = 239)	Quantitative - Quasi-experimental	↑ Exam 3 and final scores, and lower withdrawal rates
Pate et al. (2019)	Pathophysiology for Professional level Pharmacy students (N = 88)	Quantitative - Quasi-experimental	⟳ No exam score gains from ~1 wrapper; More of the lower-performing students on Exam 1 participated
Root Kustritz & Clarkson (2017)	Professional level Veterinary students (N = 59)	Quantitative - Quasi-experimental	⟳ Achievement differences with minimal wrapper participation

Study	Context	Method	Results Summary
Rosales et al. (2019)	Organic Chemistry (N = 154, Intervention = 71)	Quantitative - Quasi-experimental, interrupted time series design	↑ Final exam score ↑ 1 specific item % correct on final exam ↑ Course grades ↑ Withdrawal (vs failure) rates ↻ Study strategies (active vs passive)
Sabel et al. (2017)	Biology (Quantitative: N = 88, Qualitative: N = 20)	Mixed Methods – Convergent parallel design, quantitative data was analyzed using descriptive statistics and correlation analysis, while qualitative data was analyzed thematically	↑ Use of previous enhanced exam answer keys as a study aid for future exams ↑ Assignment scores with some use of keys → Interviewed students, who received guidance on best use of keys, performed significantly better in the course and reported more metacognitive engagement
Soicher & Gurung (2017)	Psychology (N = 25-86)	Quantitative - Quasi-experimental	↻ Exam scores ↻ Metacognition (via MAI^{e})
Stephenson et al. (2017)	Computer Science (Study 1 N = 289, Study 2 N = 752)	Quantitative - Study 1 = Quasi-experimental, interrupted time series design Study 2 = Randomized control trial, plus open-ended survey items	↻ Final exam score ↑ Students appreciated the wrappers and they reported increased use of active study strategies.
Swalve et al. (2021)	Biology (Intervention = 81) & Chemistry	Quantitative - Quasi-experimental, interrupted time series design	↻ No improvement in course grades in chemistry (3 wrappers) ↓ Course grades in biology course (7 wrappers) → Self-efficacy increased in chemistry, decreased in biology

Study	Context	Method	Results Summary
	(Intervention = 64); Control N = unknown		↓ Anxiety decreased in chemistry No information is provided on wrapper completion rates.
Thompson (2012)	Spanish (Study 1: N = 78, Study 2: N = 35)	Quantitative - Quasi-experimental	⟳ in achievement ⇀ Self-Regulated Learning, only first-year students gained (via Metacognitive self-regulation subscale of MSLQ[b])
Trogden & Royal (2019)	STEM General Chemistry (N = 27), Organic Chemistry (Intervention = 26) Calculus (N = 83), Statistics (N = 32)	Mixed Methods - Descriptive quantitative and thematic analysis for qualitative	↑ Organic Chemistry II exam scores over the semester ↑ SRL use amongst the study strategies listed ↑ Metacognition (via MAI[e])
Zimmerman et al. (2011)	Math (N = 496)	Quantitative - Randomized Control Trial	↑ Exam performance ⇀ Exam performance was even greater for more frequent self-reflectors ↑ Accuracy in item confidence ratings ↑ Pass rate on a standardized test

Note:

[a]*Grade Point Average*

[b]*Motivated Strategies for Learning Questionnaire (Pintrich et al., 1991)*

[c]*Adapted from Smith & Ellsworth (1987)*

[d]*Adapted from the Resource Questionnaire (Brown et al., 1996)*

[e]*Metacognitive Awareness Inventory (Schraw & Dennison, 1994)*

[f]*Perceived competence scale for learning (Williams & Deci, 1996)*

References

Achacoso, M. V. (2004). Post-test analysis: A tool for developing students' metacognitive awareness and self-regulation. *New Directions for Teaching and Learning*, *2004*(100), 115–119. https://doi.org/10.1002/tl.179

Adams, K. (2004). Modelling success: Enhancing international postgraduate research students' self-efficacy for research seminar presentations. *Higher Education Research & Development*, *23*(2), 115–130. https://doi.org/10.1080/0729436042000206618

Andaya, G., Hrabak, V. D., Reyes, S. T., Diaz, R. E., & McDonald, K. K. (2017). Examining the Effectiveness of a Postexam Review Activity to Promote Self-Regulation in Introductory Biology Students. *Journal of College Science Teaching*, *46*(4), 84–92. https://doi.org/10.2505/4/jcst17_046_04_84

Angell, D. K., Lane-Getaz, S., Okonek, T., & Smith, S. (2024). Metacognitive Exam Preparation Assignments in an Introductory Biology Course Improve Exam Scores for Lower ACT Students Compared with Assignments that Focus on Terms. *CBE—Life Sciences Education*, *23*(1), ar6. https://doi.org/10.1187/cbe.22-10-0212

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, *28*(2), 117–148.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman & Co.

Biernacki, P., & Waldorf, D. (1981). Snowball Sampling: Problems and Techniques of Chain Referral Sampling. *Sociological Methods & Research*, *10*(2), 123–240.

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, *27*, 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007

Brown, M. I., Doughty, G. F., Draper, S. W., Henderson, F. P., & McAteer, E. (1996). Measuring learning resource use. *Computers & Education*, *27*(2), 103–113. https://doi.org/10.1016/0360-1315(96)00017-6

Butzlaff, A., Gaylle, D., & O'Leary Kelley, C. (2018). Student Self-evaluation After Nursing Examinations: That's a Wrap. *Nurse Educator*, *43*(4), 187. https://doi.org/10.1097/NNE.0000000000000534

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

Callender, A. A., Franco-Watkins, A. M., & Roberts, A. S. (2016). Improving metacognition in the classroom through instruction, training, and feedback. *Metacognition and Learning*, *11*(2), 215–235. https://doi.org/10.1007/s11409-015-9142-6

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, *2*(1), 40. https://doi.org/10.1186/1748-5908-2-40

Century, J., Rudnick, M., & Freeman, C. (2010). A Framework for Measuring Fidelity of Implementation: A Foundation for Shared Language and Accumulation of Knowledge. *American Journal of Evaluation*, *31*(2), 199–218. https://doi.org/10.1177/1098214010366173

Chen, P., Chavez, O., Ong, D. C., & Gunderson, B. (2017). Strategic Resource Use for Learning: A Self-Administered Intervention That Guides Self-Reflection on Effective Resource Use Enhances Academic Performance. *Psychological Science*, *28*(6), 774–785. https://doi.org/10.1177/0956797617696456

Chew, K. J., Sheppard, S., Chen, H. L., Rieken, B., & Turpin, A. (2016). *Improving Students' Learning in Statics Skills: Using Homework and Exam Wrappers to Strengthen Self-regulated Learning*. 2016 ASEE Annual Conference & Exposition. https://doi.org/10.18260/p.25633

Cioffi, D. (1991). Sensory awareness versus sensory impression: Affect and attention interact to produce somatic meaning. *Cognition & Emotion*, *5*(4), 275–294. https://doi.org/10.1080/02699939108411041

Cleary, T. J., & Zimmerman, B. J. (2006). Teachers' perceived usefulness of strategy microanalytic assessment information. *Psychology in the Schools*, *43*(2), 149–155. https://doi.org/10.1002/pits.20141

Colthorpe, K., Ogiji, J., Ainscough, L., Zimbardi, K., & Anderson, S. (2019). Effect of Metacognitive Prompts on Undergraduate Pharmacy Students' Self-regulated Learning Behavior. *American Journal of Pharmaceutical Education*, *83*(4), Article 6646. https://doi.org/10.5688/ajpe6646

Colthorpe, K., Sharifirad, T., Ainscough, L., Anderson, S., & Zimbardi, K. (2017). Prompting undergraduate students' metacognition of learning: Implementing 'meta-learning' assessment tasks in the biomedical sciences. *Assessment & Evaluation in Higher Education*, *43*(2), 272–285. https://doi.org/10.1080/02602938.2017.1334872

Cook, E., Kennedy, E., & McGuire, S. Y. (2013). Effect of Teaching Metacognitive Learning Strategies on Performance in General Chemistry Courses. *Journal of Chemical Education*, *90*(8), 961–967. https://doi.org/10.1021/ed300686h

Coutinho, S. A. (2007). The relationship between goals, metacognition, and academic success. *Educate~*, *7*(1), 39–47. https://www.educatejournal.org/index.php/educate/article/view/116/134

Craig, M., Horton, D., Zingaro, D., & Heap, D. (2016). Introducing and Evaluating Exam Wrappers in CS2. *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, 285–290. https://doi.org/10.1145/2839509.2844561

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control. *Clinical Psychology Review*, *18*(1), 23–45. https://doi.org/10.1016/s0272-7358(97)00043-3

Dang, N. V., Chiang, J. C., Brown, H. M., & McDonald, K. K. (2018). Curricular Activities that Promote Metacognitive Skills Impact Lower-Performing Students in an Introductory Biology Course. *Journal of Microbiology & Biology Education*, *19*(1), 1-9. https://doi.org/10.1128/jmbe.v19i1.1324

Davis, K. C. (2021, July 26). *Examining the Efficacy of Exam Wrappers in a Computer Science Course*. 2021 ASEE Virtual Annual Conference Content Access. https://doi.org/10.18260/1-2--37120

Durlak, J. A., & DuPre, E. P. (2008). Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. *American Journal of Community Psychology*, *41*(3), 327–350. https://doi.org/10.1007/s10464-008-9165-0

Dusek, G., Yurova, Y., & P. Ruppel, C. (2015). Using Social Media and Targeted Snowball Sampling to Survey a Hard-to-reach Population: A Case Study. *International Journal of Doctoral Studies*, *10*, 279–299. https://doi.org/10.28945/2296

Edlund, J. E. (2020). Exam Wrappers in Psychology. *Teaching of Psychology*, *47*(2), 156–161. https://doi.org/10.1177/0098628320901385

El Bojairami, I., & Driscoll, M. (2019). Exam-Wrappers as a Tool to Enhance Students' Metacognitive Skills in Machine Element Design Class. *2019: Proceedings of the Canadian Engineering Education Association (CEEA-ACEG) Conference*, *138*, 1–8. https://doi.org/10.24908/pceea.vi0.13815

Ergen, B., & Kanadli, S. (2017). The Effect of Self-Regulated Learning Strategies on Academic Achievement: A Meta-Analysis Study. *Eurasian Journal of Educational Research*, *69*, 55–74.

Etikan, I. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1. https://doi.org/10.11648/j.ajtas.20160501.11

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

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

Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition*, *10*(6), 597–602. https://doi.org/10.3758/BF03202442

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation Intentions and Goal Achievement: A Meta-analysis of Effects and Processes. In *Advances in Experimental Social Psychology* (Vol. 38, pp. 69–119). Academic Press. https://doi.org/10.1016/S0065-2601(06)38002-1

Goodman, L. A. (1961). Snowball Sampling. *The Annals of Mathematical Statistics*, *32*(1), 148–170.

Grandoit, E., Bergdoll, R., Rosales, E., Turbeville, D., Mayer, S., & Horowitz, G. (2020). Exploring Organic Chemistry I Students' Responses to an Exam Wrappers Intervention. *Journal of the Scholarship of Teaching and Learning*, *20*(1), 105–129. https://doi.org/10.14434/josotl.v20i1.24825

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, *92*(1), 160–170. https://doi.org/10.1037/0022-0663.92.1.160

Hartling, X. (2022). Assessing the effectiveness of exam wrappers in a quantitative business course. *Academy of Educational Leadership Journal*, *26*(5), 1–8.

Havis, L. R. (2019). The Exam Autopsy: An Integrated Post-Exam Assessment Model. *International Journal for the Scholarship of Teaching and Learning*, *13*(1). https://doi.org/10.20429/ijsotl.2019.130104

Heckathorn, D. D., & Cameron, C. J. (2017). Network Sampling: From Snowball and Multiplicity to Respondent-Driven Sampling. *Annual Review of Sociology*, *43*(Volume 43, 2017), 101–119. https://doi.org/10.1146/annurev-soc-060116-053556

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: Journal of College Science Teaching. *Journal of College Science Teaching*, *50*(1), 69–79. https://doi.org/10.1080/0047231x.2020.12290677

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review*, *17*, 63–84. https://doi.org/10.1016/j.edurev.2015.11.002

Ibabe, I., & Jauregizar, J. (2010). Online self-assessment with feedback and metacognitive knowledge. *Higher Education*, *59*(2), 243–258. https://doi.org/10.1007/s10734-009-9245-6

Ierardi, J. A. (2014). Taking the 'Sting' Out of Examination Reviews: A Student-Centered Approach. *Journal of Nursing Education*, *53*(7), 428–428. https://doi.org/10.3928/01484834-20140619-13

Knight, J. K., Weaver, D. C., Peffer, M. E., & Hazlett, Z. S. (2022). Relationships between Prediction Accuracy, Metacognitive Reflection, and Performance in Introductory Genetics Students. *CBE—Life Sciences Education*, *21*(3), Article ar45. https://doi.org/10.1187/cbe.21-12-0341

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

LaCaille, R. A., LaCaille, L. J., & Maslowski, A. K. (2019). Metacognition, course performance, and perceived competence for learning: An examination of quiz and exam wrappers. *Scholarship of Teaching and Learning in Psychology*, *5*(3), 209–222. https://doi.org/10.1037/stl0000114

Lovett, M. C. (2013). Make exams worth more than the grade: Using exam wrappers to promote metacognition. In *Using reflection and metacognition to improve student learning* (pp. 18–52). Stylus Publishing.

Masters, J. C. (2007). Taking the Sting out of Examination Review. *Journal of Nursing Education*, *46*(8), 384. https://doi.org/10.3928/01484834-20070801-09

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, *75*(1), 33–52. https://doi.org/10.1037/0022-3514.75.1.33

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of Self-Efficacy Beliefs to Academic Outcomes: A meta-analytic Investigation. *Journal of Counseling Psychology*, *38*(1), 30–28.

Mutambuki, J. M., Mwavita, M., Muteti, C. Z., Jacob, B. I., & Mohanty, S. (2020). Metacognition and Active Learning Combination Reveals Better Performance on Cognitively Demanding General Chemistry Concepts than Active Learning Alone. *Journal of Chemical Education*, *97*(7), 1832–1840. https://doi.org/10.1021/acs.jchemed.0c00254

Mynlieff, M., Manogaran, A. L., St. Maurice, M., & Eddinger, T. J. (2014). Writing Assignments with a Metacognitive Component Enhance Learning in a Large Introductory Biology Course. *CBE—Life Sciences Education*, *13*(2), 311–321. https://doi.org/10.1187/cbe.13-05-0097

Niedwiecki, A. (2012). Teaching for lifelong learning: Improving the metacognitive skills of law students through more effective formative assessment techniques. *Capital University Law Review*, *40*(149), 149-193. https://ssrn.com/abstract=2302717

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2005). Metacognitive Monitoring Accuracy and Student Performance in the Postsecondary Classroom. *The Journal of Experimental Education*, *74*(1), 7–28.

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance, monitoring accuracy, and self-efficacy. *Metacognition and Learning*, *1*(2), 159–179. https://doi.org/10.1007/s10409-006-9595-6

O'Donnell, C. L. (2008). Defining, Conceptualizing, and Measuring Fidelity of Implementation and Its Relationship to Outcomes in K–12 Curriculum Intervention Research. *Review of Educational Research*, *78*(1), 33–84. https://doi.org/10.3102/0034654307313793

Pate, A., 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

Pintrich, P. R. (2000). Multiple Goals, Multiple Pathways: The Role of Goal Orientation in Learning and Achievement. *Journal of Educational Psychology*, *92*(3), 544–555.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)* (Report No. NCRIPTAL-91-B-004). National Center for Research to Improve Postsecondary Teaching and Learning. https://eric.ed.gov/?id=ED338122

Poorman, S. G., & Mastorovich, M. L. (2008). Using Metacognitive Strategies to Help Students Learn in Pretest and Posttest Review. *Nurse Educator*, *33*(4), 176–180. https://doi.org/10.1097/01.NNE.0000312199.81815.f4

Poorman, S. G., & Mastorovich, M. L. (2016). Using Metacognitive Wrappers to Help Students Enhance Their Prioritization and Test-Taking Skills. *Nurse Educator*, *41*(6), 282–285. https://doi.org/10.1097/NNE.0000000000000257

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for Implementation Research: Conceptual Distinctions, Measurement Challenges, and Research Agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*(2), 65–76. https://doi.org/10.1007/s10488-010-0319-7

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353–387. https://doi.org/10.1037/a0026838

Root Kustritz, M. V., & Clarkson, C. E. (2017). Use of Examination Wrappers to Direct Student Self-Assessment of Examination Preparation: A Pilot Study. *Journal of Veterinary Medical Education*, *44*(2), 338–342. https://doi.org/10.3138/jvme.0216-042R

Rosales, E., Chavarga, A., Grandoit, E., Mayer, S., Hackman, N., Elbulok-Charcape, M., Domzalski, A. C., & Horowitz, G. (2019). An Exam Wrapper Intervention in Organic Chemistry I: Impact on Course Performance and Study Behavior. *Journal of College Science Teaching*, *49*(2), 53–62. https://www.jstor.org/stable/26901368

Sabel, J. L., Dauer, J. T., & Forbes, C. T. (2017). Introductory Biology Students' Use of Enhanced Answer Keys and Reflection Questions to Engage in Metacognition and Enhance Understanding. *CBE—Life Sciences Education*, *16*(3), ar40. https://doi.org/10.1187/cbe.16-10-0298

Schellehase, K. C. (2006). Kolb's Experiential Learning Theory in Athletic Training Education: A Literature Review. *Athletic Training Education Journal*, *1*(2), 18–27. https://doi.org/10.4085/1947-380X-1.2.18

Schneider, E. F., Castleberry, A. N., Vuk, J., & Stowe, C. D. (2014). Pharmacy Students' Ability to Think About Thinking. *American Journal of Pharmaceutical Education*, *78*(8), 148. https://doi.org/10.5688/ajpe788148

Schraw, G., & Dennison, R. S. (1994). Assessing Metacognitive Awareness. *Contemporary Educational Psychology*, *19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Schuler, M. S., & Chung, J. (2019). Exam Wrapper Use and Metacognition in a Fundamentals Course: Perceptions and Reality. *Journal of Nursing Education*, *58*(7), 417–421. https://doi.org/10.3928/01484834-20190614-06

Schunk, D. H. (1995). Self-Efficacy and Education and Instruction. In J. E. Maddux (Ed.), *Self-Efficacy, Adaptation, and Adjustment: Theory, Research, and Application* (pp. 281–303). Springer US. https://doi.org/10.1007/978-1-4419-6868-5_10

Schunk, D. H., & Ertmer, P. A. (1999). Self-regulatory processes during computer skill acquisition: Goal and self-evaluative influences. *Journal of Educational Psychology*, *91*(2), 251–260. https://doi.org/10.1037/0022-0663.91.2.251

Schwinger, M., & Otterpohl, N. (2017). Which one works best? Considering the relative importance of motivational regulation strategies. *Learning and Individual Differences*, *53*, 122–132. https://doi.org/10.1016/j.lindif.2016.12.003

Sebesta, A. J., & Bray Speth, E. (2017). How Should I Study for the Exam? Self-Regulated Learning Strategies and Achievement in Introductory Biology. *CBE—Life Sciences Education*, *16*(2), ar30. https://doi.org/10.1187/cbe.16-09-0269

Sethares, K. A., & Asselin, M. E. (2022). Use of Exam Wrapper Metacognitive Strategy to Promote Student Self-assessment of Learning: An Integrative Review. *Nurse Educator*, *47*(1), 37. https://doi.org/10.1097/NNE.0000000000001026

Singer-Freeman, K., & Robinson, C. (2020a). Grand Challenges for Assessment in Higher Education. *Research & Practice in Assessment*, *15*(2), 1–20. https://www.rpajournal.com/grand-challenges-for-assessment-in-higher-education/

Singer-Freeman, K., & Robinson, C. (2020b). *Grand Challenges in Assessment: Collective Issues in Need of Solutions* (Occasional Paper No. 47; pp. 1–18). National Institute for Learning Outcomes Assessment. https://www.learningoutcomesassessment.org/wp-content/uploads/2020/11/GrandChallenges.pdf

Smith, C. A., & Ellsworth, P. C. (1987). Patterns of appraisal and emotion related to taking an exam. *Journal of Personality and Social Psychology*, *52*(3), 475–488. https://doi.org/10.1037//0022-3514.52.3.475

Soicher, R. N., & Gurung, R. A. R. (2017). Do Exam Wrappers Increase Metacognition and Performance? A Single Course Intervention. *Psychology Learning & Teaching*, *16*(1), 64–73. https://doi.org/10.1177/1475725716661872

Stephenson, B., Craig, M., Zingaro, D., Horton, D., Heap, D., & Huynh, E. (2017). Exam Wrappers: Not a Silver Bullet. *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, 573–578. https://doi.org/10.1145/3017680.3017701

Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment & Evaluation in Higher Education*, *30*(4), 325–341. https://doi.org/10.1080/02602930500099102

Swalve, N., Harwood, A., & Calhoun, E. S. (2021). 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

Tarnavsky Eitan, A., Smolyansky, E., Knaan Harpaz, I., & Perets, S. (n.d.). *Connected Papers* [Computer software]. https://www.connectedpapers.com/

Thompson, D. R. (2012). Promoting Metacognitive Skills in Intermediate Spanish: Report of a Classroom Research Project. *Foreign Language Annals*, *45*(3), 447–462. https://doi.org/10.1111/j.1944-9720.2012.01199.x

Tinnon, E. A. (2018). Reflective Test Review: The First Step in Student Retention. *Teaching and Learning in Nursing*, *13*(1), 31–34. https://doi.org/10.1016/j.teln.2017.09.001

Trogden, B. G., & Royal, J. E. (2019). Using Exam Wrappers in Chemistry, Mathematics, and Statistics Courses to Encourage Student Metacognition. *Journal on Excellence in College Teaching*, *30*(3), 71–96. https://celt.miamioh.edu/ojs/index.php/JECT/article/view/279

van Dinther, M., Dochy, F., & Segers, M. (2011). Factors affecting students' self-efficacy in higher education. *Educational Research Review*, *6*(2), 95–108. https://doi.org/10.1016/j.edurev.2010.10.003

Wagener, B. (2016). Metacognitive Monitoring and Academic Performance in College. *College Teaching*, *64*(2), 47–54. https://doi.org/10.1080/87567555.2015.1116056

Wang, X. R., Cruthirds, D. L., & Kendrach, M. G. (2018). Effect of an Individualized Post-Examination Instructor Remediation on Pharmacy Student Performance in a Biochemistry Course. *American Journal of Pharmaceutical Education*, *82*(6), 6297. https://doi.org/10.5688/ajpe6297

Williams, C. A. (2021). Exam Wrappers: It Is Time to Adopt a Nursing Student Metacognitive Tool for Exam Review. *Nursing Education Perspectives*, *42*(1), 51–52. https://doi.org/10.1097/01.NEP.0000000000000551

Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of self-determination theory. *Journal of Personality and Social Psychology*, *70*(4), 767–779. https://doi.org/10.1037/0022-3514.70.4.767

Young, A., & Fry, J. D. (2008). Metacognitive Awareness and Academic Achievement in College Students. *Journal of the Scholarship of Teaching and Learning*, *8*(2), 1–10. https://scholarworks.iu.edu/journals/index.php/josotl/article/view/1696/1694

Zimmerman, B. J. (1990). Self-Regulated Learning and Academic Achievement: An Overview. *Educational Psychologist*. https://doi.org/10.1207/s15326985ep2501_2

About the Authors

- Rebecca E. Gibbons, Assistant Director of Assessment Sarasota-Manatee campus, Office of Decision Support, University of South Florida, revaclav@usf.edu
- Deborah Hokien, Associate Dean, College of Liberal Arts and Sciences, Kutztown University, hokien@kutztown.edu
- Bryant L. Hutson, Director of Assessment, Institutional Research & Assessment, University of North Carolina at Chapel Hill, bhutson@email.unc.edu
- Heather T.D. Maness, Assistant Director of Learning Analytics and Assessment, Information Technology, University of Florida, htdaniel@ufl.edu