

# EXAM-WRAPPERS AS A TOOL TO ENHANCE STUDENTS' METACOGNITIVE SKILLS IN MACHINE ELEMENT DESIGN CLASS

Ibrahim El Bojairami & Mark Driscoll

Department of Mechanical Engineering, McGill University, Montréal, Québec, Canada

[ibrahim.elbojairami@mail.mcgill.ca](mailto:ibrahim.elbojairami@mail.mcgill.ca) & [mark.driscoll@mcgill.ca](mailto:mark.driscoll@mcgill.ca)

**Abstract** -Acquiring metacognitive skills, i.e. self-assessment of study behaviors through exam-wrappers, would potentially have concrete positive effects on students learning. The purpose of this study was to assess the link between students' performance and study behaviors, to analyze their metacognitive skills.

Exam-wrappers consisted of 4 fundamental questions and were utilized in an undergraduate engineering class, Machine Element Design, at McGill University. Data were later collected and analyzed using statistical analyses.

Results showed an overall increasing trend between the quality of reflection and students grades' improvement, with the highest being a 47.5 increase in grade for a student who scored a 10/10 reflection.

Analyzed cases led to the conclusion that students were able to use their metacognitive skills to self-assess, and accordingly adjust their study behaviors and goals, to thrive at a better performance.

**Keywords:** Metacognitive skills, Self-assessment, Self-learning, Exam-wrappers, Students Performance.

## 1. INTRODUCTION

### 1.1. Motivation for the paper

From the "reflection in action" and "reflection on action" being a critical feature of professional practice by Donald Schon's [1] to the "reflective learning" phase in the experiential learning cycle of Kolb [2], the importance of reflection in education has been extensively documented, recognized, and valued. The process of reflection has been defined as "intentionally making meaning of experiences in service of future action" [3]. In the context of education, this reflects on the idea of enhancing students' metacognitive skills, stimulating them towards the process of self-assessment. Upon doing so, students would identify what learning methodology should they keep on following, and what others they should change, in order to accomplish the best outcomes.

In this context, exam-wrappers present themselves as a plausible strategy in response to students' poor studying strategies and inability to use feedback to learn about their learning difficulties. Exam-wrappers can be defined as structured reflection activities in which students engage in after their instructor marks and returns an exam [3]. Previous studies showed that using exam-wrappers, in particular, is an excellent method of assessing student performance, metacognitive skills, and study behaviors [4], [5]. Such studies have further shown that wrappers have a huge impact on students' learning outcomes. Chew et al. found that the quantitative and qualitative results collected from wrappers done in an engineering statics class made a strong argument for the mutual benefits to both students and the teaching team to improve and enhance the teaching and learning experience [4].

Based on the premise that exam-wrappers help students practice metacognitive skills, it was decided to implement them in Machine Element Design, a Mechanical Engineering class (MECH393), at McGill university in the Fall-2019 semester. Further, reading the literature suggested that wrappers are highly likely to have a good impact on the students, and definitely no negative ones. One last factor that encouraged carrying out this exam-wrappers study is borne out of previous observations where students do not carefully read and reflect on the feedback given on their graded tests, and in some cases would not even collect the tests from the professor's office.

### 1.2. Literature review

In response to the increased failure rates [6] and the disappointing students' performance [7], Computer Science (CS) researchers and educators started to implement a wide range of strategies to reverse such effects, ergo, promote students' success. The creativity and efficacy of these strategies is summarized by creating new pedagogical approaches [8], early-warning feedback mechanisms [9], and connecting to students' contextual interests [10]. It is critical to emphasize the importance of frequently giving feedbacks to students, which in turn urges them to closely assessing their

performance, and changing their study habits accordingly, if necessary. For instance, peer instruction, a research-based pedagogical approach, includes giving feedback to students, multiple times per class [8], where some teaching tools are actually based primarily on this, and on small exercises complemented with on-the-spot feedback [11].

The idea of exam-wrappers was first introduced by Lovett as being a series of short activities that urge students to review their performance, based on the instructor's feedback, and properly altering their future learning [12]. According to Lovett, exam-wrappers should prompt students to reflect on three main components: study skills used to prepare for the exam, errors types made on the exam, and, perhaps the most critical component, modifications of study habits for better preparation for further exams [12]. Being the first to introduce this novel idea, Lovett identified three main questions for exam-wrappers to be structured upon: how did the students prepare for the exam? What kind of mistakes did they do? What can they do differently to improve their grade next time? She implemented exam-wrappers in multiple introductory level classes: biology, physics, chemistry, and calculus. She observed metacognition improvements over the academic term, especially for students enrolled more than one course that used exam-wrappers.

An earlier study that used a post-exam questionnaire, similar to what is known today as exam-wrappers, also found improvements in students' metacognitive skills [13]. This improvement was translated by an overall increase in exam results and the ability of students to adjust their study habits accordingly. One last study that used exam-wrappers was conducted by Stephenson et al. to examine the effectiveness of this tool, by implementing it in two classes, CS1 and CS2 [14]. Both studies concluded that midterm exam-wrappers are not effective for increasing mean final exam performance in first year computer science courses. They further admitted that their studies had limitations; however, they claimed that those limitations were different between both studies, and further complementary, with the strengths of one study being the weaknesses of the other. One last note they made was that even though exam-wrappers did not improve final exam performance, they still have potential benefits. Course evaluations showed that students viewed exam-wrappers as being beneficial and let them feel valued by the instructor.

In addition to having positive impact, or at least no negative one, exam-wrappers are easy to implement. According to Gezer-Templeton et al., wrappers require short time for both, students to fill and instructors to prepare and assess, mostly because it is two pages at most, of short-answer questions [15]. Further, they are

repeatable and can be adapted from course to course easily.

### 1.3. Problem definition

The shift from high-school to college/university imposes complications of developing higher-order, critical, thinking and powerful studying strategies for students. As suggested by Krathwohl et al., thinking skills progress from factual knowledge, to conceptual knowledge, to procedural knowledge, and to metacognitive knowledge [16]. As such, acquiring metacognitive skills, i.e. self-assessment and self-regulation of study behaviors, would potentially have long-lasting and concrete positive effects on students learning, especially for those facing difficulties adjusting to college level.

In the context of the MECH393 class that has been taught for the past three years by professor Mark Driscoll at McGill University, generalist and informal observations revealed that some students struggled with the discipline and motivation to change their study habits based on their performance earlier in the class. In the past, most cases revealed that students who score high grades on any midterm, assignment, or project keep the same performance throughout the semester. However, those who either fail or score low grades, get disappointed, and thus, either surrender or work even harder by following the same study behavior, ending up with a similar unsatisfactory performance.

In an attempt to understand this behavior, the main problem hypothesis was defined as: Students could improve their metacognitive skills, i.e. self-assessing their own performance, and accordingly, self-regulating their study behavior to thrive at better results.

### 1.4. Solutions considered

In an effort to overcome the explained problem, there exists many suggested methods to enhance students' metacognition. By nature, the semester is short, and students have a lot of other courses to focus on, thus, solutions to consider should be limited by time and applicability.

Student buy-in is a good approach to help overcoming student resistance by being explicit with them. Another considered method is reflective journaling, i.e. providing students with regular activities that focus on the process of learning and linking behaviors to outcomes. One last strategy considered was exam-wrappers, which is also a type of reflection. Wrappers are short questionnaires given after an exam or an assignment that reflect mainly on students' performance, identifying mistakes, and changes in study habits.

## 1.5. Selected methodology

After closely assessing the considered solutions, the first two were eliminated. This is because such studies require allocating time during each lecture, being open with students, through engaging into specific conversations, or by assigning post-classes reflective assignments, which is perhaps less applicable in an engineering class. On the other hand, exam-wrappers are easy to implement, especially because they are short, take almost no time, and are only done at some points during the semester.

The importance of exam-wrappers comes from the fact that students, after an exam, usually focus on their scores instead of focusing on how their strategies affected that score. In efforts to change this, wrappers are designed in a way to relate to exam performance by focusing on three main parts: exam preparation, mistakes made on the exam, and more essentially the changes that should be made prior to the next exam.

### Implementation

Exam-wrappers were utilized in a junior-level engineering class, Machine Element Design (MECH 393), in the Fall semester of 2018, at McGill University. This is a mandatory class for all mechanical engineering students, and during this term, 41 students were enrolled in it. The course held 2 exams: a midterm covering static and fatigue failure, shafts, and keys design; and a final covering gears, springs, bearings, screws, and fasteners design. Exam-wrappers were utilized after the midterm electronically via *Socrative*, an online application designed for effective engagement and on-the-fly assessments.

### Content

The exam-wrapper used was composed of 4 fundamental questions relating to the objectives mentioned earlier:

1. Based on how demanding you thought the course is, and how long you spent preparing for this midterm, what would be your response to the following sentence:

*I am satisfied with my performance and grade received*

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
-------------------	----------	---------	-------	----------------

2. What type of questions on the midterm presented the most challenge to you? And why do you think that?
3. After looking over your midterm, why do you think you lost the most points?

Lack of understanding of the concept
Not sure what was the question asking
Careless or calculation mistakes
Not being able to apply concepts in new contexts

4. What changes to your study habits do you plan to make when preparing for the next midterm?

### Data analysis and exam-wrapper scoring

Data were collected and analyzed using statistical analyses in Excel. To quantify and normalize the exam-wrappers, each was given a score between 1 and 10, with 10 being the highest. Trying to eliminate any subjectivity, the scoring criteria was as follows: The student starts with a 10 and then:

- If the answer to any of the two close-ended questions do not agree with the answer to any of the two open-ended ones, the student loses 1 point for each. (Incoherence)
- If the student scored below average, yet is satisfied with his/her grade, he/she loses 2 point. (Discouragement)
- If the student did not identify appropriate changes to the study habits based on the answer to question 3, 3 points are lost. (lack of self-regulation and goal adjustment)
- The last two points are lost if the reflection shows any lack of relevancy, consistency, or objectivity.

The metric to assess students' performance and improvement was defined as the grade difference between the two taken exams. This was then plotted against the reflection quality.

## 1.6. Significance of the results obtained

As previously discussed, quantitative and qualitative results collected from wrappers usually reflect on the mutual benefits to both students and the teaching team to improve and enhance the served education [4]. The significance of the results of this study lies within the ability of assessing exam-wrappers in general (to what extent was this tool useful), and its impact in MECH393 in specific (was it helpful in terms of improving students' performance). That is, collected results serve as an insight, as well as to enhance students' metacognition, self-assessment, and self-regulation skills.

## Students Standing in both Midterms

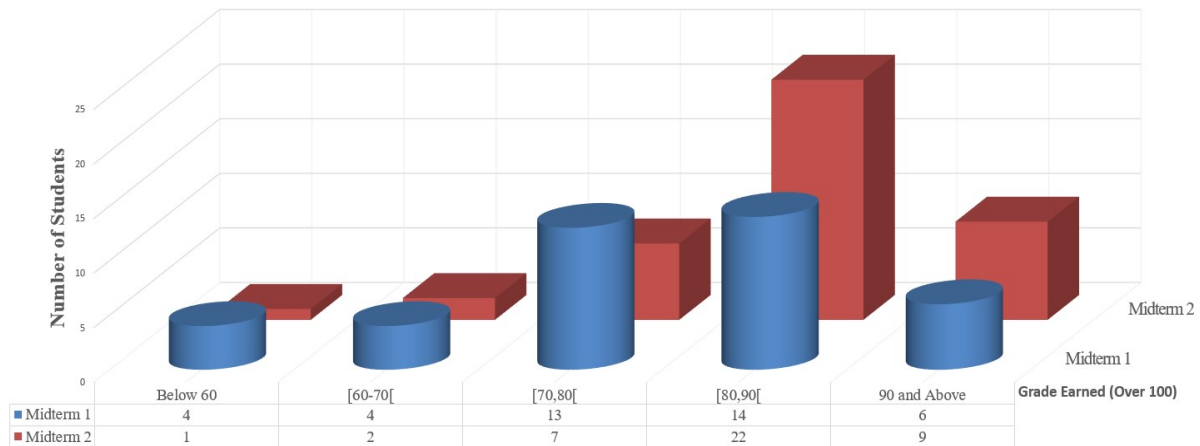


Fig. 1. Students' grades in each midterm

## 2. RESULTS AND DISCUSSION

Figure 1 shows that 80% of the students scored above 70 on the first midterm while this increased to 93% on the second midterm, with 76% of the students scored at least 80. Looking at the quality of reflection written by the students after doing the midterm, no specific correlation was observed (Fig. 2). However, in general, it seemed that most students who scored a relatively low grade, between 65% and 78%, put efforts to reflect on

the exam, and ended up with a reflection score between 6 and 8 out of 10. On the other hand, some of those who did well on the midterm, scored between 80% and 93% submitted brief reflections that were given low scores, between 3 and 5 out of 10. Figure 2 also shows that other outliers existed where 2 students scored low grades (52.5% and 71.5%), but also reflected poorly (1 and 2 reflection quality out of 10). Similarly, 4 students who did really well on their midterm (87.5%-95%), scored a very high reflection quality (9 and 10 out of 10).

## Overall Grade vs. Quality of Reflection

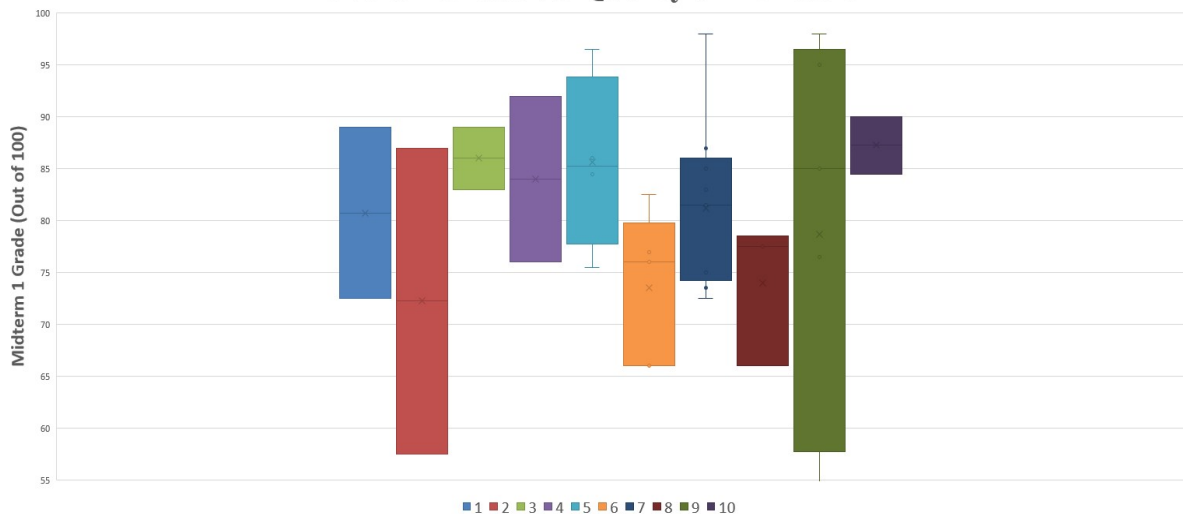


Fig. 2. Students' midterm 1 grades in relation to the quality of reflection

In considering the objective of whether students were able to self-assess their performance, this was only observed with students who scored a low grade. Analyzing their answers to the question of whether they were satisfied or not with their performance, most of

those students responded as highly unsatisfied, which was intuitive. However, when the exam-wrappers directed them at identifying the mistakes, some amazing responses were collected such as: "I think I focused more on the theory part and forgot to solve as much

exercises as possible”, “It was a fair exam but I should go back and review statics, which is where most of my mistakes came from”, ...etc. That is, when students were asked to identify specific mistakes, and were motivated to objectively answer exam-wrappers questions, they tended to spend time and effort going over their exam again and were actually able to spot specific mistakes they did. Therefore, most students, especially those who did not do well on their midterm, were able to self-assess themselves objectively.

However, this wasn't exactly the case with those who did relatively well on their exam. Although they did not fall into much mistakes, their responses were mainly: “I do not think any of the midterm's questions was challenging!”, “None of the questions was hard specifically”, “Although I did some mistakes, all questions were fair”, ...etc. It can be seen that because such students were satisfied with their performance, they did not put much effort in reflecting on their exam. Even though some of the mistakes made by those students might be attributed to major causes, such as lack of

understanding of a specific concept, students did not thrive at a better performance, and thus, did not objectively reflect the self-assessment trait.

Furthermore, to investigate the impact of exam-wrappers, students' second midterm grades, after submitting the midterm-wrappers, were plotted against the reflection quality (Fig. 3). A general increasing trend was observed where, students who submitted a relatively better reflection, scored a relatively higher grade in the second midterm. That is, for up to a 4/10 reflection quality, 4 students scored below 70%, 2 students scored between 70% and 80%, while only 3 students scored above 80% on the second midterm. As the reflection quality increased, students' grades got better where for all reflection quality of 5/10 and above, students scored above 80% with one exception of 71.25% for a reflection quality of 7/10. Moreover, 8 students who scored a grade above 90% on the second midterm submitted exam-wrappers of 7, 9, and 10 out of 10 qualities.

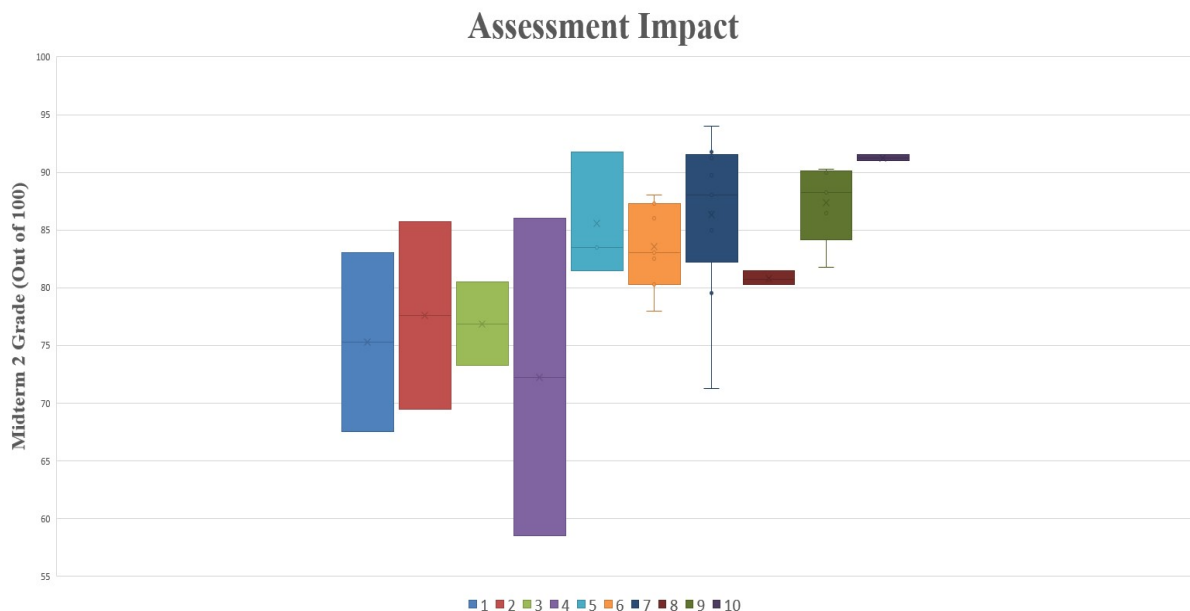


Fig. 3. Students' midterm 2 grades in relation to the quality of reflection

In considering another objective of the usefulness of exam-wrappers and whether such a tool generally has an impact or not, a trend proving its potential impact was observed (Fig. 3). After the fluctuations observed in the case of the first midterm, results were more normalized when comparing the reflection quality to the second midterm. That is, students who put more effort on the exam-wrappers ended up with higher exam scores afterwards. This was mainly attributed to the fact that students were able to reflect objectively on their mistakes, identifying those, and changing their study

behaviors to overcome those mistakes. When looking at some responses to that matter, students were able to identify the following mistakes: “careless or calculation mistakes”, “not being able to apply concepts in new contexts”, ...etc. Other more detailed responses and how students would adjust their study behaviors were: “better organizing my notes to spare time during the exam”, “spend more time on theory in preparation for the short answers”, “work in groups and solve practice problems with my friends”, ...etc.

What further proved the usefulness and impact of exam-wrappers was the other brief comments submitted by students who initially scored good grades in their first midterm. Such answers ranged from “not much” to “none”, although being have received a grade around 85%. Because those students identified their grades as satisfiable, their behavior was attributed to the fact that they felt that their studying strategy is perfect and that they will be able to achieve similar good scores in the second exam by following the same studying strategies. Conversely, especially those two students in particular, their grade dropped to 71.5% and 73.25% in the second midterm.

Thus, this study’s finding strongly suggest that exam-wrappers have a strong positive impact on students results. That is, students who put good amount of effort on wrappers, objectively, interpret their mistakes, and adjust their study behaviors accordingly, would potentially end up with better results. On the other hand, students who do not thoroughly analyze their mistakes, will not benefit from exam-wrappers, and might even detriment from this tool.

Finally, to get the full picture of how the reflections resulted in a change in students’ grades, improvement in grades between the first and second midterms were plotted against the quality of reflections (Fig. 4).

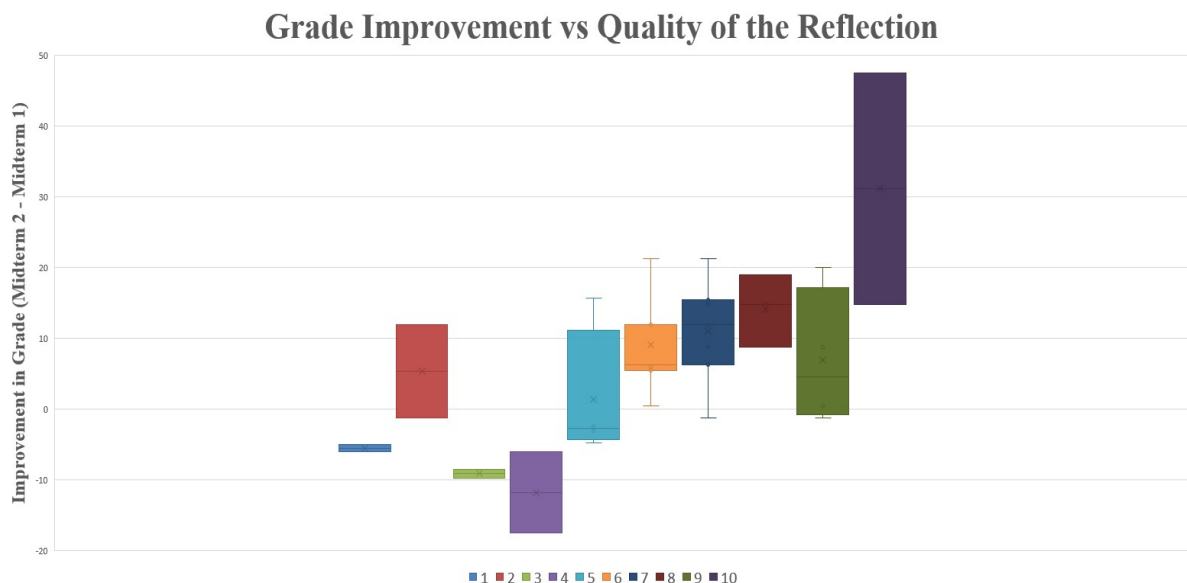


Fig. 4. Grades improvement in relation to the reflections’ quality

A positive improvement means that the student scored better on midterm 2 by the difference recorded, while a negative improvement means that the student did worse on the second midterm. Students with low reflection quality (8 students with a quality of 4 at most) show fluctuations, that is, only one student improved by 12.5 points while the rest showed scores deterioration, with one dropping by 17.5 points. Impressively, as the reflection quality got better, students showed higher grades’ improvement, with the exception of 3 students who showed a drop of 1.5, 3.25, and 5.25 points while scoring a 7, 9, and 5 reflection qualities respectively. On the other hand, great improvements were observed for reflection qualities of 7 and above, with the highest being a 47.5 increase in grade for a student who scored a 10/10 reflection (Fig. 2).

Lastly, in effort of investigating the impact of exam-wrappers on students’ metacognition, a general trend of grades improvement was observed with better qualities of reflection. Analyzing the answers to the last question

in the exam-wrapper, students who achieved a relatively high improvement responded: “I’ll try to put more emphasis on the concept instead of specific cases”, “I’ll try to understand the material more thoroughly and follow a clearer strategy to attack the problems”, “Start a lot earlier with studying and pay more attention in class”, “Read the theory more carefully for the short answers. This would help for questions that require you to know where equations came from and what fundamental theories are behind those equations”, ...etc. Generally speaking, students with such responses were able to exactly identify what should be changed regarding their study habits. By doing so, they ended up with a very high improvement between both midterms, with the highest, again, being 47.5 points increase.

Therefore, it is evident that the proper use of exam-wrappers helps students improve their metacognitive skills. This was seen as, firstly, when the professor illustrated the importance of exam-wrappers and how such a tool highly helps achieving more, and secondly,

by motivating and stimulating students to fill it as objectively as possible, they started to realize its potential positive effects. Then, by properly tailoring exam-wrappers at identifying mistakes and suggestions to change study habits accordingly, students were able to self reflect on their performance. The fact that students were able to adjust their study strategies according to the mistakes they spotted, it means that their metacognitive awareness increased, which eventually led, in this study's case, to better results.

### 3. CONCLUSIONS

The findings presented in this paper suggest that wrappers have a potential impact on students' learning in an engineering, machine element design, class. The presented analysis of student responses suggests engagement with the exam-wrappers and evidence of student awareness of effective and ineffective regulatory and learning processes. A link existed between how serious students took the exam-wrappers and how much effort they put on it, i.e. scoring high reflection quality, and the potential improvement in grade observed.

Further, students demonstrated mature reflections on their mistakes, and how they can adapt their study habits to improve in the future. This demonstrated the usefulness of exam-wrappers to obtain information about how students learn and study. This also showed that metacognitive advances are more potent when the target of learning, as well as the potential mistakes and ineffective studying strategies are pointed out to students.

One possible explanation for this strong correlation between the reflection quality and improvement in grade, especially for students who did not do relatively well on the first midterm, is that students were able to quantify the value and benefit of exam-wrappers in the development of their metacognitive skills. This is potentially the case especially when students are explicitly taught about them.

The study involved several limitations, one being that students lacked the awareness about the importance of metacognitive skills since they have not come across this term before, not even in their previous classes. It took some time and a lot of effort to convince them of how critical it is to adopt such skills. Furthermore, the class where exam-wrappers were implemented involved only two exams (a midterm and a final exam). That is, although this tool turned out to be effective in this class, due to the intensive effort put to help students realize the importance of such tools in developing metacognitive skills, it still requires further assessment in more demanding classes. This is because, the more the class would be demanding, more data can be collected, and

thus better conclusion can be drawn. Besides, due to the presence of open-ended questions, it was almost impossible to make sure that the study is objective. No matter how much effort was put to eliminate subjectivities, their presence was inevitable. As such, one suggestion would be to design the questions to be all closed-ended (check-box type of questions).

Since exam-wrapper is a relatively new concept, further research is still required to assess the exact contribution of exam-wrappers to students' performance. Although a considerable impact was proven to exist, it is essential to note that other studies found no contribution at all, where their collected data did not give any obvious conclusion. Thus, a more revised systematic approach to design wrappers seems essential to objectively assess exam-wrappers. One suggestion would be to devise an approach that involves triangulation among both, quantitative and qualitative, reported data. That is, designing wrappers in a way that allows to relate students' satisfaction, study strategies, and their scores performance. Furthermore, the development of a control group as a reference for comparison should be also in future exam-wrappers investigations. This would potentially help researchers to come up with more helpful, normalized, and credible data.

In closing, it is worth pointing out that educators are always bombarded by new teaching techniques and assessment activities to keep the education curve move forward. One of the appealing reflective exercises that improve students' performance is exam-wrappers due to its ease of implementation and potential positive incomes. Moreover, metacognitive skills are very essential for students' education, which can be potentially build using exam-wrappers. Our hypothesis was that by prompting students to reflect on their exams, they would be able to identify effective study strategies. When they adapt those strategies, they would end up with improved performance in later exams. Exam-wrappers were found to be effective to that matter where a strong relation was observed between reflection quality and grade improvement. Thus, exam-wrapper was proven to be an effective tool to improve metacognitive knowledge by improving self-assessment, goal setting, and self-regulation skills.

### Acknowledgements

This study was supported by the eLATE Teaching and Learning Improvement Funds from the Faculty of Engineering at McGill University.

## References

- [1] H. S. Schwartz and D. A. Schon, "The Reflective Practitioner: How Professionals Think in Action.," *Adm. Sci. Q.*, 1987.
- [2] D. Kolb, *Experiential Learning: Experience as the Source of Learning and Development*. 1984.
- [3] J. A. Turns, B. Sattler, K. Yasuhara, J. L. Borgford-Parnell, and C. J. Atman, "Integrating Reflection into Engineering Education," *2014 ASEE Annu. Conf. Expo. Indianapolis*, 2014.
- [4] K. J. Chew, H. L. Chen, B. Rieken, A. Turpin, and S. Sheppard, "Improving students' learning in statics skills: Using homework and exam wrappers to strengthen self-regulated learning," *ASEE Annu. Conf. Expo. Conf. Proc.*, vol. 2016–June, 2016.
- [5] M. Craig, D. Horton, D. Zingaro, and D. Heap, "Introducing and Evaluating Exam Wrappers in CS2," *Proc. 47th ACM Tech. Symp. Comput. Sci. Educ. - SIGCSE '16*, pp. 285–290, 2016.
- [6] C. Watson and F. W. B. Li, "Failure rates in introductory programming revisited," in *Proceedings of the 2014 conference on Innovation & technology in computer science education - ITiCSE '14*, 2014.
- [7] R. Lister *et al.*, "A multi-national study of reading and tracing skills in novice programmers," in *Working group reports from ITiCSE on Innovation and technology in computer science education - ITiCSE-WGR '04*, 2004.
- [8] D. Zingaro and L. Porter, "Peer Instruction in computing: The value of instructor intervention," *Comput. Educ.*, 2014.
- [9] C. Watson, F. W. B. Li, and J. L. Godwin, "No tests required: comparing traditional and dynamic predictors of programming success," in *Proceedings of the 45th ACM technical symposium on Computer science education - SIGCSE '14*, 2014.
- [10] M. Guzdial, "Exploring hypotheses about media computation," in *Proceedings of the ninth annual international ACM conference on International computing education research - ICER '13*, 2013.
- [11] A. R. Basawapatna, K. H. Koh, and A. Repenning, "Using scalable game design to teach computer science from middle school to graduate school," in *Proceedings of the fifteenth annual conference on Innovation and technology in computer science education - ITiCSE '10*, 2010.
- [12] M. C. Lovett, "Make exams worth more than the grade," *Using Reflect. metacognition to Improv. student Learn. Across Discip. across Acad.*, vol. 18, 2013.
- [13] M. V. Achacoso, "Post-test analysis: A tool for developing students' metacognitive awareness and self-regulation," *New Dir. Teach. Learn.*, 2004.
- [14] B. Stephenson, M. Craig, D. Zingaro, D. Horton, D. Heap, and E. Huynh, "Exam Wrappers: Not a Silver Bullet," *Proc. 2017 ACM SIGCSE Tech. Symp. Comput. Sci. Educ. - SIGCSE '17*, pp. 573–578, 2017.
- [15] P. G. Gezer-Templeton, E. J. Mayhew, D. S. Korte, and S. J. Schmidt, "Use of Exam Wrappers to Enhance Students' Metacognitive Skills in a Large Introductory Food Science and Human Nutrition Course," *J. Food Sci. Educ.*, vol. 16, no. 1, pp. 28–36, 2017.
- [16] Krathwohl, "A Revision of Blooms Taxonomy," *Am. J. Psychol.*, 2009.