ORIGINAL ARTICLE

# Use of Exam Wrapper in Internal Medicine Residency Training in Two Tertiary Private Hospitals: A Pre-experimental Study

Janice Jill K. Lao, MD[1] and Erlyn A. Sana, PhD[2]

[1]*Philippine General Hospital, University of the Philippines Manila*
[2]*National Teacher Training Center for the Health Professions, University of the Philippines Manila*

## ABSTRACT

**Background and Objective.** Self-assessment and metacognition can be practiced with an exam wrapper (EW). EW is a structured, metacognitive, and self-regulated learning strategy that involves guided self-reflection on an exam already taken to improve study habits. This research describes how internal medicine (IM) residents at two tertiary private hospitals performed in written examinations using an EW. The relationship between the residents' metacognition, the exam wrapper, and exam performance was also determined.

**Methods.** This study employed a pre-experimental pre- and post-test design. The EW was constructed and tested for validity and reliability. It included (1) a description of study habits, (2) accuracy in self-efficacy perception and exam score prediction, (3) perceived reasons for exam mistakes, and (4) future study plans of residents. A complete enumeration of 24 IM residents was conducted. Respondents completed the Metacognitive Awareness Inventory (MAI) at the beginning of the study. The intervention consisted of (1) residents taking Exam 1: Gastroenterology, followed by EW; (2) Exam 2: Endocrinology and EW; then (3) Exam 3: Oncology, EW, and MAI. Scores were compared using a paired t-test or analysis of variance (ANOVA). The relationships between metacognition scores, the EW, and exam performance were determined using the Pearson correlation coefficient. The level of significance was set at $p < 0.05$.

**Results.** The final EW comprises 16 items, with overall indices of content validity ratio of 0.72 and item-rated content validity of 0.8. The internal consistency coefficient is 0.65 (Kuder-Richardson 20). Nineteen out of 24 residents (79.17%) completed the study. Mean exam percentage scores were 57.97%, 42%, and 51.16% for Exams 1, 2, and 3, respectively. Exam 2 differed significantly from the other two exams ($p = 0$ and $p = 0.04$). EWs for the first two exams were not significantly different and revealed: (1) top study habits included studying right before an exam and skimming the textbook; (2) 68.42% vs. 63.16% accuracy of self-efficacy perception; (3) 26.32% vs. 31.58% accuracy of grade estimation; (4) 31.58% vs. 26.32% accuracy of error analysis; (5) most errors were due to not reading about the topic, and (6) most planned to "read more." Mean MAI scores were 36.79 ± 9.10 (pretest) and 36.05 ± 9.44 (post-test) ($p = 0.81$). All correlations were not statistically significant.

**Conclusion.** Residents performed poorly during exams, crammed their studies, preferred low-impact learning strategies, and lacked self-reflection skills and metacognition monitoring. Time issues related to reading or studying were common. There is no significant relationship between metacognition score and exam wrapper use or exam performance in IM residency trainees.

*Keywords: exam wrapper, metacognition, self-regulation, Internal Medicine residents*

Corresponding author: Janice Jill K. Lao, MD
Division of Nephrology
Department of Medicine
Philippine General Hospital
University of the Philippines Manila
Taft Avenue, Ermita, Manila 1000, Philippines
Email: jklao@up.edu.ph
ORCiD: https://orcid.org/0009-0004-3837-9703

## INTRODUCTION

Passing a written specialty board examination is still the norm to obtain a specialty certificate, locally or internationally. The Philippine College of Physicians (PCP) serves as the certifying body for internal medicine (IM) practice in the country. According to its training handbook, written examinations are integral to a resident's summative assessment, performance in the residency in-service written examination or RITE is considered one of the performance indicators of a training institution, and performance in the Philippine Specialty Board of Internal Medicine (PSBIM) written examination is deemed as the final test of a trainee's achievement of the recommended learning outcomes.[1]

Self-assessment in the form of exam wrappers has been employed in some medical schools to improve written examination performance.[2] An exam wrapper (EW) is a structured metacognitive and self-regulated learning strategy that involves a guided self-reflection of an exam already undertaken to improve study habits.[3,4] It prompts students to learn from the exam rather than only for the exam.[4] Increased metacognitive awareness in students and better performance on subsequent examinations were reported in that initial study among university education students.[5] EWs were also found to be useful in recognizing effective study habits among higher education students across different courses, including health sciences courses.[4,6-9]

An EW follows Zimmerman's self-regulation strategy of learning (SRL) that is divided into three (3) stages: (1) forethought (appraising pre-task self-efficacy, establishing goals and strategies); (2) performance (doing the task with self-monitoring); and (3) self-reflection (done after the task). It allows students to control their learning and includes processes of self-assessment and reflection.[10] By guiding students through the process of reflection up to the creation of a strategy for improvement, we as faculty are also motivating them to be better learners.[2]

Another important component of an EW, and the SRL strategy, is metacognition. Metacognition refers to higher-order thinking that involves active control or regulation over the cognitive processes engaged in learning and overseeing whether a cognitive goal has been met.[11] Self-assessment increases metacognition, making learners more proficient at evaluating their progress toward completing a task, an important skill of self-regulated learning.[12] Effective metacognition is a critical element in the development of self-awareness and the creation of a culture that promotes self-directed, lifelong learning.[13] It is a strong predictor of academic success as students become more strategic in their learning, such as focusing more on learning new information rather than dwelling on known information.[14,15] Several studies have shown the value of good metacognitive skills in academic performance.[15-22]

Research indicates that while metacognition scores and exam performance improve with the use of EWs, results are inconsistent across studies.[2,21,23-25] Among health sciences students, no studies have explored the correlation between EW use and metacognition scores.[2,7,25] However, students generally perceive EWs as beneficial.[2,4] Enhancing study habits and metacognition not only can boost cognitive knowledge but might also foster self-regulation skills essential for lifelong learning. EWs can be effective in developing metacognition, self-regulation, and self-assessment skills, contributing to successful adult learning alongside self-determined learning.[26] Study on the relationship between self-regulation, metacognition, and the national surgical in-training examination performance of US surgical residents showed a significant correlation among the three.[18] Even in studies with weaker correlations, metacognition predicted academic performance when controlling for intelligence or prior academic success.[19,20] These findings underscore the potential for improving metacognitive skills to enhance exam performance, particularly for students with less than exemplary academic backgrounds, and this type of trainee is exactly what we have in our institution.

Despite exhaustive research, there is a lack of studies assessing the effectiveness of exam wrappers (EWs) among graduate medicine trainees, who are generally more mature and should exhibit greater metacognitive awareness than undergraduates.[22] Higher order of learning, including metacognition, is expected from an IM resident to train their decision-making and critical thinking skills. Pitfalls were also identified in previous EW and metacognitive studies, like using a metacognition assessment tool that was not validated, employing an EW without a proper feedback activity, or answering EWs only once.[21,24,27] These issues were thus addressed in the conduct of this study. This study aimed to describe how IM residents in two tertiary private hospitals regulated themselves to ensure satisfactory performance in written examinations through a validated exam wrapper. Additionally, we looked at the relationship between metacognition and exam wrapper use, and between metacognition and exam performance. We hypothesized that the use of exam wrappers would increase metacognition and eventually exam scores.

## MATERIALS AND METHODS

### Study Design

This was an interventional study composed of two phases. The initial phase (Phase 1) of creating the EW adopted a descriptive quantitative design. The interventional phase (Phase 2) of administering the EW followed a quantitative pre-experimental pretest-posttest design.[28]

### Study Setting

Phase 1 was done in a local tertiary public hospital, while Phase 2 was done in two local tertiary private hospitals. The two private hospitals were selected based on their 2023 RITE 1 scores, 2023 PSBIM passing rate, and ease of

communication with the training core team. As both are private hospitals, confounding factors inherent to the type of training institution that might affect the trainee's cognitive and affective domains and be a source of variation in their metacognition scores, EW answers, and examination scores were minimized.

## Population and Sampling Technique

Both phases used total enumeration sampling, with participants providing informed consent and given the option to withdraw at any time. Phase 1 involved 19 senior residents during their comprehensive examination, while Phase 2 included 15 residents from Hospital 1 and nine from Hospital 2 (n = 24). Ultimately, 19 residents (four first year, nine second year, six third year) completed the Phase 2 study, resulting in a 21% dropout rate (5 of 24) (Figure 1). Each Phase 2 participant was assigned a code for anonymity, accessible only to the researcher. Phase 2 participants had an average age of 29.74 ± 2.51 years. Most were from medical schools with >90% passing rates on the October 2024 medical board exam, with many ranking in the 40-50% range of their batch. The average board rating was 78.84 ± 2.84, and participants had graduated from medical school an average of 4.21 ± 1.36 years before residency. During the study, 94.73% had to give lectures or presentations, 36.84% participated in quiz competitions, and 15.79% were absent for at least a day.

## Materials and Data Collection Procedure

### *Phase 1: Exam Wrapper Validation and Reliability Testing*

Several EWs used in different studies were reviewed.[3-6,8,9,21,23,24,29-31] Some questions were adapted and modified to specifically cater to IM residents and to include all three self-regulation processes of forethought, performance, and self-reflection described by Zimmerman.[10] The EW was organized into three parts: Part I focused on exam preparation, exploring respondents' study habits, time management, and self-efficacy with questions like, "Did you feel prepared for the exam?" and "How much time did you spend studying?" Part II encouraged residents to reflect on their performance, identify mistakes, and understand strategies for improvement, including questions like, "How did your score compare to your expectations?" Part III required creating a study plan for future exams, promoting the forethought process by integrating insights from Parts I and II. Table 1 summarizes the constructs and item distribution of the pilot EW (Appendix A).

Participants received the pilot EW, an EW evaluation form, and a brief study description and self-regulation strategy definition after their exam. The evaluation form included guide questions about the SRL strategy, which could be answered with "yes" or "no," with allowed space for participants to indicate relevant question numbers. The last page contained questions about the questionnaire's format, length, wording, and space for additional comments. A focus group discussion (FGD) was held afterward to clarify their responses. Answers from the evaluation sheet were encoded in Microsoft Excel®, with a point awarded if an item was deemed to represent an SRL construct.

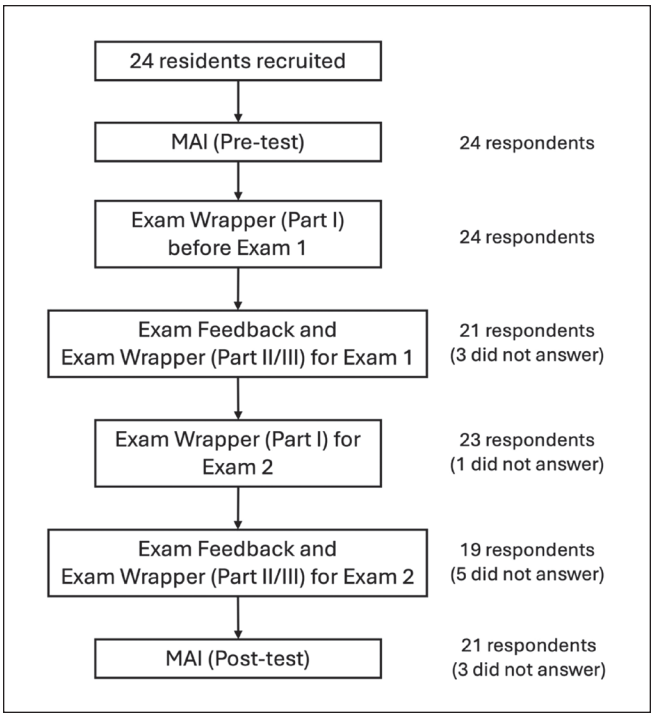


**Figure 1.** Flowchart of respondents for Phase 2.

**Table 1.** Blueprint of the Pilot Exam Wrapper (EW)

| Constructs | Number of items |
|---|---|
| *I: Exam preparation* | 9 |
| *II: Evaluation of exam performance* | 5 |
| *III: Study plan for the next exam* | 3 |
| *Total* | 17 |

Content validity (CV) for self-regulated learning (SRL) phases was measured using the content validity ratio (CVR) formula: CVR = (Nr – N/2)/(N/2), where "Nr" is the number of raters who answered "yes" and N is the total number of raters.[32] A CVR of 0.42 or higher is valid for 20 raters according to the Lawshe table.[33] The item-rated CV index (I-CVI) was calculated as I-CVI = (Nr)/(N), with revisions made if I-CVI is between 0.70-0.79, and items with I-CVI <0.7 eliminated.[34] The overall CV index was determined by averaging the CVR and/or I-CVI.[34,35] Reliability was assessed using the Kuder-Richardson Formula 20 (KR-20) for internal consistency, with a target reliability coefficient of 0.70 or higher.[35,36] Coefficient interpretations ranged from very low (0.00-0.20) to very high (0.80-1.00).[37] Based on validity and reliability testing, a final version of the EW (Appendix B) was revised and encoded into a Google Form for easier data analysis. Each exam wrapper is divided into two Google

Forms, one for Part I and another for Parts II and III. The link is shared with each institution's coordinator. Each exam has its own exam wrapper link.

### *Phase 2: Assessing the Effect of Exam Wrapper*

The metacognition of all residents was determined using the true-false version of the Metacognitive Awareness Inventory (MAI) (Appendix C), a 52-item questionnaire focused on metacognitive knowledge (17 items) and metacognitive regulation (35 items).[38] The knowledge domain is divided into declarative (8 items), procedural (4 items), and conditional (5 items) knowledge, while the regulation domain includes planning (7 items), information management strategies (10 items), comprehension monitoring (7 items), debugging strategies (5 items), and evaluation (6 items). Each correct "true" response earns one point, with a maximum score of 52. The modified version demonstrated acceptable internal consistency (Cronbach's alpha = 0.7), with the knowledge and regulation domains showing consistency scores of 0.7 and 0.8, respectively.[7] The true-false version of the MAI was encoded into a Google Form, and the link was shared with each institution's study coordinator.

Syncing of long exam schedules was done before the start of questionnaire administration. The included exams were in Gastroenterology (Exam 1), Endocrinology (Exam 2), and Oncology (Exam 3), as these were the exams that still had to be administered for the year for both institutions. Both institutions administered the same exam questions in electronic form. Exams, however, were not given simultaneously due to some institution-specific circumstances and the sudden cancellation of work, but the chronology was followed.

The MAI questionnaire was administered at the beginning and end of the study period (before Exam 1 and after Exam 3). The revised EW (Appendix B) was given after Exam 1 and Exam 2 in two parts: Part I immediately after the exam and Parts II and III after a scheduled review 1-2 weeks later. Written exam performance for Exams 1, 2, and 3, along with demographic data (age, training year, medical school, ranking, board rating, and years since graduation), was obtained in coded format. Additionally, the institution provided information on conference schedules, lectures, CME activities, submission deadlines, and any absences during the study period, as these could affect examination preparation. Residents are required to take one week of vacation leave, scheduled by the institution.

A summary of both Phase 1 and Phase 2 is shown in Figure 2.

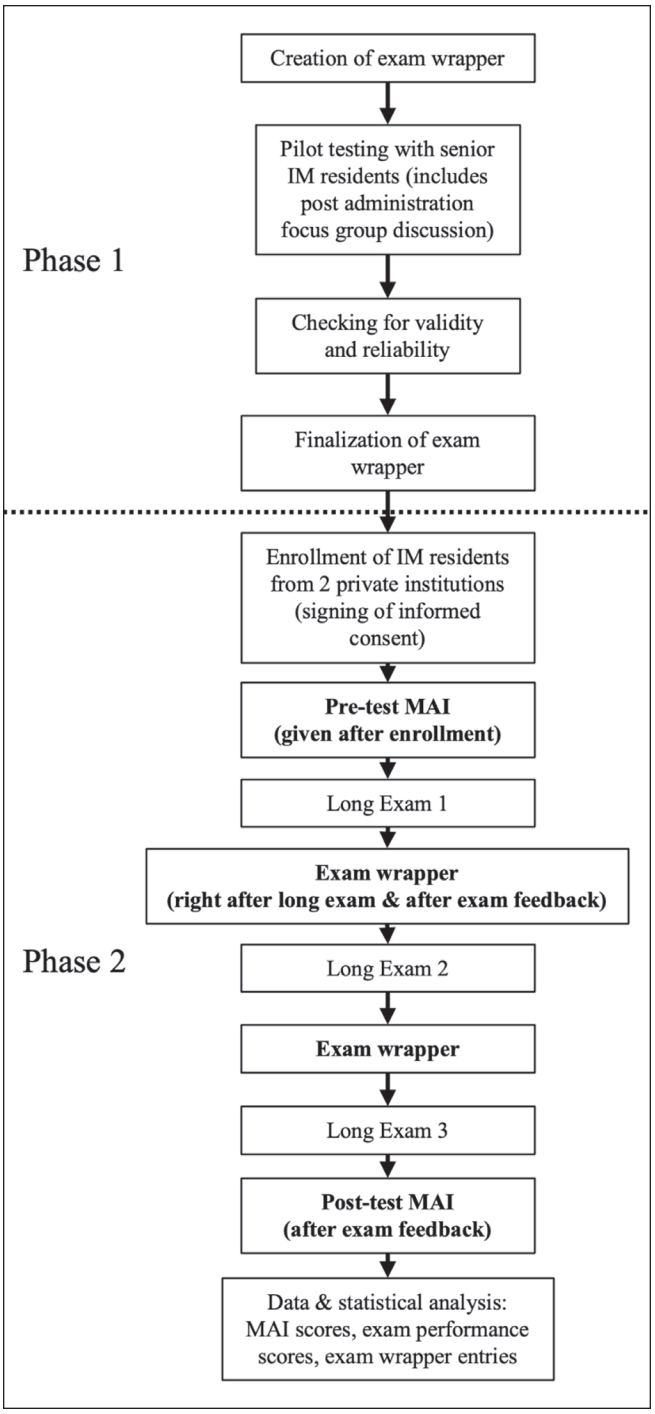


**Figure 2.** Study flowchart.

## Analysis of Data

Baseline variables and answers to the EW were analyzed through frequencies and percentage distributions, means and standard deviations, as appropriate. Free text data collected from the EW were independently reviewed, categorized, and counted. Categories were based on the presence of any uniform themes among the answers (e.g., time management issues, etc.). A rubric for evaluating the EW use (Appendix D) for Exam 2 was also created and employed to assess the quality of EW use by the residents. Metacognition scores were analyzed with a paired t-test, while exam performance scores were compared using ANOVA at $p < 0.05$. Pearson correlation coefficients at $p < 0.05$ determined relationships between changes in metacognition scores, EW rubric scores, and exam performance.

### Study Limitations

The small sample size limits the study's generalizability. The pre-experimental design lacks a comparison group and randomization, making it impossible to definitively establish a cause-and-effect relationship from the exam wrapper treatment. Verifying the truthfulness of responses to the EW and true-false MAI format was challenging despite reminders. Other training activities, such as conferences and lectures, that could confound results were noted but did not include specific timelines or themes.

### Ethical Considerations

This study was submitted to the University of the Philippines Manila Research Ethics Board (UPM REB) and each hospital's Technical Review Board (TRB) and Ethical Research Committee (ERC) for approval. Its social value lies in improving the quality of internists, which is expected to enhance patient care and national health. Written informed consent was read, signed, and possessed by all enrolled residents, and our copy was securely stored until the study concludes. The residents are not considered a vulnerable group and are expected to benefit from the study's findings. There are no foreseeable risks to participants. The study complies with the Data Privacy Act of 2012, ensuring no participant names are used in analysis or reporting.[39] Participants could decline to answer questions or withdraw anytime. The purposive sampling method used is fair, and care was taken to treat all participants equally, regardless of their background. Participants were informed about the study's methods and results, and full disclosure of any researcher interests was provided in the consent forms. Honesty in responses was emphasized to maintain transparency.[40]

### Conflict of Interest

There is no perceived conflict of interest for the investigator involved in this study. Neither the main investigator nor the thesis panelists were directly involved in the making of the exams, checking of the exams, or recording of grades submitted to the study.

## RESULTS

### Validation of Exam Wrapper for Residency Training in Internal Medicine

The ratings of all 19 senior residents were subjected to validity testing as their answers were all deemed complete and clear. All phases of the self-regulation strategy were elucidated in the pilot test EW (Appendix A) as shown in Table 2 according to their content validity scores, except for Part I #4 (CVR 0.37), Part III #2 (CVR 0.37), and Part III #3 (CVR -0.47).

**Table 2.** Content Validity and Interpretations of the Exam Wrapper on the Phases of Self-regulation Learning Strategy

| Items | Nr* | CVR** | Interpretation | I-CVI*** | Interpretation |
|---|---|---|---|---|---|
| ***Part I*** | | | | | |
| 1 | 17 | 0.79 | Retain | 0.89 | Appropriate |
| 2 | 15 | 0.58 | Retain | 0.79 | Needs Revision |
| 3 | 15 | 0.58 | Retain | 0.79 | Needs Revision |
| 4 | 13 | 0.37 | Eliminate | 0.68 | Eliminate |
| 5 | 17 | 0.79 | Retain | 0.89 | Appropriate |
| 6 | 15 | 0.68 | Retain | 0.79 | Needs Revision |
| 7 | 18 | 0.89 | Retain | 0.95 | Appropriate |
| 8 | 16 | 0.68 | Retain | 0.84 | Appropriate |
| 9 | 19 | 1.00 | Retain | 1.00 | Appropriate |
| ***Part II*** | | | | | |
| 1 | 19 | 1.00 | Retain | 1.00 | Appropriate |
| 2 | 17 | 0.79 | Retain | 0.89 | Appropriate |
| 3 | 19 | 1.00 | Retain | 1.00 | Appropriate |
| 4 | 19 | 1.00 | Retain | 1.00 | Appropriate |
| 5 | 16 | 0.68 | Retain | 0.84 | Appropriate |
| ***Part III*** | | | | | |
| 1 | 19 | 1.00 | Retain | 1.00 | Appropriate |
| 2 | 13 | 0.37 | Eliminate | 0.68 | Eliminate |
| 3 | 5 | -0.47 | Eliminate | 0.37 | Eliminate |
| ***Overall CVI*** | | 0.70 | | 0.85 | |

* *Number of residents that evaluated the item as relevant*
** *CVR or Content Validity Ratio = (Nr–N/2)/(N/2) with 19 evaluators at (N=19)*
*** *I-CVI or item-level content validity index = (Nr)/(N)*

The details of the questions with poor content validity scores are the following:

Part I #4: What percentage of your exam preparation was spent studying alone or with others?

Part III #2: What else can we, your consultants, do to help support your learning and preparation for the next exam?

Part III #3: Do you find this exam wrapper activity useful? Why or why not?

Part I #4 was removed and incorporated as "group study" under Part I #5. Questions in Part III, which offered useful feedback for the trainer, were retained. Questions 2, 3, and 6 from Part I had an I-CVI that fell within the range of 0.70-0.79, leading to minor revisions. The questionnaire initially received a CVI of 0.70 (CVI of relevance) and 0.85 (I-CVI) before editing, and a CVI of 0.72 and 0.86 after removing Part I #4. If excluding Part III, the CVI improved to 0.81 and 0.91.

Questions 1 and 9 from Part I and question 1 from Part III were found to have a high CVR for the forethought phase (0.79, 1.00, and 1.00, respectively). The specific questions are enumerated below:

- Did you feel that you were prepared for the exam?
- How well do you think you did in this exam?
- Based on your responses to the questions above, name at least three things you plan to do differently in preparing for the next exam.

Questions 2, 3, 5, and 6 of Part I and questions 1 to 5 of Part II were all found to have significant CVR (>0.42) for the performance phase and are enumerated below:

- Which of the following best describes your study pattern for this exam?
- Approximately how much time did you spend preparing for this exam?
- Which of the following did you do to prepare for the exam?
- How much of your study plan were you able to carry out in this exam?
- How did your actual score compare to how you thought you did on the exam after taking it?
- How did this exam score compare to your last exam score?
- As you look over your graded exam, analyze where/how you lost points.
- Do you know all of the ways that you can get help in studying?
- Do you feel that you have made a good assessment of your learning habits and know how to adjust your approach?

Lastly, questions 7 and 8 from Part I and questions 1, 3, and 5 from Part II significantly represented the self-reflection phase (CVR >0.42) and are enumerated below:

- Do you think these study plans improved your study habits for this exam?
- If you were unable to carry out all or most of your study plan from the previous exam wrapper, what were the reasons?
- How did your actual score compare to how you thought you did on the exam after taking it?
- As you look over your graded exam, analyze where/how you lost points.
- Do you feel that you have made a good assessment of your learning habits and know how to adjust your approach?

**Table 3.** Blueprint of the Final Exam Wrapper (EW)

| Constructs | Number of items |
|---|---|
| ***I: Exam preparation*** | 8 |
| ***II: Evaluation of exam performance*** | 5 |
| ***III: Study plan for the next exam*** | 3 |
| ***Total*** | 16 |

For internal consistency, the KR-20 of the ratings given was found to be 0.65. Though below the 0.70 research target, this still falls within the moderate to high reliability range of 0.6 to 0.8. This was taken into consideration in the revision of the EW.

Out of 19 panelists, 17 (89%) found the EW acceptable in length, clarity, organization, format, and choice adequacy. Suggestions included adopting a digital format, improving bullet design, adding more space for answers, making grammatical corrections, and clarifying the item on point loss (Part II #3), which may confuse some. These inputs, along with content validity and reliability results, were considered in the reformatting of the new version of the exam wrapper. The final EW (Appendix B) was decreased to 16 items, with its blueprint shown in Table 3.

## Performance of Residents in the Written Examinations and Effects of Exam Wrapper on their Study Habits, Perceived Causes of Exam Mistakes, and Corrective Plans of Residents

The mean percentage scores for Exams 1, 2, and 3 were significantly different ($p = 0.003$), with Exam 1: Gastroenterology having the highest mean and Exam 2: Endocrinology the lowest (Table 4). Significant differences were found between Exam 1 vs Exam 2 ($p = 0$) and Exam 2 vs Exam 3 ($p = 0.04$), while Exam 1 and Exam 3 showed no significant difference ($p = 0.18$). A trend of increasing mean scores with higher training levels was observed. The number of passing residents per exam was 10 (52.63%) for Exam 1, 3 (15.79%) for Exam 2, and 5 (26.32%) for Exam 3. Overall, 11 residents (57.89%) passed at least one exam, six (31.56%) passed at least two, and only one (5.26%) passed all three. Pass rates varied by medical school and ranking, with one resident failing all exams despite being from a top-ranking school according to PSBIM passing rates. But excluding that resident, the rest who belonged to a top-performing medical school (five residents) were indeed able to pass at least one

**Table 4.** Means and Standard Deviations of Residents in the Exam (n=19)

| Variable | Overall, n = 19 Mean (SD) | Level I, n = 4 Mean (SD) | Level II, n = 9 Mean (SD) | Level III, n = 6 Mean (SD) | *p*-value |
|---|---|---|---|---|---|
| ***Exam 1 in Gastroenterology*** (total item: 70) | 57.97 (16.52) | 41.07 (4.42) | 53.65 (10.13) | 75.71 (7.17) | 0.003 (using ANOVA) |
| ***Exam 2 in Endocrinology*** (total item: 50) | 42.00 (9.30) | 37.00 (5.29) | 43.11 (10.87) | 43.67 (11.20) | |
| ***Exam 3 in Oncology*** (total item: 50) | 51.16 (15.62) | 38.50 (5.26) | 53.11 (15.46) | 56.67 (16.08) | |

**Table 5.** Comparison of Study Habits Employed for Exams 1 and 2 (n=19)

| Variables in the Exam Wrapper | Exam 1 Mean (SD) or Frequency (%) | Exam 2 Mean (SD) or Frequency (%) | *p*-value |
|---|---|---|---|
| ***Study hours*** | 5.38 (5.83) | 9.82 (18.34) | 0.33 |
| ***Study pattern*** | | | 0.74 |
| spaced out | 6 (31.58%) | 7 (36.84%) | |
| right before exam | 13 (68.42%) | 12 (63.16%) | |
| ***Other learning strategies utilized**** | | | |
| skimming the textbook | 12 (63.16%) | 13 (68.42%) | 0.74 |
| first thorough reading of the textbook | 4 (21.05%) | 2 (10.53%) | 0.39 |
| rereading the textbook | 3 (15.79%) | 2 (10.53%) | 0.64 |
| reviewing notes | 11 (57.89%) | 7 (36.84%) | 0.20 |
| other medias | 2 (10.53%) | 3 (15.79%) | 0.64 |
| self-quizzing | 6 (31.58%) | 4 (21.05%) | 0.48 |
| study actual cases | 7 (36.84%) | 5 (26.32%) | 0.50 |
| reading medical journals/guidelines | 2 (10.53%) | 2 (10.53%) | 1.00 |
| drawing concept maps/algorithms | 3 (15.79%) | 0 (0%) | 0.08 |
| group study | 0 (0%) | 2 (10.53%) | 0.16 |
| ***High impact strategies employed*** | 12 (63.16%) | 12 (63.16%) | 0.77 |

** The percentage of learning strategies used adds up to >100% because each student was allowed to check more than one of the provided options.*

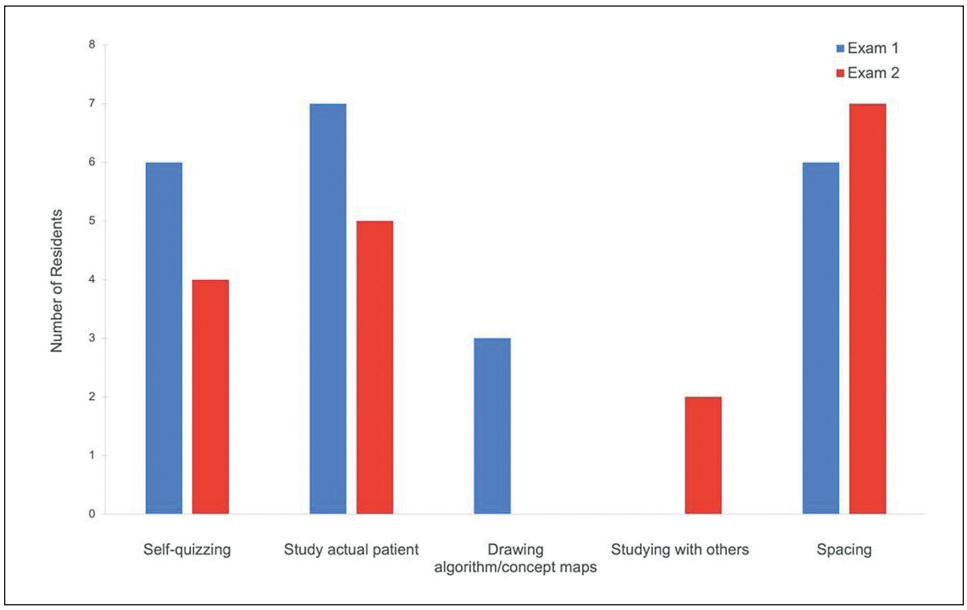


**Figure 3.** High-impact learning strategies cited by residents.

exam, and this includes the resident who was able to pass all three exams. It is also interesting to note that the three residents with a medical board rating of 83 and above were able to pass at least two out of their three exams.

The EWs were able to show the study habits (Table 5) of the involved residents. Residents spent only <10 hours on studying, with most preferring to study right before an exam and studying alone. Most residents just skimmed their textbook, followed by reviewing their notes for both EWs.

The high-impact strategies used by the 12 residents for Exams 1 and 2 are illustrated in Figure 3. Spacing was the most common strategy for both exams, while group study was

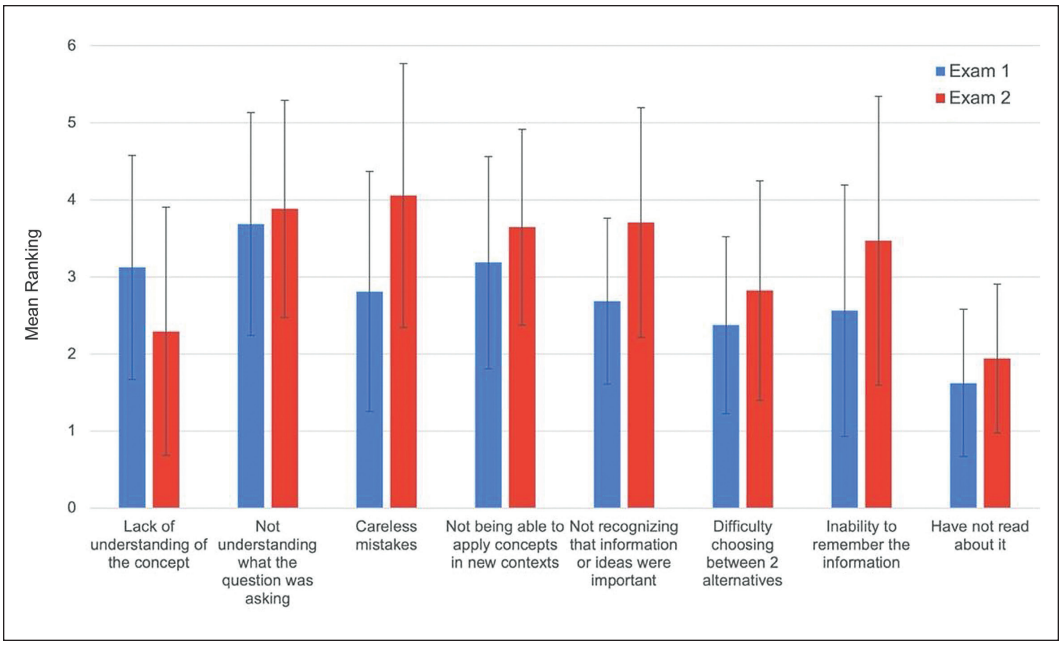


**Figure 4.** Mean ranking of specific causes of exam mistakes on Exam 2 compared to Exam 1.

only used in Exam 2. Six residents (31.58%) adopted more high-impact strategies for Exam 2, whereas five (26.32%) used more low-impact strategies. A paired t-test showed no significant difference (p = 0.77) between the two exams regarding high-impact strategy usage. Of the six residents who increased high-impact strategies, only two (33.33%) passed Exam 2—one using self-quizzing and experiential learning, and the other spacing and experiential learning. Among the four residents who passed Exam 2, one did not use any high-impact strategies, while the others used spacing, self-quizzing, or a combination of self-quizzing and case studies.

Only three residents (15.79%) claimed to be prepared for both Exam 1 and Exam 2, but none were consecutively prepared for both. All three who claimed readiness for Exam 1 passed it, while all three failed Exam 2. Reviewing their past scores, two of the three who prepared for Exam 1 passed their previous Gastroenterology exam, and two of the three who prepared for Exam 2 passed their previous Endocrinology exam. Self-efficacy accuracy was 68.42% for Exam 1 and 63.16% for Exam 2 (p = 0.74). Grade estimation was accurate for five (26.32%) residents in Exam 1 and six (31.58%) in Exam 2 (p = 0.49). Seven residents (36.84%) improved their grade estimation for Exam 2, while six (31.58%) got worse.

Most residents classified their exam errors as being due to not being able to read for both exams. Figure 4 shows the ranking of the types of errors for both EWs.

Help seeking knowledge was found to be >50% for both EWs (52.63% and 57.89%), with an increasing trend. One resident from each EW (not the same person) claimed to not know the ways to seek help, yet did not give any answer when asked about how his/her consultant can help. Specific suggestions on what kind of help the consultant staff can provide to improve their exam performance from both EWs are shown in Table 6.

The top response for the proposed study plan for both EWs is "reading more" (Table 7). Only 11 residents (61.11%) attempted to carry out their study plan from their first EW, with a mean execution of 20.32% ± 26.85%, ranging from 1 to 80%. Six (54.55%) of these residents perceived improvement in study habits. Reasons for non-execution of the study plan are shown in Table 8, with time issues topping the list [6 out of 8 (75%)].

**Table 6.** Responses Given by Students to the *"What else can we, your consultants, do to help support your learning and preparation for the next exam?"* Question (N = 19)

| | Exam 1 | Exam 2 |
|---|---|---|
| *More practice quizzes* | 1 | 4 |
| *More lectures* | 2 | 4 |
| *Teaching rounds* | 1 | 1 |
| *Use of updated guidelines* | 2 | 0 |
| *Give clinical scenarios/case-based learning* | 1 | 1 |
| *Allot protected review/study sessions* | 1 | 3 |
| *Time lectures with upcoming exam* | 1 | 0 |
| *Give coverage/important chapters* | 1 | 2 |
| *Time management strategies* | 1 | 0 |
| *Exam taking strategies* | 1 | 0 |
| *Handouts* | 0 | 1 |
| *Use PSBIM-like questions* | 0 | 1 |
| *None* | 7 | 4 |

**Table 7.** Proposed Study Plans for the Next Exam (N = 19)

| | Exam 1 | Exam 2 |
|---|---|---|
| *Read/study more* | 5 | 9 |
| *Concentrate/focus on studying* | 0 | 2 |
| *Repeat review/re-read* | 0 | 2 |
| *Read/study every time it is permitted or everyday* | 3 | 2 |
| *Make time/detailed study schedule* | 4 | 6 |
| *Study fast* | 0 | 1 |
| *Distinguish important concepts* | 0 | 1 |
| *Understand concepts properly before moving on/ analyze more* | 4 | 4 |
| *Prioritize topics* | 1 | 0 |
| *Remember/memorize* | 0 | 1 |
| *Self-quizzing* | 4 | 2 |
| *Study previous exam* | 2 | 0 |
| *Study tables and diagrams more* | 2 | 0 |
| *Note writing/summaries for easier rereading* | 2 | 3 |
| *Read on handled cases* | 1 | 2 |
| *Integrate clinical scenarios* | 1 | 1 |
| *Flashcards* | 0 | 1 |
| *Study CPGs* | 0 | 1 |
| *Rest/sleep well prior to testing* | 2 | 1 |
| *Read questions and answers carefully and thoroughly/identify context clues* | 5 | 2 |
| *Do not overthink* | 1 | 0 |
| *Ask guidance from consultants* | 0 | 1 |
| *None* | 1 | 0 |

**Table 8.** Reasons for Non-execution of the Study Plan

| |
|---|
| *No focus at all* |
| *Time management/time issues/other workloads* |
| *Did not expect the questions to be difficult* |
| *Too tired* |
| *No study plans* |

**Table 9.** Reasons Given by Respondents for Why They Deemed the Exam Wrapper Helpful

| |
|---|
| *Reflection of study habits shortcomings* |
| *Improves self-awareness/form of self-evaluation* |
| *Evaluate test-taking skills* |
| *Reflect/look back to make adjustments, improvements, and learn from mistakes* |
| *Helps identify gaps in knowledge or study techniques* |
| *In depth assessment of misconception in handling the exams* |

The study plans were also assessed according to appropriateness by comparing them with the study habits and error types they enumerated. Four (4) EWs were excluded in this analysis because of erroneous error analysis in one or both EWs. 81.25% of study plans in the EW for Exam 1 and 93.75% for Exam 2 were deemed appropriate (p = 0.30). Perceived knowledge and confidence in adjustments of learning strategies were noted to decrease after the second exam (68.42% vs 47.37%) despite slightly more residents (11 vs 13) included high-impact learning strategies in their second study plan, with spacing topping both EWs.

Eighteen out of 19 respondents (94.74%) reported the EW activity as helpful, while one resident answered "maybe." The different reasons given by residents on why it is helpful are shown in Table 9. The resident who answered "maybe" stated "still not able to study differently" as the reason.

## Metacognition of Residents Using the Metacognitive Awareness Inventory (MAI) before and after a Series of Exam Wrappers Use

The mean MAI scores were 36.79 ± 9.10 for the pre-test and 36.05 ± 9.44 for the post-test, showing no significant difference (p = 0.81) (Table 10). Level I residents had the lowest scores (28.50 ± 7.72; 33.50 ± 14.48), while Level III residents scored the highest (39.83 ± 9.37; 37.50 ± 8.92). Six residents (31.58%) improved their scores, while 10 (52.63%) scored lower. Level I residents had higher post-test scores, whereas Levels II and III had lower scores. In specific MAI domains, the highest scores were in debugging strategies (87.37%) and the lowest in evaluation (59.65%). Post-test scores increased for declarative knowledge and information management, but decreased for procedural knowledge, conditional knowledge, planning, and comprehension monitoring. No significant differences were found in any domains (p <0.05).

## Correlation of Metacognition of Residents and Exam Wrapper Use

Using the rubric for EW (Appendix D), a mean score of 25.88 ± 2.91 was obtained out of a maximum score of 36, maximum of 3 points per category. The residents obtained the lowest mean score when assessed regarding study plan execution (1.53 ± 0.51) and the highest for the type or level of error made during Exam 2 (2.88 ± 0.34), as most of them were recall and related errors. The mean scores of each component of the rubric are shown in Table 11.

Correlating the EW score with the MAI scores using the Pearson correlation coefficient showed weak positive

**Table 10.** Means and Standard Deviations of Residents in the Metacognitive Awareness Index Questionnaire (n=19)

| Variable | Overall, n = 19; Mean (SD) | Level I, n = 4; Mean (SD) | Level II, n = 9; Mean (SD) | Level III, n = 6; Mean (SD) | *p*-value |
|---|---|---|---|---|---|
| *MAI (pre)* | 36.79 (9.10) | 28.50 (7.72) | 38.44 (8.05) | 39.83 (9.37) | 0.81 (using paired t-test) |
| *MAI (post)* | 36.05 (9.44) | 33.50 (14.48) | 36.22 (8.21) | 37.50 (8.92) | |

**Table 11.** Exam Wrapper Scores (n = 19)

| Category | Mean Score (max 3 points each), (SD) |
|---|---|
| ***Self-efficacy Knowledge*** | 2.37 (0.96) |
| ***Time Management*** | |
| Study time | 2.11 (0.94) |
| Study spacing | 1.74 (0.99) |
| ***Learning Habits/Strategies*** | |
| Type of learning strategies | 1.84 (0.90) |
| Knowledge of ways to seek help in studying | 2.53 (0.61) |
| Knowledge of how to adjust one's strategies | 2.32 (0.67) |
| Score Prediction Accuracy | 2.16 (0.76) |
| ***Error Analysis*** | |
| Completeness | 2.47 (0.90) |
| Level of Error | 2.89 (0.32) |
| ***Study Plan*** | |
| Comprehensiveness | 1.68 (0.75) |
| Depth of self-reflection | 2.00 (0.58) |
| Execution | 1.53 (0.51) |

correlation with MAI difference (0.21) but showed weak negative correlation with both pretest (-0.36) and posttest MAI (-0.37). All correlations were not statistically significant.

### Correlation of Metacognition of Residents and Exam Performance

Pearson correlation showed a weak positive correlation (0.35) between pretest MAI and average exam performance, which drops to 0.30 when only Exam 2 and Exam 3 are included. There was no correlation between post-test MAI and mean exam performance, and considering just Exam 2 and Exam 3 resulted in a weak negative correlation (-0.11). Additionally, the MAI difference had a weak negative correlation with mean exam performance (-0.24). The pre-test MAI consistently showed weak positive correlations with all exams (0.31, 0.28, 0.26), while post-test MAI and MAI difference had weak positive correlations with Exam 1 (0.16, 0.05) but weak negative correlations with the others. None of these correlations were statistically significant at $p < 0.05$.

## DISCUSSION

### Validation of Exam Wrapper for Residency Training in Internal Medicine

The items in the pilot EW (Appendix A) were mostly valid (CVR >0.42, I-CVI >0.80), with one item from Part I and two from Part III deemed invalid, and three items in Part I needing revisions (Table 2). All SRL phases had representative items. The EW showed high reliability at 0.65 (KR-20), below the desired 0.70. It should be noted that there were only 19 respondents in the pilot study, one person short of the recommended 20-30 subjects for reliability testing of pilot studies.[36] Revisions were made based on resident feedback and validation, and reliability study results to enhance item clarity and relevance, eventually improving both validity and reliability. It is also important to note that both validity and reliability are properties of the score and not the instrument itself, and can change with the setting.[41]

### Performance of Residents in the Written Examinations and Effects of Exam Wrapper on their Study Habits, Perceived Causes of Exam Mistakes, and Corrective Plans of Residents

The mean exam scores of the IM residents were low, averaging 40-60% (Table 4), with most failing according to exam MPL. This aligns with their rankings in the 40-50% range of their medical school batch. The pre-training backgrounds varied among exam passers, but about 50% attended top-ranking medical schools in terms of PSBIM performance. A medical board rating of 83 and above seemed predictive of good exam results. Studies showed a link between prior academic performance and exam success, regardless of exam wrapper use.[20,24] The improvement in scores from Level I to Level 3 is expected, reflecting greater competency in more senior residents. Exam 2 had significantly lower scores, likely due to its higher difficulty, as indicated by the lowest passing rate.

The exam wrapper revealed poor study habits among residents, who spent at most 10 hours studying for exams (Table 5), despite them being 3-4 weeks apart. Although not significant, there was indeed an increasing trend of study hours, less cramming, and group studying noted between Exam 1 and Exam 2. This could be due to: (1) positive effects of the exam wrapper, (2) residents having more time during their rotation, (3) greater interest in Exam 2 topics, or (4) stronger colleague connections leading to group study. In a similar study among undergraduate biology students, 74.7% reported studying earlier and scheduling more time for subsequent exams.[42] Among the four residents who answered applying experiential learning as a learning strategy (one for Exam 1, three for Exam 2), three of them answered ≤2 hours spent on studying. It might be that they did not count the time spent with their actual patient when they answered this part of the exam wrapper. Additional clarification should be done for future exam wrappers.

The beneficial effect of using more high-impact learning strategies was not elucidated in this study, as the passing rate and grade improvement from Exam 1 to Exam 2 showed little difference. Among the few who passed Exam 2 (4 out of 19), one did not use any high-impact strategies. Factors such as the exam's difficulty, topic interest, and study time may have had a greater impact than these strategies. Additionally, some respondents may be unaware of these strategies. A baseline knowledge assessment and workshops could help enhance their effectiveness.

The self-efficacy for subspecialty examination of included residents was very low, only 15.79% for both exams, and no resident claimed to be prepared for both exams. This low self-efficacy likely stems from limited study time, averaging less

than 10 hours (Table 4). Consistent with a learner's previous experiences being one of the strongest cues for self-efficacy, the residents who claimed to be prepared for the exam indeed passed or almost passed their previous exam of the same subspecialty, except for one whose previous exam data was not available.[43] Repeated failures can further erode self-efficacy and motivation to study. Therefore, mentoring sessions could help improve their outlook and address issues affecting their grades not covered by the evaluation.

Self-efficacy accuracy (68.42% vs 63.16%) and grade estimation (26.32% vs 31.58%) did not significantly differ between EWs (p = 0.74, 0.49). The difficulty level of Exam 2 may have affected outcomes, as all residents who felt prepared failed. Conversely, 4 out of 19 who felt unprepared in Exam 2 passed, indicating that exam difficulty may have caused self-doubt. This aligns with the Dunning-Kruger effect, where those who got high grades underestimate themselves, and the opposite for those with low grades.[44] This reflects poorly on the quality of self-reflection of these residents. To improve self-reflection skills, suggestions include: (1) frequent multi-source feedback; (2) stern mentoring for overconfident, poor performers; and (3) supportive mentoring for insecure but competent residents to encourage improvement without giving superficial reassurance.[45]

As expected from the low study hours, the most common cause of error was attributed to not being able to read what is being asked for, although there was a decreasing ranking trend from Exam 1 (1.63 ± 0.96) to Exam 2 (1.94 ± 1.00) (Figure 4). The slightly higher number of residents doing worse in error analysis accuracy after Exam 2, compared to those who improved (3 vs 2), further proved the comments made during the pilot study on how this part of the EW seems confusing. It might be possible that some did not answer this part properly or seriously. Further improvement of this part should be considered. It also showed again how these residents might find it toiling to reflect on past actions.

The relatively low help-seeking knowledge (52.63% and 57.89%) for both EWs indicates a disconnect with the training team. Hospital 1 holds regular quarterly feedback between the training team and the residents, yet 6 of them (40%) answered that they do not know how to seek help. Are the questions being asked during said feedback about their exams lacking or not specific enough? Hospital 2, on the other hand, did not hold regular monthly or quarterly feedback, but they do give feedback after an exam. It is, however, not known if this feedback includes provision of strategies to the trainees for their next exam. It is good to note, however, that more residents mentioned ways to help them compared to the first EW (12 vs 15) (Table 6), which might suggest improvement in self-reflection skills and forethought skills.

Given that the most common error was not reading about the topic, it is not surprising that the most common strategy they wrote is to "read more" (Table 7). It is also not surprising that their top reason for non-execution of their study plan is lack of time (Table 8). During the study period, all but one resident had conference or lecture reporting assignments, seven had quiz competitions, and all had areas to do rounds on and must go on 24-hour duty at least every three days. Time issues have been a recurring top problem identified in several exam wrapper studies.[6,7,42] One study noticed that the undergraduate students who had insufficient monitoring of their learning strategies tended to blame a lack of time studying as the main reason for failing an exam.[42] It was also noted to be the main consideration in the allied health undergraduate students' choice of learning strategy.[7] Study time indeed is a tricky problem to resolve, as you need to balance letting the trainees have more clinical experiential learning, but you must also allot time for them to study their textbooks, notes, or other references, and this time must be when they are not yet too exhausted to study. Another thing that might have worsened the "lack of time" in these two institutions is the lack of applicants. Hospital 1 can train a maximum of seven residents per year, and Hospital 2 can train six, totaling 39 spots. However, only 24 residents participated at the study's start despite everyone agreeing to participate. This could have potentially increased individual workloads as residents took on more duties and had quicker rotations for reporting tasks. Hospital 2 hired hospitalists to help, but without data on patient load, the effectiveness of this measure is unclear. Additionally, some residents participated in quiz competitions, which might have added to their burden if the topics differed from their exams. The presence of a protected study time was also not part of the data gathered by this study. Whether the personal opinions or reasons listed here are indeed significant factors in the study time issues of residents warrants further validation.

The appropriateness of the study plan reflects the quality of self-reflection and forethought, with over 80% of residents providing suitable plans for errors or problems. This is consistent with their high scores in debugging strategies on the MAI, but contrasts with their lower evaluation scores. Many errors stemmed from not knowing the topic, making it easy to suggest "reading more" in their plans. A deeper analysis using the EW rubric revealed a mean score of 1.68 ± 0.75 for comprehensiveness and 2.00 ± 0.58 for self-reflection depth. This suggests a lack of thorough self-evaluation or seriousness in their responses. While some included effective learning strategies, execution was often lacking. Like the findings among undergraduate biology students, increasing study time was manageable, but adopting new strategies proved challenging.[42]

Interestingly, the perceived knowledge about learning strategies and how to adjust them decreased on their next EW. Given that Exam 2 was significantly different from the other two and deemed more difficult, this might have lowered their morale. It might also be that they were more honest in answering this part than during their first EW. There might also be failure from the coordinators to fully teach them how to adjust their learning strategies after Exam 1. Despite the possibility of the latter, slightly more residents (11 vs 13)

employed more high-impact strategies in their second study plan, most being about spacing as a direct answer to their main issue about study time.

Ninety-four percent (94.74%) of residents perceived the EW as helpful, mainly because it helped them reflect and make improvements in their study habits and test-taking skills (Table 9). Such a perception did not differ regardless of the exam performance of the residents, unlike those seen in undergraduates.[3,30] This might reflect the maturity of post-graduate trainees and possibly less importance given by residents to academic grades versus actual clinical experience.

### Metacognition of Residents Using the Metacognitive Awareness Inventory (MAI) before and after a Series of Exam Wrappers Use

Results indicated a trend toward a lower mean metacognition score after taking 2 EWs, though not significantly different (Table 10). This trend was also observed among first-year allied health students, but this study found no significant differences in specific domains.[7] A possible explanation is the smaller sample size of 19 respondents in this study compared to 52 in the other study. Additionally, scores on the pre-test MAI may have been exaggerated, as noted in research on undergraduate STEM students.[24] Among residency levels, Level I showed higher post-test MAI scores compared to Levels II and III. Factors influencing this may include honesty in MAI responses, effects of clinical rotations and training, the impact of exam performance, self-pressure among Level III residents, and the maturation of self-reflection with experience.

### Correlation of Metacognition of Residents and Exam Wrapper Use

The result of a nonsignificant weak negative correlation in this study between the EW score and the MAI (pre) and MAI (post) scores might be evidence of exaggeration of answers. Recall that an exam wrapper is itself a structured metacognitive and self-regulated learning strategy.[3,4] It is thus expected that a rubric designed to assess the quality of self-regulation will end up having a positive correlation with a metacognition assessment tool, if there will indeed be any correlation. The frequency and interval of exam wrapper administration might also be a factor, as there might be a very short period for any significant positive change in metacognition to occur during the study period.

### Correlation of Metacognition of Residents and Exam Performance

Results indicated weak positive correlations between pre-test MAI and exam performance, which were not statistically significant. This aligns with previous studies suggesting that better metacognition leads to better academic performance.[15-24] In contrast, existing studies on exam wrappers (EWs) show positive correlations with post-test MAI scores, not pre-test MAI.[23,24] In our study, the correlation between exam performance and post-test MAI varied, showing weak negative but nonsignificant results for Exams 2 and 3, which followed EW use. Exam 2 had the lowest mean grade and the highest failure rate among residents, while Exam 1 had the highest mean grade (Table 4). This suggests that despite lower exam scores, residents may have achieved unexpectedly higher post-test MAI scores with EWs. Additionally, those with higher scores in Exams 2 and 3 were mostly still considered to have failed, likely influencing their post-test MAI responses. Similar to findings in undergraduate studies, where high performers report frequent metacognitive strategy use, the participants in this study may have experienced the opposite effect.[24] This may also explain the negative correlation between MAI differences and exam performance. But as mentioned earlier, this inconsistent correlation might also be due to the exaggeration of MAI scores by our participants.

## CONCLUSION

An EW with acceptable validity indices (CVR 0.72, I-CVI 0.86) and reliability coefficient (KR-20 0.65) was created after some revisions. Both institutions have the majority of their residents performing poorly during exams, who were also found to cram their study time, to prefer the usage of low-impact learning strategies, and have poor self-reflection skills and poor metacognition monitoring. Time issues related to reading or studying were the top problem identified, and the target of most study plans enumerated by the residents. There is no significant relationship between metacognition score and exam wrapper use or exam performance in IM residency trainees. Rather than a tool to improve exam performance or metacognition, an exam wrapper seems to serve more as a self-reflection practice tool and a tool that can help the training committee improve on their teaching-learning activities, including the scheduling of these activities, to fit the type of learners they have and to address the issue of "study time." It is thus a good tool for training institutions to adopt as a means of self-reflection practice for their residents and to open channels for feedback about their usual teaching-learning activities.

### Recommendations

Time management, especially in relation to studying, appears to be a major perceived reason for poor exam performance according to this study. Training institutions might need to invest in workshops or mentoring strategies to train their trainees about this skill. Training institutions might also need to set an actual study period to ensure that studying occurs during the training period. Given the residents' poor self-reflection skills, exercises like EW could be advantageous, but a protected time for self-reflection is also suggested to encourage honesty in answering. Although the study did not demonstrate a significant correlation between EW and metacognition scores, further research

exploring specific interventions within the EW framework may also yield promising results. This includes improvements in how performance feedback of residents is done, emphasis given on the importance of exam performance, encouraging help-seeking behaviors, gathering suggestions for exam improvements, and considering workshops on effective learning strategies. EW revisions may be needed regarding item definitions, the inclusion of new learning strategies like AI, and simplifying error analysis. Further investigation is needed into the negative correlation between quality EW use and MAI scores. Exploration on the timing of the exam wrapper should also be sought to give enough time for these residents to imbibe the results of their self-reflection and to learn a new learning strategy. It will also be good to explore if it is indeed time issues, poor metacognition monitoring, or exhaustion that is playing a more significant role in study motivation.

## Acknowledgments

The authors would like to thank the invaluable help of the study's hospital coordinators, Drs. Marian Joy Lopez and John Kenneth Molina, for recruiting and reminding the enrolled residents to answer the questionnaires and to submit them on time. They were also the ones who transmitted the coded data sheet of baseline information for the study. The authors would also like to thank Drs. Marian Joy Lopez, Daniel Guevarra, and John Paulo Vergara for creating the examinations used for both training institutions during the duration of this study, and for transmitting the examination grades.

## Statement of Authorship

Both authors certified fulfillment of ICMJE authorship criteria.

## Author Disclosure

Both authors declared no conflicts of interest.

## Funding Source

The study was funded by the authors.

# REFERENCES

1. Philippine College of Physicians. Handbook on Outcome Based Education-Competency Based Traiing Program (OBE-CBTP) in Internal Medicine and Accreditation Process [Unpublished]. 2021. pp. 8-38.
2. Kannan KK, Muthammal R. Exam wrapper and metacognition for undergraduate surgery students in exam preparation. Int Surg J. 2020;7(12):4083. doi: 10.18203/2349-2902.isj20205361.
3. Gezer-Templeton PG, Mayhew EJ, Korte DS, Schmidt SJ. Use of exam wrappers to enhance students' metacognitive skills in a large introductory food science and human nutrition course. J Food Sci Educ. 2017 Jan;16(1):28-36. doi: 10.1111/1541-4329.12103.
4. Pate A, Lafitte EM, Ramachandran S, Caldwell DJ. The use of exam wrappers to promote metacognition. Curr Pharm Teach Learn. 2019 May 1;11(5):492-8. doi: 10.1016/j.cptl.2019.02.008. PMID: 31171251.
5. Achacoso MV. Post-test analysis: A tool for developing students' metacognitive awareness and self-regulation. New Dir Teach Learn. 2004 Dec 1;2004(100):115-9. doi: 10.1002/tl.179.
6. Butzlaff A, Gaylle D, Kelley CO. Student self-evaluation after nursing examinations: That'sa wrap. Nurse Educ. 2018 Jul 1;43(4):187-90. doi: 10.1097/NNE.0000000000000534. PMID: 29652698.
7. Cale AS, Hoffman LA, McNulty MA. Promoting metacognition in an allied health anatomy course. Anat Sci Educ. 2023 May;16(3):473-85. doi: 10.1002/ase.2218. PMID: 35951462.
8. Sethares KA, Asselin ME, Mahoney D, Nicotera J, Chung J, Schuler M. Description and comparison of exam wrapper learning strategy use in baccalaureate and associate degree nursing students: A descriptive study. Nurse Educ Today. 2021 Aug 1;103:104961. doi: 10.1016/j.nedt.2021.104961. PMID: 34049121.
9. Vemu S, Denaro K, Sato BK, Fisher MR, Williams AE. Moving the needle: Evidence of an effective study strategy intervention in a community college biology course. CBE Life Sci Educ. 2022;21(2):ar24. doi: 10.1187/cbe.21-08-0216. PMID: 36112626: PMCID: PMC9727618.
10. Zimmerman BJ. Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. Am Educ Res J. 2008 Mar;45(1):166-83. doi: 10.3102/0002831207312909.
11. Livingston JA, Metacognition: An overview [Internet]. 2003 [cited 2022 Dec]. Available from: https://eric.ed.gov/?id=ED474273
12. Siegesmund A. Increasing student metacognition and learning through classroom-based learning communities and self-assessment. JMBE. 2016 May;17(2):204-14. doi: 10.1128/jmbe.v17i2.954. PMID: 27158301: PMCID: PMC4858356.
13. Medina MS, Castleberry AN, Persky AM. Strategies for improving learner metacognition in health professional education. Am J Pharm Educ. 2017 May 1;81(4):78. doi: 10.5688/ajpe81478. PMID: 28630519: PMCID: PMC5468716.
14. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. J Pers Soc Psychol. 1999 Dec;77(6):1121. PMID: 10626367.
15. Coutinho SA. The relationship between goals, metacognition, and academic success. Educate. 2007 Jun 30;7(1):39-47.
16. Aurah CM. The effects of self-efficacy beliefs and metacognition on academic performance: A mixed method study. Am J Educ Res. 2013 Jan;1(8):334-43. doi: 10.12691/education-1-8-11.
17. de Boer H, Donker AS, Kostons DD, Van der Werf GP. Long-term effects of metacognitive strategy instruction on student academic performance: a meta-analysis. Educ Res Rev. 2018 Jun 1;24:98-115. doi: 10.1016/j.edurev.2018.03.002.
18. Contessa J, Kyriakides TC, Nadzam G. Does assessment of self-regulated learning and metacognition in surgical residents provide insight to performance on high stakes standardized examinations? A pilot study. Clin Med Rev Case Rep. 2015;2(8):47. doi: 10.23937/2378-3656/1410047.
19. Ohtani K, Hisasaka T. Beyond intelligence: A meta-analytic review of the relationship among metacognition, intelligence, and academic performance. Metacognition Learn. 2018 Aug;13:179-212. doi: 10.1007/s11409-018-9183-8.
20. Plants RT. The relationship of motivation and metacognition to academic performance in graduate medical education [Dissertation]. Diss Abstr Int. 2000;61(3-A):884.
21. Soicher RN, Gurung RA. Do exam wrappers increase metacognition and performance? A single course intervention. PLAT. 2017 Mar;16(1):64-73. doi: 10.1177/1475725716661872.
22. Zulkiply N. Metacognition and its relationship with students' academic performance. Int J Learn. 2009;15(11):97-106.
23. Carpenter TS, Beall LC, Hodges LC. Using the LMS for exam wrapper feedback to prompt metacognitive awareness in large courses. JoTLT [Internet]. 2020 Aug 12 [cited 2022 Dec 6];9(1). Available from: https://scholarworks.iu.edu/journals/index.php/jotlt/article/view/29156
24. Hodges LC, Beall LC, Anderson EC, Carpenter TS, Cui L, Feeser E, et al. Effect of exam wrappers on student achievement in multiple,

large STEM courses. J Coll Sci Teach. 2020 Sep 1;50(1):69-79. doi: 10.1080/0047231X.2020.12290677.
25. Schuler MS, Chung J. Exam wrapper use and metacognition in a fundamentals course: Perceptions and reality. J Nurs Educ. 2019 Jul 1;58(7):417-21. doi: 10.3928/01484834-20190614-06. PMID: 31242311.
26. Kellenberg F, Schmidt J, Werner C. The adult learner: Self-determined, self-regulated, and reflective. Signum Tempor. 2017;9(1):23. doi: 10.1515/sigtem-2017-0001.
27. Burman NJ, Boscardin CK, Van Schaik SM. Career-long learning: Relationship between cognitive and metacognitive skills. Med Teach. 2014 Aug;36(8):715-23. doi: 10.3109/0142159X.2014.909010. PMID: 24796358.
28. Sana EA. Evaluation in Health Science Education. In: Sana EA. Teaching and Learning in the Health Sciences. Quezon City: University of the Philippines Press; 2013. pp. 182-202.
29. Badir A, O'Neill R, Liao J, Papkov GI. Exam Wrappers, Reflection, and Student Performance in Engineering Mechanics–Part II. In: 2019 ASEE Annual Conference & Exposition [Internet]. 2019 Jun 15 [cited 2022 Dec 6]. Available from: https://www.academia.edu/68825868/Exam_Wrappers_Reflection_and_Student_Performance_in_Engineering_Mechanics_Part_II
30. Edlund JE. Exam wrappers in psychology. Teach Pscyhol. 2020 Apr;47(2):156-61. doi: 10.1177/0098628320901385.
31. Liao J, Kunberger T, Papkov GI, Badir A, O'Neill R, Nguyen LD. Exam Wrappers, Reflection, and Student Performance in Engineering Mechanics. In: 2018 ASEE Annual Conference & Exposition [Internet]. 2018 Jun 23 [cited 2022 Dec 6]. Available from: https://peer.asee.org/exam-wrappers-reflection-and-student-performance-in-engineering-mechanics
32. Nikolopoulou K. What is content validity? Definition & examples [Internet]. 2023 [cited 2024 Feb]. Available from: https://www.scribbr.com/methodology/contentvalidity/
33. Zeraati M, Masoudi Alavi N. Designing and validity evaluation of Quality of Nursing Care Scale in Intensive Care Units. J Nurs Meas. 2014 Dec 1;22(3). doi: 10.1891/1061-3749.22.3.461. PMID 25608432.
34. Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi-Majd H, Nikanfar A-R. Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. J Caring Sci. 2015 Jun 1;4(2):165-78. doi: 10.15171/jcs.2015.017. PMID: 26161370. PMCID: PMC4484991.
35. Bolarinwa OA. Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. Niger Postgrad Med J. 2015 Oct 1;22(4):195-201. doi: 10.4103/1117-1936.173959. PMID: 26776330.
36. Price PC, Jhangiani R, Chiang ICA. Reliability and validity of measurement. In: Research methods in psychology, 2nd ed [Internet]. 2015 [cited 2024 Feb]. Available from: https://opentextbc.ca/researchmethods/chapter/reliability-and-validity-of-measurement/
37. Liliana RA, Raharjo W, Jauhari I, Sulisworo D. Effects of the online interactive learning media on student's achievement and interest in physics. Univers J Educ Res. 2020 Mar;8(3):59-68. doi: 10.13189/ujer.2020.081507.
38. Schraw G, Dennison RS. Assessing metacognitive awareness. Contemp Educ Psychol. 1994 Oct 1;19(4):460-75. doi: 10.1006/ceps.1994.1033.
39. National Privacy Commission Republic Act 10173 – Data Privacy Act of 2012 [Internet]. 2012 [cited 2024 Feb]. Available from: https://www.privacy.gov.ph/data-privacy-act/
40. Philippine Health Research Ethics Board, National Ethical Guidelines for Research Involving Human Participants [Internet]. 2022 May [cited 2022 Dec 6]. Available from: https://www.pchrd.dost.gov.ph/publications/2022-national-ethical-guidelines/
41. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. Am J Med. 2006 Feb 1;119(2):166.e7-16. doi: 10.1016/j.amjmed.2005.10.036. PMID: 16443422.
42. Stanton JD, Neider XN, Gallegos IJ, Clark NC. Differences in metacognitive regulation in introductory biology students: when prompts are not enough. CBE Life Sci Educ. 2015 Jun 1;14(2):ar15. doi: 10.1187/cbe.14-08-0135. PMID: 25976651: PMCID: PMC4477731.
43. Cook DA, Artino Jr AR. Motivation to learn: an overview of contemporary theories. Med Educ. 2016 Oct;50(10):997-1014. doi: 10.1111/medu.13074. PMID: 27628718: PMCID: PMC5113774.
44. Dunning D, Heath C, Suls JM. Flawed self-assessment: Implications for health, education, and the workplace. Psychol Sci Public Interest. 2004 Dec;5(3):69-106. doi: 10.1111/j.1529-1006.2004.00018. PMID: 26158995.
45. Rahmani M. Medical trainees and the Dunning–Kruger effect: when they don't know what they don't know. J Grad Med Educ. 2020 Oct;12(5):532-4. doi: 10.4300/JGME-D-20-00134.1. PMID: 33149817: PMCID: PMC7594774.

# APPENDICES

**Appendix A.** Pilot Study Exam Wrapper and Evaluation Form

**A. Pilot Study Exam Wrapper**

**Part I:**

(1) Did you feel that you were prepared for the exam? Yes ___ No ___ (Carpenter, Beall, & Hodges, 2020)
(2) Approximately how much time did you spend preparing for this exam? Please give a best estimate in hours, do not use ranges (e.g., 3.5 not 3-4). _____ (Achacoso, 2004) (Badir, O'Neill, Liao, & Papkov, 2019) (Butzlaff, Gaylle, & O'Leary Kelley, 2018) (Edlund, 2020) (Pate, Lafitte, Ramachandran, & Caldwell, 2019) (Soicher & Gurung, 2017)
(3) Which of the following best describes your study patterns for this exam? (Badir, O'Neill, Liao, & Papkov, 2019) (Pate, Lafitte, Ramachandran, & Caldwell, 2019) (Vemu, Denaro, Sato, & Williams, 2022)
___ (a) I spaced out my study session over multiple days/weeks
___ (b) I did most of my studying right before the test
(4) What percentage of your exam preparation was spend studying alone or with others? (Badir, O'Neill, Liao, & Papkov, 2019) (Butzlaff, Gaylle, & O'Leary Kelley, 2018) (Sethares, et al., 2021)
(a) % alone ___ (b) % with others ___
(5) Which of the following did you do to prepare for the exam? (Check all that apply; Mark the ones that you did differently for this exam with an asterisk "*") (Achacoso, 2004) (Badir, O'Neill, Liao, & Papkov, 2019) (Butzlaff, Gaylle, & O'Leary Kelley, 2018) (Edlund, 2020) (Pate, Lafitte, Ramachandran, & Caldwell, 2019) (Sethares, et al., 2021) (Vemu, Denaro, Sato, & Williams, 2022) (Soicher & Gurung, 2017)
___ (a) Skimming the textbook
___ (b) Reading the textbook thoroughly for the first time
___ (c) Re-reading the textbook
___ (d) Reviewing notes/summarized references (includes notes during rounds)
___ (e) Reviewing other media types; Specify what media: _____
___ (f) Self-quizzing/testing
___ (g) Study actual patient case
___ (h) Reading medical journals/guidelines
___ (i) Drawing concept maps or algorithms
___ (j) Others (please specify): _____
(6) How much of your study plan from the last exam wrapper were you able to carry out in this exam? ___%
(7) Do you think those study plans improved your study habits for this exam?
Yes ___ No ___ Did not carry it out ___ Did not take exam wrapper 1 ___
(8) If you were unable to carry out all or most of your study plan from the previous exam wrapper, what were the reasons?
(9) How well do you think you did in this exam? (Give an estimate of your possible score) (Achacoso, 2004)

**Part II:**

(1) How did your actual score compare to how you thought you did on the exam after taking it? (Achacoso, 2004) (Soicher & Gurung, 2017) ___ Higher ___ About the same ___ Lower
(2) How did this exam score compare to your last exam score? (Badir, O'Neill, Liao, & Papkov, 2019) (Soicher & Gurung, 2017)
___ Higher ___ About the same ___ Lower
(3) As you look over your graded exam, analyze where/how you lost points. Fill in the blanks with the number of points you lost due to each of the following: (Badir, O'Neill, Liao, & Papkov, 2019) (Butzlaff, Gaylle, & O'Leary Kelley, 2018) (Carpenter, Beall, & Hodges, 2020) (Edlund, 2020) (Pate, Lafitte, Ramachandran, & Caldwell, 2019) (Sethares, et al., 2021) (Soicher & Gurung, 2017)
___ (a) Lack of understanding of the concept
___ (b) Not understanding what the question was asking
___ (c) Careless mistakes
___ (d) Not being able to apply concepts in new contexts
___ (e) Not recognizing that information or ideas were important
___ (f) Difficulty choosing between 2 alternatives
___ (g) Inability to remember the information
___ (h) Have not read about it
___ (i) Others (please specify): _____
(4) Do you know all the ways that you can get help in studying? Yes ___ No ___ (Carpenter, Beall, & Hodges, 2020)
(5) Do you feel that you have made a good assessment of your learning habits and know how to adjust your approach?
Yes ___ No ___ (Carpenter, Beall, & Hodges, 2020)

**Part III:**

(1) Based on your responses to the questions above, name at least three things you plan to do differently in preparing for the next exam. (Be as specific as possible) (Badir, O'Neill, Liao, & Papkov, 2019) (Butzlaff, Gaylle, & O'Leary Kelley, 2018) (Edlund, 2020) (Pate, Lafitte, Ramachandran, & Caldwell, 2019) (Sethares, et al., 2021) (Soicher & Gurung, 2017)
(2) What else can we, your consultants, do to help support your learning and preparation for the next exam? (Achacoso, 2004) (Badir, O'Neill, Liao, & Papkov, 2019) (Butzlaff, Gaylle, & O'Leary Kelley, 2018) (Edlund, 2020)
(3) Do you find this exam wrapper activity useful? Why or why not?

**B. Pilot study evaluation form for the domains of the self-regulation learning strategy**

**Instruction:** Please check the appropriate column (Yes or No) if the components under the specific phases of the Self-regulated Learning Theory were elucidated in the different parts of the exam wrapper. Said process may not have been directly asked but is a vital process for one to be able to come up with an answer to the question.

| Part I/II/III | Yes | No | If Yes, write the item # |
|---|---|---|---|
| **Forethought Phase** | | | |
| Task analysis: fragmenting of the task and establishing the strategies | | | |
| • Goal setting (taking into account standards and intended level of "perfection") | | | |
| • Strategic planning (selecting an action plan and choosing the strategies that are needed) | | | |
| Self-motivation beliefs: personal variables that generate and maintain the motivation to perform the task | | | |
| • Self-efficacy (learner's belief about their capability to perform the task) | | | |
| • Outcome expectations (beliefs about the probability to success in the task) | | | |
| • Task value (relevance of the task for personal goals) | | | |
| • Interest (liking for the task) | | | |
| • Goal orientation (beliefs about one's learning purposes) | | | |
| **Performance Phase** | | | |
| Self-observation: process of comparing what one is doing against the expert model | | | |
| • Metacognitive monitoring (cognitive process of assessing one's performance) | | | |
| • Self-recording (keeping a record of the actions for a latter analysis) | | | |
| Self-control: process of maintaining the concentration and interest through strategies | | | |
| • Task strategies (use of specific tactics related to the task) | | | |
| • Self-instruction (self-given instructions about the task) | | | |
| • Imagery (mental organization of the information) | | | |
| • Time management (planning the use of time during a task) | | | |
| • Environmental structuring (creating an environment that facilitates learning) | | | |
| • Help-seeking (asking for help when needed) | | | |
| • Interest incentives (self-given messages to remind the goals) | | | |
| • Self-consequences (enhancing the feeling of progress through self-praise and self-rewards) | | | |
| **Self-reflection Phase** | | | |
| Self-judgment: process in which one assess one's work | | | |
| • Self-evaluation (assessment of performance based on one's assessment criteria and modulated by one's performance level goal) | | | |
| • Causal attribution (self-explanations about the reasons for success or failure) | | | |
| Self-reaction: reactions to self-judgments | | | |
| • Self-satisfaction/affect (affective and cognitive reactions produced by self-judgments) | | | |
| • Adaptive/defensive decisions (will to perform the task in the future and to activate learning strategies) | | | |

**Appendix B.** Exam Wrapper

**Part I:**
Code: _____

(1) Did you feel that you were prepared for the exam? Yes ___ No ___
(2) Approximately how much time did you spend preparing for this exam? Please give a best estimate in hours, do not use ranges (e.g., 3.5 not 3-4). _____
(3) How did you divide the number of hours you studied? Over multiple days/weeks ___ All right before the test ___
(4) Which of the following did you do to prepare for the exam? (Check all that apply; Mark the ones that you did differently for this exam with an asterisk "*")
___ Skimming the textbook
___ Reading the textbook thoroughly for the first time
___ Re-reading the textbook
___ Reviewing notes/summarized references (includes notes during rounds)
___ Reviewing other media types; Specify what media: _____
___ Self-quizzing/testing
___ Study actual patient case
___ Reading medical journals/guidelines
___ Drawing concept maps or algorithms
___ Studying with others/group study
___ Others (please specify): _____
(5) How much of your study plan from the last exam wrapper were you able to carry out in this exam? Write "0" if you did not implement it or did not take an exam wrapper last time. ___%
(6) If you implemented your study plan, do you think your study habits improved for this exam?
Yes ___ No ___ Did not carry it out ___
(7) If you were unable to carry out all or most of your study plan from the previous exam wrapper, what were the reasons?
(8) How well do you think you did in this exam? (Give an estimate of your possible score)

**Part II:**
Code: _____

(1) How did your actual score compare to how you thought you did on the exam after taking it?
___ Higher
___ About the same
___ Lower
(2) How did this exam score compare to your last exam score?
___ Higher
___ About the same
___ Lower
(3) As you look over your graded exam, analyze where/how you lost points. Fill in the blanks with the number of points you lost due to each of the following:
___ (a) Lack of understanding of the concept
___ (b) Not understanding what the question was asking
___ (c) Careless mistakes
___ (d) Not being able to apply concepts in new contexts
___ (e) Not recognizing that information or ideas were important
___ (f) Difficulty choosing between 2 alternatives
___ (g) Inability to remember the information
___ (h) Have not read about it
___ (i) Others (please specify): _____
(4) Do you know all the ways that you can get help in studying? Yes ___ No ___
(5) Do you feel that you have made a good assessment of your learning habits and know how to adjust your approach?
Yes ___ No ___

**Part III:**
(1) Based on your responses to the questions above, name at least three things you plan to do differently in preparing for the next exam. (Be as specific as possible)
(2) What else can we, your consultants, do to help support your learning and preparation for the next exam?
(3) Do you find this exam wrapper activity useful? Why or why not?

**Appendix C.** Metacognitive Awareness Inventory *(Modified True/False Format)*

**This questionnaire is available for free. The original version was created by Schraw and Dennison (Schraw & Dennison, 1994) but was later modified to the true-false format by several authors for lesser survey fatigue (see Review of Related Literature for discussion). It is further modified here to include a short description of its constructs and instructions on how to answer it.*

**Description:** This is a 52-item questionnaire to assess your metacognition awareness about learning, divided into 2 domains: knowledge and regulation. Metacognitive knowledge is knowing what you know, including knowledge of your skills, abilities, interests, and preferences, on how you address your learning tasks, on what learning strategies are applicable for a certain condition, and why that is so. Metacognitive regulation pertains to the cognitive process during the learning process, from the planning stage, organizing the information you know and how you tap them, monitoring your learning or monitoring the precision of the learning strategy that you chose to use, and evaluating your performance after a learning activity and the effectiveness of the learning strategy used.

**Instruction:** Read the following statements and put a (√) under the true column if you exercise said activity or under the false column if you do not practice it.

| | True | False |
|---|---|---|
| 1. I ask myself periodically if I am meeting my goals | | |
| 2. I consider several alternatives to a problem before I answer | | |
| 3. I try to use strategies that have worked in the past | | |
| 4. I pace myself while learning in order to have enough time | | |
| 5. I understand my intellectual strengths and weaknesses | | |
| 6. I think about what I really need to learn before I begin a task | | |
| 7. I know how well I did once I finish a test | | |
| 8. I set specific goals before I begin a task | | |
| 9. I slow down when I encounter important information | | |
| 10. I know what kind of information is most important to learn | | |
| 11. I ask myself if I have considered all options when solving a problem | | |
| 12. I am good at organizing information | | |
| 13. I consciously focus my attention on important information | | |
| 14. I have a specific purpose for each strategy I use | | |
| 15. I learn best when I know something about the topic | | |
| 16. I know what an attending expects me to learn | | |
| 17. I am good at remembering information | | |
| 18. I use different learning strategies depending on the situation | | |
| 19. I ask myself if there was an easier way to do things after I finish a task | | |
| 20. I have control over how well I learn | | |
| 21. I periodically review to help me understand important relationships | | |
| 22. I ask myself about the material before I begin | | |
| 23. I think of several ways to solve a problem and choose the best one | | |
| 24. I summarize what I've learned after I finish | | |
| 25. I ask others for help when I don't understand something | | |
| 26. I can motivate myself to learn when I need to | | |
| 27. I am aware of what strategies I use when I study | | |
| 28. I find myself analyzing the usefulness of strategies while I study | | |
| 29. I use my intellectual strengths to compensate for my weaknesses | | |
| 30. I focus on the meaning and significance of new information | | |
| 31. I create my own examples to make information more meaningful | | |
| 32. I am a good judge of how well I understand something | | |
| 33. I find myself using helpful learning strategies automatically | | |
| 34. I find myself pausing regularly to check my comprehension | | |
| 35. I know when each strategy I use will be most effective | | |

| | True | False |
|---|---|---|
| 36. I ask myself how well I accomplish my goals once I'm finished | | |
| 37. I draw pictures or diagrams to help me understand while learning | | |
| 38. I ask myself if I have considered all options after I solve a problem | | |
| 39. I try to translate new information into my own words | | |
| 40. I change strategies when I fail to understand | | |
| 41. I used the organizational structure of the text to help me learn | | |
| 42. I read instructions carefully before I begin a task | | |
| 43. I ask myself if what I'm reading is related to what I already know | | |
| 44. I reevaluate my assumptions when I get confused | | |
| 45. I organize my time to best accomplish my goals | | |
| 46. I learn more when I am interested in the topic | | |
| 47. I try to break studying down into smaller steps | | |
| 48. I focus on overall meaning rather than specifics | | |
| 49. I ask myself questions about how well I am doing while I am learning something new | | |
| 50. I ask myself if I learned as much as I could have once I finish a task | | |
| 51. I stop and go back over new information that is not clear | | |
| 52. I stop and reread when I get confused | | |

**Questions per domains:**
Declarative knowledge – #5, 10, 12, 16, 17, 20, 32, 46
Procedural knowledge – #3, 14, 27, 33
Conditional knowledge – #15, 18, 26, 29, 35
Planning – #4, 6, 8, 22, 23, 42, 45
Information management strategies – #9, 13, 30, 31, 37, 39, 41, 43, 47, 48
Comprehension monitoring – #1, 2, 11, 21, 28, 34, 49
Debugging strategies – #25, 40, 44, 51, 52
Evaluation – #7, 19, 24, 36, 38, 50

**Appendix D.** Rubric for Second Exam Wrapper Use

| | 1 | 2 | 3 |
|---|---|---|---|
| ***Knowledge of one's ability to take the exam/ self-efficacy (Part 1 Q1)*** | Feels prepared but obtained a failing score or not prepared but obtained a passing score | | Feels prepared and obtained a passing score or did not prepare and obtained a failing score |
| **Time Management** | | | |
| ***Study time (Part 1 Q2)*** | Decrease in study hours | Maintained previous study hours | Increase in study hours |
| ***Timing of study period in relation to examination (Part 1 Q3)*** | Right before the test | | Spaced out over multiple days/weeks |
| **Learning Habits/Strategies** | | | |
| ***Type of learning strategies (Part 1 Q4)*** | Mostly low impact strategies | | Mostly high impact strategies |
| ***Knowledge of ways to seek help in studying (Part 2 Q4, Part 3 Q2)*** | Not knowledgeable but not willing to ask for help | Not knowledgeable but willing to seek help | Knowledgeable |
| ***Knowledge of how to adjust one's strategies (Part 2 Q5, Part 3 Q1 & 2)*** | Not knowledgeable but not willing to ask for help | Not knowledgeable but willing to seek help | Knowledgeable and suggested ways to correct it |
| ***Score Prediction Accuracy (Part 1 Q6)*** | >50% difference | 10-50% difference | Same (<10% difference) |
| **Error Analysis (Part 2 Q3)** | | | |
| ***Completeness*** | Incomplete | | Complete |
| ***Level of Error*** | Mostly due to application errors or unable to decide between two choices | Mostly due to comprehension errors | Mostly due to recall errors, carelessness, not reading about it, or not recognizing important information |
| **Study Plan** | | | |
| ***Comprehensiveness (Part 3 Q1)*** | Too generalized or vague | | Specific and comprehensive |
| ***Depth of self-reflection (Part 3 Q1)*** | Fail to internalize the reasons of his/her mistakes and identified inappropriate steps to correct them | Has good grasp of the reasons of his/her mistakes and identified means to mitigate them but are not the best solutions | Fully grasped the reason of his/her mistakes and appropriately identified means to mitigate them |
| ***Execution (Part 1 Q5)*** | Did not execute it | Partially executed it | Able to fully execute it |