How does quiz activity affect summative assessment outcomes? An analysis of three consecutive years’ data on self-directed learning

Chi Eun Oh¹, Hyunyong Hwang²

¹Department of Pediatrics, Kosin University College of Medicine, Busan, Korea
²Department of Laboratory Medicine, Kosin University College of Medicine, Busan, Korea

Background: We investigated how quiz activities can improve summative assessment outcomes by analyzing the relationship between them.

Methods: We used 217 first-year medical students’ medical informatics data from 3 consecutive years. We analyzed summative assessment outcomes between quiz completion and incompletion groups, one-time and multiple-time quiz learning groups, and three combined comparisons between subgroups of quiz learning activity frequencies: 1 versus 2, 3, 4, and 6 (group 1), 1 and 2 versus 3, 4, and 6 (group 2), and 1, 2, and 3 versus 4 and 6 (group 3). We then analyzed correlations between the final quiz scores and summative assessment outcomes.

Results: The summative assessment means for students who completed quizzes and those who did not were 87.16±8.73 and 83.22±8.31, respectively (p=0.001). The means for the one-time and multiple-time quiz learning groups were 86.54±8.94 and 88.71±8.10, respectively (p=0.223). The means for combined subgroups were not significantly different between groups (p>0.05), although a statistically significant increasing trend was found from groups 1 to 3 (0.223>0.203>0.075 using the t-test and 0.225>0.150>0.067 using the Mann-Whitney test, respectively). Summative assessment scores were not significantly correlated with quiz scores (r=0.115, p=0.213).

Conclusions: Quizzes helped students who used self-directed learning obtain better summative assessment outcomes. Formative quizzes presumably did not provide students with direct knowledge, but showed them their weak points and motivated them to work on areas where their knowledge was insufficient.

Keywords: Correlation of data; Formative assessment; Quiz; Self-directed learning; Summative assessment

Introduction

Doctors are required to continuously update and improve their medical skills and knowledge based on changes in the field that lead to better practices in medicine, and to voluntarily learn what is necessary to meet the requirements for medical professionals [1,2]. Therefore, individuals have to decide their own learning needs to plan and implement these processes for their successful lifelong professional development [1]. From this perspective, self-directed learning has always been a cornerstone of the ideal student learning methods [3].

Received: July 4, 2022; Revised: September 1, 2022; Accepted: September 6, 2022

Corresponding Author: Hyunyong Hwang, MD, PhD

Department of Laboratory Medicine, Kosin University College of Medicine, 262 Gamcheon-ro, Seo-gu, Busan 49267, Korea
Tel: +82-51-990-6783  Fax: +82-51-990-3010  E-mail: terminom@hanmail.net

© 2022 Kosin University College of Medicine

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Students need to be able to accurately evaluate their competency of curricular topics and modulate learning goals accordingly to be successful self-directed learners [4]. Formative assessment is very important as an instructional tool in this regard [5], because it can be considered not only as an assessment resource but also a guide for students to recognize areas where they experience difficulties in their acquisition process by tracing their own progress in self-directed learning [6]. In general, the term “formative assessment” includes any activities that happen between trainers and trainees after an assessment [7]. Therefore, formative assessments are designed to help students improve learning by providing summative assessment familiarization and feedback that guides student learning [8].

Although there are various kinds of formative assessment tools, quizzes are used most often [9-11]. This is because they promote motivational completion by increasing medical student and teacher interactions [12]. Repeated testing enhances long-term information retention compared to repeated studies [13]. This implies that testing is not only an assessment tool, but also plays a significant role in student learning. Furthermore, information retrieval demonstrated by taking tests is a key to effective long-term information retention [14]. As a result, when quizzes are implemented as a method of test-enhanced learning, they can be useful for students’ learning complex sets of medical facts [15].

Quizzes where students are expected to strengthen their learning by completing the quiz activity are called formative quizzes. Formative quizzes have been reported to improve summative assessment outcomes [16-18]. Although many studies have reported the usefulness of quizzes for improving learning achievements, they have not evaluated various quiz activity conditions with respect to the quiz’s relationship to the summative assessment outcome. In other words, critical points within the quizzes should be included when determining positive learning effects. Currently, only quiz scores or quiz activity completion are used when analyzing the relationship between the quizzes and summative assessment outcome.

Much information about the relationship between formative assessments and summative assessments has already been reported in previous studies. However, the association between various quiz activities such as quiz activity frequency and sequential quiz activity with learning accomplishments has not been sufficiently validated. Therefore, we evaluated the effects of various analytical quiz activities on summative assessment outcomes in this study [19]. Several critical quiz activity points, such as quiz activity frequency and score trends for sequential quiz activities, were analyzed to determine the factors that played a role in improving summative assessment outcomes. We attempted to explore how quiz activity affected summative assessment outcomes by analyzing the data from 3 consecutive years to compare the average and final quiz scores, quiz activity frequency, and summative assessment outcomes.

Methods

Ethical statements: This study was approved by the Institutional Review Board of Kosin University Gospel Hospital (KUGH 2021-11-004). Informed written consent was exempted. Data remained confidential throughout this study.

1. Enrolled students

We provided a web-based instruction (WBI) platform during a “medical informatics” course for students in their first year of medical school. Three consecutive data from the course could be retrieved and we analyzed the data retrospectively for the analyses of this study.

2. Quiz

Learning goals provided before class time were established based on the ultimate achievements required for medical students after completion of the medical informatics course. Formative quizzes were provided for students to test and review what they learned during their medical informatics class time. Each quiz consisted of 16 questions based on the learning goals given to students in advance. Written feedback was not provided in any of the formative question results. This was intended for students to recognize their weak points in learning and implement further study voluntarily. Instead, a forum site was provided for students to ask any questions during their self-directed learning.

We analyzed 217 first-year medical students’ records from a medical informatics class taken from 2019 to 2021. Students were asked voluntarily to complete quizzes created using Moodle version 3.0 software (Martin Dougiamas, Perth, Australia; http://www.moodle.org/) (Fig. 1). Stu-
dents either installed the Moodle app, a WBI platform, on their smartphone or used the WBI website online through a computer [20].

3. Summative assessment
After completing the medical informatics course, students used a summative assessment as their final examination. The assessment’s level of difficulty differed each year, which could introduce analytical bias. To account for this, we raised the top score to 100 and adjusted other scores accordingly. The summative assessments consisted of 20 to 25 questions.

4. Summative assessment outcome for quiz completion versus quiz incompletion groups
We divided students into quiz completion and quiz incompletion groups and calculated the summative assessment outcome means for each group. We then compared the statistical differences between the two groups.

5. Quiz activity frequency and summative assessment outcome
We divided students into six subgroups according to quiz completion frequency, from 0 to 6. Once students answered 16 quiz questions in each attempt, they were considered to have completed one round of quiz activity. We analyzed the quiz completion frequency for each student and investigated whether it caused a better summative assessment outcome. We compared quiz final score means with summative assessment outcomes for each quiz frequency group.

6. Score trends for sequential quiz activity
We combined subgroups of various quiz learning activity frequencies to create three comparison combinations: 1 versus 2, 3, 4, and 6 (group 1), 1 and 2 versus 3, 4, and 6 (group 2), and 1, 2, and 3 versus 4 and 6 (group 3). We then compared the summative assessment outcome with the three groups.

7. Correlation between the summative assessment outcome and final quiz scores
We compared the final quiz scores with the summative assessment outcomes. When students performed more than one quiz learning activity, the last activity’s score was used for this analysis.

![Quiz navigation](image)

**Fig. 1.** Formative quizzes provided for students’ self-directed learning. Formative quizzes were made using the “Quiz” function on the Moodle platform. Students could select questions and were allowed unlimited attempts to solve them. The highest grade was applied for a question with multiple attempts. EMR, electronic medical record; EHR, electronic health record.
8. Statistical analysis
We used the \( t \)-test to analyze the mean differences between quiz completion and incompletion groups and between one-time quiz completion and multiple-time quiz completion groups. We used the Mann-Whitney test for non-parametric mean difference analyses. Parametric and non-parametric analyses of the mean differences depending on quiz activity frequency were analyzed using the one-way analysis of variance (ANOVA) test and Kruskal-Wallis test, respectively. We used Pearson correlation to evaluate any summative assessment outcome correlation with final quiz scores. Statistical analyses were performed using SPSS version 25 (IBM Corp., Armonk, NY, USA). Differences were considered statistically significant at \( p<0.05 \).

Results
A total of 217 students’ data were obtained for 3 consecutive years and used for analyses. The number of students enrolled for this study in 2019, 2020, and 2021 were 74, 72, and 71, respectively.

1. Summative assessment outcome for quiz completion versus quiz incompletion groups
The quiz completion group (n=119) and quiz incompletion group (n=98) summative assessment results, including standard deviations (SDs), were 87.16±8.73 and 83.22±8.31, respectively. They were significantly different (\( p=0.001 \)) (Fig. 2).

2. Quiz activity frequency and summative assessment outcome
We divided the quiz completion group into two quiz frequency groups, the one-time quiz learning group and multiple-time (2, 3, 4, and 6 times) quiz learning group. The groups consisted of 85 and 34 students, and the means±SD were 86.54±8.94 and 88.71±8.10, respectively. These results are not statistically different (\( p=0.223 \)) (Table 1). The summative assessment scores among groups were not significantly different (\( p=0.376 \) on one-way ANOVA, \( p=0.335 \) on Kruskal-Wallis test).

3. Score trends for sequential quiz activity
The number of combined subgroups is not large enough for parametric analysis, so we used a non-parametric method (Mann-Whitney test). The mean in the combined subgroups was not significantly different between groups (\( r>0.05 \)) (Table 2). Although we did not calculate a statistically significant \( p \)-value in either the parametric or non-parametric statistical analyses, the \( p \)-values showed decreasing trend from groups 1 to 3 (0.223>0.203>0.075 on \( t \)-test and 0.225>0.150>0.067 on Mann-Whitney test).

4. Correlation between the summative assessment outcome and final quiz scores
Each student’s final quiz score was compared to their summative assessment score. The mean scores were significantly different using the \( t \)-test (\( t=-3.377, p=0.001 \)).

![Fig. 2. Summative assessment outcomes for quiz completion and incompletion groups. The thick horizontal line in the middle of the box is the mean for each group's summative assessment score. The mean scores were significantly different using the \( t \)-test (\( t=-3.377, p=0.001 \)).](image)

**Table 1.** Summative assessment outcomes depending on quiz learning activity frequency

<table>
<thead>
<tr>
<th>Frequency of quiz learning activity</th>
<th>No. of patients</th>
<th>Mean±SD</th>
<th>95% CI for mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>98</td>
<td>83.22±8.31</td>
<td>(81.56–84.89)</td>
</tr>
<tr>
<td>1</td>
<td>85</td>
<td>86.54±8.94</td>
<td>(84.61–88.47)</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>87.86±7.76</td>
<td>(84.32–91.39)</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>87.29±10.44</td>
<td>(77.63–96.94)</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>92.00±4.90</td>
<td>(85.92–98.08)</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>100.00</td>
<td>-</td>
</tr>
</tbody>
</table>

SD, standard deviation; CI, confidence interval.
mative assessment score. Results show the summative assessment scores were not significantly correlated with the last quiz scores ($r=0.115$, $p=0.213$) (Fig. 3).

**Discussion**

Formative quizzes were implemented for self-directed learning and formative assessment, which are usually expected to accompany feedback from trainers. The purpose of formative quizzes in this study was not the same as those in traditional learning environments as described in the methods section. The formative quizzes were devised for students to recognize what they needed to improve upon for further studying. This was the main intention of the formative quizzes. It is assumed that any students who participated in supplementary study to compensate for a lack in knowledge during the formative quiz activities could obtain better scores on summative assessments.

Given that students were informed during the first-class period of the course that the formative quiz scores would not be included in their final grades, it is believed that the formative quizzes were mainly utilized by students as a measuring tool for their status of learning. We recognize the possibility that there were some students who had previous knowledge of the content before answering the formative quizzes. However, the formative quiz could still guide students’ learning via formative questions regardless of any prior exposure to the quiz content.

There was a significant summative assessment outcome difference between students who completed their quizzes and those who did not. This is concordant with previous studies [17,18]. The summative assessment measures the extent of learning while the formative assessments are a tool to help guide students toward their learning goals.

There is a plethora of evidence that formative assessments are associated with positive learning outcomes [6,21,22]. Feedback during formative assessments was assumed to be a core component for positive summative assessment outcomes [21]. Medical students want feedback in a timely manner, either verbally, aurally, through video, or via self-assessment [23-26]. Feedback should be different depending on the information or skills the student needs. With the intent to enable self-directed learning, we focused on students using quizzes to identify their knowledge gaps. The quiz questions focused on key medical informatics concepts and knowledge.

In general, formative assessments are low stress and stu-
Students do not feel threatened and judged when taking them \[27\]. The formative quizzes did not have a completion time and scores were not included in the final grades. Since the quizzes were not mandatory, it is likely that they did not overwhelm the students; it is plausible that the students who completed them were more motivated to learn the material than those who did not since there were no other included incentives regarding their participation \[28\]. We assumed the students who completed the quizzes were more apt to voluntarily and vigorously reflect upon their learning goals during the formative assessment portions of the course.

Repeated exposure to testing enhances self-efficacy on tests \[29\]. Information provided repeatedly over time is more easily retained than when all information is offered at once \[30\]. Information retention is more enhanced on delayed tests rather than repeated studying \[31\]. As the quiz frequency increased, the summative assessment scores increased (Table 1). Although the summative assessment outcome differences in each group were not statistically significant and the number of students in the high frequency groups was not high enough for statistical analysis, we still observed an increasing trend in the summative assessment score. The time intervals between quizzes spanned up to 100 days (data not shown). We created several combinations of quiz frequency groups to reinforce analytical power related to the low student numbers in high frequency groups. Although not statistically significant, the summative outcome scores increased as the quiz frequency combinations increased to the highest frequency combination (group 3) \(p=0.075\) and \(p=0.067\) in \(t\)-test and Mann-Whitney test, respectively (Table 2). This increase might be caused by the repeated testing over time, which likely facilitated information retention and helped students better prepare for subsequent testing.

Taking quizzes repeatedly was self-directed because there was no other participation incentive. When incentives were introduced, the number of students who scored well on the quizzes did not correspond to the number of students who scored well on the summative assessment in the previous study \[10\]. This result may be explained by the assumption that self-directed learning is strengthened when it is implemented voluntarily, without any external pressure.

As repeated tests have shown over time, we assumed that the final quiz score would correspond to the summative assessment outcome. Interestingly, there was no statistically significant correlation between the final quiz score and the summative assessment outcome (Fig. 3). This strongly implies that the formative quiz scores are not a direct predictor of summative assessment outcomes. Instead, the main role of formative quizzes is not just to provide knowledge, but to also enable students to understand in which areas improvement was needed. Although positive correlations between quiz scores and summative assessment outcomes have been reported in previous studies \[32-36\], the educational conditions in other studies were not fully comparable to those of the current one, which could explain the differences in the correlational results. In our study, students were not concerned about quiz scores because they were not included in the final grades nor did the students feel judged by their tutors, two factors that appeared to act positively on the students’ self-directed learning.

This study had some limitations. First, it focused specifically on the quiz activity’s relationship to the summative assessment outcome. It did not consider other factors such as individual tutor feedback or other available resources. Second, students took the quizzes voluntarily, which does not always mean that students who abstained from taking the quizzes did not participate in another form of self-directed learning. We didn’t analyze other possible self-directed learning actions unrelated to formative quizzes. Third, since this study was retrospectively analyzed with data recorded during a medical informatics course over 3 consecutive years, it does not include any information asking for direct responses of students regarding the degree of compliance to the intention of the formative quizzes. Fourth, we showed a significant difference in summative assessment outcomes between the quiz completion and quiz incompletion groups. This result could be reinforced through the further analysis of the difference in summative assessments depending on the students’ academic performances of all learning activities. It could not be determined whether the significant difference of summative assessments between the two groups was caused solely by the quiz activity or if it was affected by the excellence in the students’ learning abilities in total learning activities.

In conclusion, self-directed learning using quizzes is thought to be useful for a better summative assessment outcome regardless of frequency and the final score ob-
Students who performed better on their summative assessments are assumed to have improved in their weaker areas through the quiz learning activities. From this perspective, students used the quiz learning activities to overcome their lack of knowledge. The quizzes themselves did not provide direct knowledge, but instead revealed their weaker points, and were believed to have motivated them to voluntarily make up for those insufficiencies. More devices for self-directed learning need to be developed and recommended to students to help them voluntarily improve their performance. Furthermore, it is suggested that additional studies be conducted in order to analyze the difference in the summative assessment outcome according to the level of students’ excellence in their academic performance.

Article information

Conflicts of interest
Chi Eun Oh and Hyunyong Hwang are editorial board members of the journal but were not involved in the peer reviewer selection, evaluation, or decision process of this article. No other potential conflicts of interest relevant to this article were reported.

Funding
None.

Author contributions
Conceptualization: HH. Data curation: CEO, HH. Formal analysis: CEO, HH. Methodology: CEO, HH. Project administration: HH. Resources: CEO, HH. Validation: CEO, HH. Visualization: CEO, HH. Writing - original draft: CEO. Writing - review & editing: HH. Approval of final manuscript: all authors.

ORCID
Chi Eun Oh, https://orcid.org/0000-0002-0439-8170
Hyunyong Hwang, https://orcid.org/0000-0003-0662-3041

References

17. McNulty JA, Espiritu BR, Hoyt AE, Ensminger DC, Chan-


