

Medical Students Academic Performance Assessment in Physiology Courses Using Formative and Summative Quizzes at SMBB Medical College Karachi, Pakistan

Muhammad Noman Rashid^{1,*}, Ali Muhammad Soomro², Aftab Hussain Abro³,
Syeda Beenish Noman⁴

¹Department of Physiology, Research and Ethics Department Shaheed Mohtarma Benazir Bhutto Medical College, Karachi, Pakistan

²Department of Physiology, University of Sindh, Jamshoro, Pakistan

³Department of Community Health Sciences, Shaheed Mohtarma Benazir Bhutto Medical College, Karachi, Pakistan

⁴Department of Obstetrics and Gynecology, Aga University Hospital, Karachi, Pakistan

Email address:

drnomanrashid@outlook.com (M. N. Rashid)

*Corresponding author

To cite this article:

Muhammad Noman Rashid, Ali Muhammad Soomro, Aftab Hussain Abro, Syeda Beenish Noman. Medical Students Academic Performance Assessment in Physiology Courses Using Formative and Summative Quizzes at SMBB Medical College Karachi, Pakistan. *Advances in Applied Physiology*. Vol. 2, No. 1, 2017, pp. 10-17. doi: 10.11648/j.aap.20170201.12

Received: April 17, 2017; **Accepted:** May 16, 2017; **Published:** August 28, 2017

Abstract: Background Formative practice quizzes have become common resources for self-evaluation and focused reviews of course content in the medical curriculum. In the current study two separate studies were conducted to (1) compare the effects of a single or multiple voluntary practice quizzes on subsequent summative examinations and (2) examine when students are most likely to use practice quizzes relative to the summative examinations. Material and Methods In the first study, providing a single online practice quiz followed by instructor feedback had no effect on examination average grades compared to the previous year or student performances on similar questions. However, there were significant correlations between student performance on each practice quiz and each summative examination ($r = 0.42$ and $r = 0.24$). When students were provided multiple practice quizzes with feedback (second study). Results there was a weak correlation between the frequency of use and performance on each summative examination ($r = 0.17$ and $r = 0.07$). The frequency with which students accessed the practice quizzes was greatest the day before each examination. In both studies, there was a decline in the level of student utilization of practice quizzes overtime. Conclusion we concluded that practice quizzes provide some predictive value for performances on summative examinations. Second, making practice quizzes available for longer periods prior to summative examinations does not promote the use of the quizzes as a study strategy because students appear to use them mostly to assess knowledge one to two days prior to examinations.

Keywords: Formative Assessment, Physiology Education, Practice Tests, Examinations, Summative Assessment, Study Tools, Online Examination

1. Introduction

There is extensive literature addressing the learning processes that encode knowledge and incorporate experiences. The broader theories of self-regulated learning and feedback intervention (Butler and Winne, 1995; Hattie and Timperley, 2007; Evans, 2013) include the memory processes involving retrieval, where students actively reconstruct their knowledge

through recall on practice examinations and self-testing (Kulik et al., 1984; Cook et al., 2006; Roediger and Karpicke, 2006; Streips, 2007; Larsen et al., 2008; Kromann et al., 2009; Pyc and Rawson, 2010; Karpicke and Blunt, 2011; Roediger et al., 2011). Formative self-assessments also serve the purpose of informing students of their level of mastery of specific subject areas prior to taking summative examinations. Ideally, formative self-assessment not only fulfills the needs of the

learner, but it also fulfills a teacher's obligations to the student by providing appropriate resources to aid in learning. Additional incentives come from the accrediting institution, the Liaison Committee on Medical Education (LCME), which in the Educational Standard ED-31 directs that a course, "...should provide alternate means (e.g., self-testing, teacher consultation) that will allow medical students to measure their progress in learning" (LCME, 2010). Because of the efficiency and ease of distribution made possible through the internet and other online resources, the use of formative practice quizzes has proliferated in recent years. Question databases are increasingly popular resources for medical students, who use them for self-evaluation and to provide more focused reviews in specific subjects. Several papers have demonstrated that students who participate in taking formative practice quizzes tend to perform better on their summative evaluations (Olson and McDonald, 2004; Krasne et al., 2006; Kibble, 2007; Dobson, 2008; Velan et al., 2008; Carrillo-de-la-Pena et al., 2009; Leaf et al., 2009; Kibble, 2011; Kibble et al., 2011; Logan et al., 2011) although some authors have failed to find strong associations between participation on formative practice quizzes and performance on summative examinations (Johnson, 2006; Urtel et al., 2006; Palmer and Devitt, 2008). Some of the discrepancies in the literature may be related to variations in the paradigms tested. For example, the studies cited above varied in the length of time practice quizzes were made available, the scope of the practice quizzes and whether they were repeated, the level of similarity of assessments, whether feedback was provided, the frequency of practice quizzes, sample size, and characteristics of students (year of study), whether or not practice quizzes were optional, whether incentives were provided for taking practice quizzes, as well as differences in the curriculum and summative examinations. In the current study the researchers therefore, undertook the objectives to compare effects of different paradigms of formative practice quizzes on student performances in summative examinations within the same curriculum. Although our studies occurred in different years and with different students, the curriculum and the faculty who taught the course were the same during the study years. The current study further tested the effectiveness of formative quizzes by examining individual performances on questions that were similar in both the practice quiz and summative examination. Finally, we sought to extend the literature on formative testing by determining optimal times to provide formative practice quizzes. This article is divided according to each of the respective studies to include methods specific to the study and the results of that study.

2. Methods

2.1. Participants

The participants were first and second year medical students (100 per year) participated in various courses of Human Physiology at the Shaheed Mohtarma Benazir Bhutto Medical College, Lyari, Karachi.

2.2. Examination Applications

A web-based examination database and applications was developed as part of the Shaheed Mohtarma Benazir Bhutto Medical College Medical Education Network to provide on-line examinations (McNulty et al., 2011). The examination interface allows students to cross out answers, add notes, and submit answers for later review in the event they were unsure of the answer (McNulty et al., 2007). The in-house examination applications were constructed.

All questions utilized for the practice quizzes and summative examinations were written by experienced faculty teaching in the course, thereby establishing content validity. This also ensured that style and difficulty of the questions were similar in both the formative and summative assessments. Questions present in the practice quizzes covered the same subject content assessed during summative examinations. The reliability of examinations is based on more than 10 years of data showing low interassay variability of discrimination factors of individual questions. Moreover, correlations of individual student performances across examinations are high (ranging between $r = 0.68$ and $r = 0.78$). All examinations and quizzes were composed of both factual and conceptual, clinical questions (Burns, 2010).

3. Data Collection and Analysis

The frequency of access of practice quizzes by students was obtained from server logs and the performances of individual students on practice quizzes and summative examinations were extracted from the examination report applications. All data were entered into Excel spreadsheets and the names of students deleted prior to further analyses to preserve anonymity. The data were analyzed by Student's *t*-test, regression analysis, and analysis of variance (ANOVA). The study was reviewed and approved by the Institutional Review Board at the Shaheed Mohtarma Benazir Bhutto Medical College, Lyari, Karachi, Pakistan.

4. Study #1: Associations Between Single Practice Quiz and Summative Evaluation

The first study tested the hypothesis that giving a formative practice quiz and follow-up oral review would improve the overall outcome on summative examinations and the outcome on specific questions. The study addressed the following specific questions controlling within our limits a number of variables including similar examination questions, similar faculty, and similar.

Curriculum in both years of the study: (1) is there an association between performance on a single practice quiz and the summative examination? (2) Does a single practice quiz improve the performance on summative examinations? And (3) does a single practice quiz improve the performance on similar questions in summative examinations?

4.1. Methods Specific to Study #1

Examination method. There were three summative examinations given each year of the course. In both years of the study, the first two examinations had the same number of questions with 12 identical questions used in both years (six questions per examination). In Year 2 of Study #1,

Prior to each of the first two examinations, voluntary practice quizzes with 50 questions each were provided to students using the on-line examination applications. Three days prior to each examination, students were given the opportunity to complete the practice quiz during a 24-hour period. The students were instructed that the practice quizzes were intended to inform them of their level of preparedness and to treat these practice quizzes as if they were taking a summative examination (e.g., do not look up answers). The day following release of the practice quiz (when it had been closed), the first author reviewed each practice quiz question in the lecture hall to explain the questions and answers. These review sessions were entirely voluntary and attendance was estimated to be >90%. In both years of the study, students had access to several other on-line class tutorials for self-evaluations.

4.2. Results of Study #1

Table 1 shows that the average scores of the two examinations for both years of the study were similar (ANOVA; $P>0.05$) indicating that the introduction of practice quizzes prior to the examinations had no effect on summative performances. For the 12 identical questions included in both years (six in Examination 1; six in Examination 2), students performed better on four questions, worse on four questions, and about the same on four questions in the year that they were provided a practice quiz. Although the practice quiz

was voluntary, only three students (2%) elected not to take it for the first examination and another five (4%) did not complete the first practice quiz. These numbers increased for the second practice quiz when 17 students (12%) elected not to take the quiz and another 15 students (11%) did not complete it. Regression analysis (excluding students who did not complete the practice quizzes) showed significant correlations ($P<0.01$) between individual scores on the practice quizzes and both summative examinations, although the effect size for these correlations was medium for the first examination ($r = 0.42$) and small for the second ($r = 0.24$). There was improvement in the average score from the first practice quiz (mean 563.2) to the second practice quiz (x 570.9), excluding those students who did not complete the examination. There was a large effect size for the correlations of individual student performances (grades) on both summative examinations ($r = 0.68$). The effect size for correlations of individual student performances on the two practice quizzes was moderate ($r = 0.32$). To better understand the effectiveness of a practice quiz on the outcome of a summative examination, six additional questions were included on the first practice quiz that were very similar in content and/or concept to questions on the first examination. Table 2 lists the phrases identifying the content and/or concept that were entered into the examination database. The results from Table 2 show that the class did better on two of these questions in the summative examination, worse on two questions, and about the same on remaining two questions. On one question (Q 1, Table 2), 26 students who answered it correctly on the practice quiz, missed the question on the summative examination even though the stem was worded exactly the same, and three of the five choices were also the same.

Table 1. Results of Both Summative Examinations for Each Year of the Study.

Yearly Summative Examination	N	Examination mean% (\pm SD)	Identical Questions (06 / Examination)					
Examination 1			01	02	03	04	05	06
Year 1	100	84.0(\pm 7.9)	40	09	14	10	20	07
Year 2 with practice quiz	100	83.1(\pm 7.5)	32	17	10	24	10	07
Examination 2								
Year 1	100	81.6(\pm 8.1)	14	29	11	20	11	15
Year 2 with practice quiz	100	83.4(\pm 8.4)	18	10	30	20	10	12

Summative examinations show that the introduction of a practice quiz prior to each examination in Year 2 had no significant effects on summative performances (ANOVA; $P>0.05$).

Table 2. Counts of the Total Number of Students Who Missed Questions that were Similar on the Practice Quiz and High Stakes Summative Examination.

Concepts Tested		Total Missed	Missed both	Correct/ Missed
Practice	Introduction	15	04	16
Summative		20		
Practice	Cell Biology	34	07	26
Summative		31		
Practice	Locomotion	32	10	05
Summative		13		
Practice	Respiratory	55	10	11
Summative		17		
Practice	Cardiovascular	34	12	12
Summative		31		

Counts are also provided for students who missed the question on both the practice quiz and the summative examination (Missed both), and those who answered it correctly on the practice quiz, but missed it on the summative examination (Correct/missed).

5. Study #2: Associations Between Multiple Practice Quizzes and Summative Evaluations

The first study included only a single practice quiz with follow-up review by faculty. The second study tested the hypothesis that there are associations between the frequency with which students take multiple practice quizzes and their performance on summative examinations. The study design addressed the following specific questions: (1) how frequently did students use the practice quizzes? (2) When did they use them relative to the time of the summative examinations? And (3) was there an association between use of practice quizzes and performance on summative examinations?

5.1. Methods Specific to Study #2

In this study, practice quizzes were not made available to the students before the first course examination to obtain a base line for individual student performance without the availability of practice quizzes. Following the first examination, a practice quiz was released to students immediately following each lecture. The practice quizzes were created and made available to students through Moodle, an open-source course management system that was also used as the course forum for asynchronous interactions with faculty and other students. Students were informed at the beginning of the course that the formative quizzes were provided for self-assessment and the class was notified through email when new quizzes were posted online. The authors developed a total of 14 practice quizzes that were specific to each of the 14 anatomy lectures given during the final two-thirds of the course. The number of questions per quiz ranged from 8 to 41 (mean 520). The quizzes were composed of multiple-choice questions in a format similar to questions students encountered on their summative examinations. Students used their school-provided username and password to log into Moodle to access the practice quizzes. Once a practice quiz was released, it was available to students throughout the remainder of the course,

and they were free to take it as many times as they wanted. Instructive feedback was provided by the server for correct and incorrect answers on submission of each practice quiz. Server logs included the date and time which students accessed and “submitted” each quiz.

5.2. Results Specific to Study #2

Daily access logs showed that students who accessed the practice quizzes typically “submitted” the quiz to receive feedback on the questions. Accordingly, data are presented only on the frequency of “submitted” quizzes. The logs for each of the practice quizzes (Table 3) revealed several trends. First, the frequency with which students accessed the quizzes was greatest the day before the examination with a large number even on the day of the examination. Second, the total.

Number of quizzes accessed decreased during the final third of the course compared to the second third even though there were two more class days during the third part of the course. The decline in the submission of practice quizzes from the second to the third examinations was evident in the increased number of students who decided not to submit any practice quizzes and the greater number of students who decreased their use of practice quizzes for the third examination (Table 4). Only 15 students viewed all 14 of the practice quizzes. There was a significant, but weak, association between the frequency with which students submitted the practice quizzes and their performance on the second examination ($r = 0.17$; $P = 0.037$). The average examination score for those students who completed 6–7 of the self-quizzes (mean 588.5) was significantly higher than the average examination score (mean 585.1) for those students who submitted 0–1 practice quizzes ($t = 52.79$; $df = 5111$; $P = 0.006$; $d = 0.52$). There were no significant associations between the frequency with which students took the practice quizzes and their performance on the third examination in the course ($r = 0.07$; $P > 0.05$). This absence of significant associations on the third examination corresponded to a larger number of students ($n = 562$, Table 4) who submitted fewer practice quizzes prior to Examination 3 than they did for Examination 2. The examination averages of these students did not change from Examination 2 to Examination 3. By comparison, the 24 students who decided to submit more practice quizzes prior to Examination 3 exhibited an average increase of 1.7% on Examination 3 ($P = 0.038$; $t = 51.84$; $df = 523$; $d = 0.77$ —paired Student’s t -test, one-tailed).

Table 3. The Frequency with Which Students Accessed Each Practice Quiz Daily Preceding the Second and Third Summative Examinations.

SECOND EXAMINATION															
Cardio vascular system	09	08	08	05	04	09	10	12	15	05	18	13	45	07	11
Respiratory system	11	12	45	42	11	10	09	11	25	42	13	70	11	12	10
Urinary system	42	13	18	11	20	21	12	13	25	54	11	19	25	21	11
Gastro intestinal system	10	12	25	85	26	55	12	14	23	24	26	11	14	45	44
THIRD EXAMINATION															
Reproductive System	03	11	15	17	25	15	11	06	15	40	24	21	12	13	14
Central Nervous system	14	57	20	22	49	25	22	13	14	25	23	23	14	12	44
Peripheral nervous system	12	10	31	22	20	12	55	21	23	12	14	15	24	22	45
Special senses	10	13	25	14	12	11	15	22	36	21	12	14	15	25	55

Weekends are depicted in *italics and underlined*; Summative examinations included content areas for that part of the course (i.e., examinations were not cumulative).

Table 4. Metrics on the Number of Practice Quiz Submissions by Students Prior to Each of the Summative Examinations in Study #2.

	EXAMINATION 2	EXAMINATION 3
Total number of submissions	867	697
Number of students with 0 submissions	52	82
Number of students with 1–2 submissions	24	07
Number of students with seven submissions	36	25
Number of students who increased use of practice quizzes from examination 2 to examination 3.	00	24
Number of students who decreased use of practice quizzes from examination 2 to examination 3.	00	62
Number of students who did not change use of practice quizzes from examination 2 to examination 3.	00	64
Average grade of students who increased use of practice quizzes	85.3	87.0*
Average grade of students who decreased use of practice quizzes.	86.7	86.9
Average grade of students who did not change their use of practice quizzes	86.4	86.6

There was a tendency for students to decrease their use of practice quizzes, but students who increased use of practice quizzes had an increased average grade on the second examination. $Ap = 0.039$; $t = 1.84$; $df = 23$; $d = 0.77$ (paired Student's *t*-test, one-tailed).

6. Discussion

Our results demonstrate that student performances on practice quizzes can be associated with performances on their summative evaluations. Participation in voluntary practice quizzes, regardless of performance, also was weakly associated with better performance on summative examinations. These observations on performance and participation on practice quizzes have been reported in several previous studies (Johnson, 2006; Kibble, 2007; Dobson, 2008; Velan *et al.*, 2008; Carrillo-de-la-Pena *et al.*, 2009; Leaf *et al.*, 2009; Kibble, 2011). Outcomes that have been highlighted as advantages of practice quizzes are their importance as “predictors” of performance on summative examinations (Krasne *et al.*, 2006; Kibble, 2007; Kibble *et al.*, 2011) and their positive effects on learning (Olson and McDonald, 2004; Velan *et al.*, 2008). Predictive values of various assessments from Medical College Admission Test scores and undergraduate education to outcomes in medical school (such as the physiology course) include significant correlations with student performances in their national board examinations and subsequent residencies (Ripkey *et al.*, 1998; Holtman *et al.*, 2001; Peterson and Tucker, 2005; Hamdy *et al.*, 2006; Donnon *et al.*, 2007; Gandy *et al.*, 2008; Poole *et al.*, 2012). From these results and those of others, we suggest that the positive associations between practice quizzes and summative evaluations are related broadly to the overall academic achievement of the students. This conclusion is based on the large effect size for correlations we found between summative examinations testing different material. Kibble *et al.* (2011) also reported equally high correlations for examinations testing unrelated content. From the results of our first study, it is difficult to conclude that students learned their physiology better as a result of the formative practice quizzes. First, the class averages on examinations and the performance on identical questions

were not significantly different when summative examinations were preceded by a practice quiz. Second, there were no significant effects on the performance of students when very similar questions were included on both the formative and summative examinations. One question had identical stems and two of the five choices were the same; yet seven students who missed it on the practice quiz also missed it on the examination while 26 other students missed the question even though they answered it correctly on the practice quiz. This would suggest that students who missed it on the practice quiz learned from their mistakes, but the other 26 students either never learned it (and guessed correctly on the practice quiz) or knew the material for the practice quiz and forgot it 2 days later when they took the examination. The latter explanation seems less likely because most students ‘study time intensifies as examinations draw nearer. This is supported by the fact that averages on the practice quizzes tended to be lower than those on the examinations, which raises the issue of construct validity of the practice quiz. However, significant correlations across practice and summative examinations suggest that the lower practice scores were related to students intensifying their studies closer to the time of the summative examinations. A common concern among faculty and administrators is question exposure, or the possibility that reuse of questions may lead to its recognition by students on future examinations. Although we did not test performances on identical questions given in the practice and summative examinations, our results from very similar questions suggest that question exposure is not an important issue, even when separated by only a couple of days. Wood (2009) also showed that repeat examinees did not appear to remember questions they had seen before.

The present result showing that student “learning,” on aggregate, did not improve following a single practice quiz with follow-up review conflicts with literature on the benefits of active retrieval enhancing retention (the “testing effect”), which has been demonstrated in a variety of learners and conditions (Roediger and Karpicke, 2006; Pyc and Rawson, 2010; Wissman *et al.*, 2011; Carpenter and Kelly, 2012). In their meta-analysis of the effects of practice on test scores,

Kulik *et al.* (1984) reported that larger effect sizes

occurred: (1) when identical forms of tests were used; (2) when frequency of testing was increased; and (3) in students of higher “ability.” Conceivably, our negative finding was related to only a single practice quiz before each examination. However, Velan et al. (2008) found that repeated attempts at formative practice quizzes did not increase the examination scores. Several studies have reported direct associations between the voluntary use of formative practice quizzes and achievement (i.e., students who participated tended to have better grades; Johnson, 2006; Kibble, 2007; Velan et al., 2008; Carrillo-de-la-Pe~na et al., 2009; Kibble, 2011). In our second study where multiple practice quizzes were available for the students, we found a significant, but weak, association between the frequency with which students took the practice quizzes and their performance on their examination, but only on one of the two examinations. The lack of associations on the third examination may have been due to the large number of students who submitted fewer practice quizzes prior to the third examination than they did for the second examination. However, positive effects of taking practice quizzes were suggested by our finding that students who increased their use of practice quizzes also increased their scores on examinations compared to those students who either voluntarily decreased use or did not change their level of use. In spite of any trends, it is not possible to conclude any causal effects between use of practice quizzes and performance on examinations. As Johnson (2006) concluded, “It was not clear if quiz use caused achievement or achievement caused quiz use.”

A common finding in this and previous studies is that significant numbers of students elect not to participate in these voluntary exercises. In the current study, as many as 29% of the students elected not to take any of the practice quizzes in the second study. Participation rates for voluntary quizzes were even lower in other studies, ranging between 40 and 70% (Olson and McDonald, 2004; Johnson, 2006; Kibble, 2007; Carrillo-de-la-Pe~na et al., 2009). There was 12% nonparticipation

Even when students were regularly reminded that non participation was linked to poor examination results (Kibble, 2011). In a recent survey of medical students, McNulty et al. (2012) reported that practice quizzes were one of the least commonly used study strategies in the majority of basic science courses.

The process of self-selection for any voluntary exercise presents a limitation because a selection bias cannot be excluded from the interpretation of results. The reasons for not participating are not known, although it has been speculated that it relates to motivation and competency of the students (Johnson, 2006; Leaf et al., 2009). Not surprisingly, participation on formative quizzes increased when credit was given for participation (Kibble, 2007). Nonparticipation extends to another trend; a decline in the rate of participation over time. This decline was observed in both of this studies and has been reported to occur with practice quizzes (Kibble et al., 2011) and the use of other voluntary course resources

whether it is an audience response system (Hoyt et al., 2010) or an online dissector (Mc Nulty et al., 2004). These declines may be the result of students discovering that a specific resource does not fit their learning style and/or increasing demands on their time as the course progresses (Smythe and Hughes, 2008). One objective of our study was to determine optimal times to give practice quizzes. When practice quizzes were made available for several days prior to an examination, the highest frequency of access occurred repeatedly the day before the examination. For five of the 14 practice quizzes, the second highest rate of access occurred on the day of the examination. This pattern suggests to us that many of the students were using the practice quizzes to assess their knowledge in specific content areas or “fact check.” These results suggest that students were not using the practice quizzes to learn their anatomy. It is noteworthy that the timing of the practice quiz in our first study coincided with the time most students self-select to take them.

To summarize, because formative self-testing resources are rapidly becoming an important component of the medical curriculum, refined measures of their effectiveness are needed, especially under the variety of medical education settings and formats (e.g., lectures, small groups, clinical activities). One purpose of formative assessment is to inform students of their learning progress and to identify learning errors that require intervention. This purpose assumes students will use practice quizzes on a regular basis to monitor their learning and correct knowledge errors discovered through these assessments. This study indicated that voluntary utilization of practice quizzes declines over time, and that students tend to use these assessments one day prior to summative examinations to check their content knowledge. Checking their knowledge just prior to summative examinations provides little time for students to engage in deeper examination of their misconceptions and suggests that students only engage in “cramming” to correct errors in their learning and do not use practice quizzes to their fullest formative purpose. From our results, practice quizzes with feedback interventions did not seem to promote the appropriate strategies associated with self regulated learning theories (e.g., planning, internalization of standards of learning, self-evaluation, etc.; Butler and Winne, 1995).

Formative assessment also has the purpose of improving instruction by allowing faculty to gather information regarding student learning and to take instructive action to correct misunderstandings evident in the formative assessment results. Study #1 showed that voluntary participation was high if the practice quiz was given 2–3 days before the summative examination and included opportunity to receive faculty feedback in the classroom prior to the summative examination. However, this time frame provides minimal opportunity for faculty to use students’ performance data to correct students’ misconceptions at a deeper level beyond discussions on which answers were correct or incorrect, or for faculty to adjust their instruction practice to facilitate student learning. Future studies should begin to explore why medical students are not using practice quiz

resources to their fullest formative value and how to facilitate students' use of these resources as well as, how medical school faculty can use these resources more constructively to promote student learning and improve instructional practices.

In addition to the limitations associated with the possible bias of self-selection noted above, we did not collect information on the reasons students elected not to participate in the practice quizzes. Interpreting the effects of a single practice quiz on summative performances (Study #1) has limitations because of the many variables affecting those performances.

Our study also would have benefited from an instrument to measure student motivation and level of engagement in the course (Urtel *et al.*, 2006) to determine the degree to which these attributes are associated with participation in voluntary activities such as the practice quizzes.

In conclusion, although frequency of taking practice quizzes provides some predictive value, the associations with performance on summative assessments can be small. Any causal relationships between the two should be viewed cautiously. Finally, intrinsic factors that might explain participation in voluntary formative quizzes include the level of motivation and engagement of the student in the course or clerkship. In future studies, we will attempt to measure the qualitative parameters that influence medical student participation in voluntary formative assessments.

References

- [1] Burns ER. 2010. "Anatomizing" reversed: Use of examination questions that foster use of higher order learning skills by students. *Anat Sci Educ* 3:330–334.
- [2] Butler DL, Winne PH. 1995. Feedback and self-regulated learning: A theoretical synthesis. *Rev Educ Res* 65:245–281.
- [3] Carpenter SK, Kelly JW. 2012. Tests enhance retention and transfer of spatial learning. *Psychon Bull Rev* 19:443–448.
- [4] Carrillo-de-la-Pe~na MT, Baille's E, Caseras X, Mart_inez A, Ortet G, P_erez J. 2009. Formative assessment and academic achievement in pre-graduate students of health sciences. *Adv Health Sci Educ Theory Pract* 14:61–67.
- [5] Cook DA, Thompson WG, Thomas KG, Thomas MR, Pankratz VS. 2006. Impact of self-assessment questions and learning styles in web-based learning.
- [6] A randomized, controlled, crossover trial. *Acad Med* 81:231–238.
- [7] Dobson JL. 2008. The use of formative online quizzes to enhance class preparation and scores on summative exams. *Adv Physiol Educ* 32:297–302.
- [8] Donnon T, Paolucci EO, Violato C. 2007. The predictive validity of the MCAT for medical school performance and medical board licensing examinations: A meta-analysis of the published research. *Acad Med* 82:100–106.
- [9] Evans C. 2013. Making sense of assessment feedback in higher education. *Rev Educ Res* 83:70–120.
- [10] Gandy RA, Herial NA, Khuder SA, Metting PJ. 2008. Use of curricular and extracurricular assessments to predict performance on the United States Medical Licensing Examination (USMLE) Step 1: A multi-year study. *Learn Assist Rev* 13:27–35.
- [11] Hamdy H, Prasad K, Anderson MB, Scherpbier A, Williams R, Zwierstra R, Cuddihy H. 2006. BEME systematic review: Predictive values of measurements obtained in medical schools and future performance in medical practice. *Med Teach* 28:103–116.
- [12] Hattie J, Timperley H. 2007. The power of feedback. *Rev Educ Res* 77:81–112. Holtman MC, Swanson DB, Ripkey DR, Case SM. 2001. Using basic science subject tests to identify students at risk for failing Step 1. *Acad Med* 76: S48–S51.
- [13] Hoyt A, McNulty JA, Gruener G, Chandrasekhar A, Espiritu B, Ensminger D, Price R Jr, Naheedy R. 2010. An audience response system may influence student performance on anatomy examination questions. *Anat Sci Educ* 3:295–299.
- [14] Johnson GM. 2006. Optional online quizzes: College student use and relationship to achievement. *Can J Learn Tech* 32:61.
- [15] Karpicke JD, Blunt JR. 2011. Retrieval practice produces more learning than elaborative studying with concept mapping. *Science* 331:772–775.
- [16] Kibble J. 2007. Use of unsupervised online quizzes as formative assessment in a medical physiology course: Effects of incentives on student participation and performance. *Adv Physiol Educ* 31:253–260.
- [17] Kibble JD. 2011. Voluntary participation in online formative quizzes is a sensitive predictor of student success. *Adv Physiol Educ* 35:95–96.
- [18] Kibble JD, Johnson TR, Khalil MK, Nelson LD, Riggs GH, Borrero JL, Payer AF. 2011. Insights gained from the analysis of performance and participation in online formative assessment. *Teach Learn Med* 23:125–129.
- [19] Krasne S, Wimmers PF, Relan A, Drake TA. 2006. Differential effects of two types of formative assessment in predicting performance of first-year medical students. *Adv Health Sci Educ Theory Pract* 11:155–171.
- [20] Kromann CB, Jensen ML, Ringsted C. 2009. The effect of testing on skills learning. *Med Educ* 43:21–27.
- [21] Kulik JA, Kulik CC, Bangert RL. 1984. Effects of practice on aptitude and achievement test scores. *Educ Res J* 21:435–447.
- [22] Larsen DP, Butler AC, Roediger HL III. 2008. Test-enhanced learning in medical education. *Med Educ* 42:959–966.
- [23] LCME. 2010. Liaison Committee on Medical Education. Functions and Structure of a Medical School: Standards for Accreditation of Medical Education Programs Leading to the M. D. Degree. June, 2010, Washington, DC: Liaison Committee on Medical Education. 28 p.
- [24] Leaf DE, Leo J, Smith PR, Yee H, Stern A, Rosenthal PB, Cahill-Gallant EB.
- [25] Pillinger MH. 2009. SOMOSAT: Utility of a web-based self-assessment tool in undergraduate medical education. *Med Teach* 31: e211–e219.
- [26] Logan JM, Thompson AJ, Marshak DW. 2011. Testing to enhance retention in human anatomy. *Anat Sci Educ* 4:243–248.

- [27] McNulty JA, Halama J, Espiritu B. 2004. Evaluation of computer-aided instruction in the medical gross anatomy curriculum. *Clin Anat* 17:73–78.
- [28] McNulty JA, Sonntag B, Sinacore J. 2007. Test-taking behaviors on a multiple-choice exam are associated with performance on the exam and with learning style. *Med Sci Educ* 17:52–57.
- [29] McNulty J, Chandrasekhar A, Hoyt A, Gruener G, Espiritu B, Price R. 2011. Computer-based testing in the medical curriculum: A decade of experiences at one school. *J Educ Comput Res* 45:287–295.
- [30] McNulty JA, Ensminger DC, Hoyt AE, Chandrasekhar AJ, Gruener G, Espiritu B. 2012. Study strategies are associated with performances in basic science courses in the medical curriculum. *J Educ Learn* 1:1–11.
- [31] Olson BL, McDonald JL. 2004. Influence of online formative assessment upon student learning in biomedical science courses. *J Dent Educ* 68:656–659.
- [32] Palmer EJ, Devitt PG. 2008. Limitations of student-driven formative assessment in a clinical clerkship. A randomised controlled trial. *BMC Med Educ* 8: 29.
- [33] Peterson C, Tucker RP. 2005. Medical gross anatomy as a predictor of performance on USMLE step 1. *Anat Rec* 283B: 5–8.
- [34] Poole P, Shulruf B, Rudland J, Wilkinson T. 2012. Comparison of UMAT scores and GPA in prediction of performance in medical school: A national study. *Med Educ* 46:163–171.
- [35] Pyc MA, Rawson KA. 2010. Why testing improves memory: Mediator effectiveness hypothesis. *Science* 330:335.
- [36] Roediger HL, Karpicke JD. 2006. Test-enhanced learning: Taking memory tests improves long-term retention. *Psychol Sci* 17:249–255.
- [37] Roediger HL, Agarwal PK, McDaniel MA, McDermott KB. 2011. Testenhanced learning in the classroom: Long-term improvements from quizzing. *J Exp Psychol Appl* 17:382–395.
- [38] Ripkey DR, Swanson DB, Case SM. 1998. School-to-school differences in Step 1 performance as a function of curriculum type and use of Step 1 in promotion/ graduation requirements. *Acad Med* 73: S16–S18.
- [39] Smythe G, Hughes D. 2008. Self-directed learning in gross human anatomy: Assessment outcomes and student perceptions. *Anat Sci Educ* 1:145–153.
- [40] Streips UN. 2007. Quizzes are for learning not just evaluation. *Med Sci Educ* 17: s7–s8.
- [41] Urtel MG, Bahamonde RE, Mikesky AE, Udry EM, Vessely JS. 2006. On-line quizzing and its effect on student engagement and academic performance. *J. Scholarsh Teach Learn* 6:84–92.
- [42] Velan GM, Jones P, McNeil HP, Kumar RK. 2008. Integrated online formative assessments in the biomedical sciences for medical students: Benefits for learning. *BMC Med Educ* 8:52.
- [43] Wissman KT, Rawson KA, Pyc MA. 2011. The interim test effect: Testing prior material can facilitate the learning of new material. *Psychon Bull Rev* 18: 1140–1147.
- [44] Wood TJ. 2009. The effect of reused questions on repeat examinees. *Adv Health Sci Educ Theory Pract* 14: 465–473.