Evaluation of the Effectiveness of Online Peer-based Formative Assessments (PeerWise) to Enhance Student Learning in Physiology: A Systematic Review Using PRISMA Guidelines

Ahmed Sayed Khashaba
Riyadh Elm University, Kingdom of Saudi Arabia

To cite this article:

Evaluation of the Effectiveness of Online Peer-based Formative Assessments (PeerWise) to Enhance Student Learning in Physiology: A Systematic Review Using PRISMA Guidelines

Ahmed Sayed Khashaba

Abstract

Due to the emerging demands on shifting focus towards the development of more student-centered and engaging learning experiences, this systematic review elucidates the effectiveness of PeerWise introduction into the blended learning model in Physiology education based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines. Twenty electronic databases were utilized to access related studies between years 2010 to April 2020. A total of eight recent articles on PeerWise in physiology were analyzed. Three studies were conducted among medical students, and five studies were among other courses (i.e., Pharmacy, Biomedical Science, Optometry, and Human Physiology). Majority of the study designs were of cross-sectional quantitative and qualitative studies. Data extracted from the articles include (i) the pattern of PeerWise usage, (ii) the association between PeerWise and academic achievement, (iii) the level of student engagement, (iv) the quality of questions created and (v) students’ perceptions. Four emerging themes were identified among students' perceptions; (i) learning competency, (ii) fun learning experience, (iii) engagement with peers, and (iv) motivation. Methodological quality and risk of biased were assessed; and research gaps, limitations and recommendations were addressed. The present review serves as a guide for new authors to refine their knowledge and improve future research in the topic area.

Introduction

Physiology is one of the foundation sciences for medicine and other allied health sciences courses. However, students find learning physiology to be challenging, mainly on how the subject was delivered, the nature of the subject which requires causal reasoning and highly integrative as well as the characteristics of the students that learn by memorizing contents and unable to correlate clinically (Michael, 2007). Moreover, non-medical students perceived physiology to have minimal clinical relevance to their professions (Andrew et al., 2015) in which may contribute to a superficial-learning approach without deeper learning.

In response to these challenges, various innovations, and trends in medical education that have been undertaken globally to improve learning experiences which include student-centered learning (SCL), problem-based learning (PBL), case-based learning (CBL), team-based learning (TBL) and integrated learning. Student-centered learning has grown much interest over the past few decades for its effectiveness in training students, especially in Anatomy and Physiology subjects (Arroyo-Jimenez et al., 2005). In SCL, students ideally take the initiative and responsibility for their learning (Benedict, Schonder & Mcgee, 2013).

For this review, it was of interest to investigate one widely used and promising SDL tool which is Online Peer-based Formative Assessments (PeerWise) above others, due to its concept of contributing student pedagogy (CSP) that requires minimal facilitation from the instructors yet effective in promoting active student engagement and sharing of resources with an element of technology integration. PeerWise is a free online program (http://peerwise.cs.auckland.ac.nz/) that is implemented in a unit course as a support for the traditional face-to-face teaching where instructors/teachers have students to create multiple-choice questions (MCQs) related to a course and then answer, rate and comment on questions submitted by their peers throughout the course of study, either as a compulsory form of assessment with a participation grade or voluntarily at their own pace (Luxton-Reilly & Denny, 2010). This application provides opportunities for students to become active learners by creating their questions, provide the explanation for correct and wrong answers, critically review and
evaluate questions and answers prepared by their peers, provide feedback, as well as rate the quality of the questions of which will develop a deeper understanding on the course contents and stimulate higher-order cognitive skills (Hardy et al., 2014; Mcqueen et al., 2014). Above all, students will be rewarded with badges, trophies, and leaderboards for their active participation which turns out to be a motivation factor for them to strive better in the next course.

The design of online assessment activity is crucial for effective learning and properly aligning the assessment to the learning outcomes can produce a constructive learning practice (Biggs & Tang, 2011). To assert that online assessment served the intended purpose, the value of the tool must be determined in the assessment activity by assessing learner’s learning in the subject (Baleni, 2015). Hence, there is wide interest to investigate the relationships between PeerWise usage and learning outcomes across a wide range of academic disciplines. Previous studies found a positive correlation between PeerWise participation and examination scores (Bottomley & Denny, 2011; Rhodes, 2013; Mcqueen et al., 2014; Mckenzie & Roodenburg, 2017) and its benefits are greater for students of lower ability (Hardy et al., 2014; Mcqueen et al., 2014) which could be attributed to the peer feedback and comments that encouraged active discussion. Previous study also suggested that students were able to develop high-quality questions and were able to rate the quality of questions created by others (Bottomley and Denny, 2011). Furthermore, students generally report a positive attitude towards learning via PeerWise as they found it enjoyable and engaging tool, easy to use, and helpful for revision before the examination (Mckenzie & Roodenburg, 2017; Singh 2014).

With regards to the evaluation of PeerWise in physiology courses, available data seem to be limited and incomplete with the current existing studies focused primarily on the association with academic performance and student perceptions. Therefore, the main focus of this review is to report comprehensively on the role of PeerWise in supporting student’s independent learning in physiology by further investigating the study characteristics, study designs, method of assessments, study limitations, and ultimately its effects on cognitive and affective performances. The review is based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The scope of this review is limited to empirical studies that utilized PeerWise as part of the blended learning in the subjects of physiology and pathophysiology published between the years of 2010 to 2020. This review aims to explore the effectiveness of PeerWise to enhance student learning in physiology by looking into the (i) current trend/pattern of PeerWise usage in physiology curriculum; (ii) association between PeerWise and academic performance; (iii) level of student engagement, (iv) quality of questions created, and (v) student perceptions on the value of PeerWise.

To the best of the author’s knowledge, this is the first systematic review that provides qualitative and quantitative data relating to the association between PeerWise and academic performance, the impact of its usage on student engagement, students’ perceptions on its value as well as the assessment on the methodological quality and risk of bias in individual studies. Moreover, this review addresses the research gaps, limitations, and recommendations from the selected studies to serve as a guide that can be used to improve future work in the topic area. As such, this review will be beneficial for new authors to refine their knowledge and develop new research ideas on their subject area of interest. Overall, it is hoped to add to the existing PeerWise body of knowledge and fills the gap in the current literature.

Method

The Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines was adopted throughout this review to provide a robust and comprehensive framework of a systematic review (Moher et al., 2015). Ethical approval for this study was obtained from the Institutional Review Board of Riyadh Elm University (FRP/2020/211/109/108).

Eligibility Criteria

i. All original studies investigating the effectiveness of PeerWise in physiology education and its associated students’ perceptions.
ii. Publication years between 2010 to April 2020.
iii. All articles that include physiology or pathophysiology courses/topics.
iv. Articles in English language.
v. Articles being an empirical study.
vi. Articles published in scholarly peer-reviewed journals.
Information Sources

Studies were identified by searching relevant papers between years 2010 to April 2020 via following electronic databases; Elsevier, Springer, Pubmed Central, Science Direct, Scopus, Clinical Key, Physiology.org, Wiley, SAGE, Google Scholar, Research Gate, Academia.edu, NCBI website, BMJ, IJCMR.com, Semanticscholar.org, Australasian Journal of Educational Technology (AJET), Journal of Prospectus in Applied Academic Practice (JPAAP), tandfonline.com and MedEd Publish. Other additional relevant studies were further hand-searched from the reference lists of retrieved studies.

Search

The search in the databases was carried out using the following keywords search strategy: “PeerWise + Physiology” or “PeerWise + Effectiveness” or “Physiology + Peer-base Assessments” or “PeerWise + Students’ Perceptions” or “PeerWise + Learning Physiology.”

Study Selection & Data Collection Process

Each title and abstract was first screened to meet the inclusion criteria. Following that, the full-text articles were further accessed for eligibility. Data from studies that meet all the eligibility criteria were recorded and analyzed in the excel spreadsheet (see Figure 1).

![Figure 1. PRISMA Flow Diagram of the Study Selection Process](image)

Data Items

Data extracted from each study include: (i) the demographic background of the studies, (ii) assessment methods, (iii) pattern of Peerwise usage; (iv) the association between PeerWise activity and academic performance (v) the level of student engagement, (vi) quality of questions, and (vii) students’ perceptions on the value of PeerWise.
Risk of Bias in Individual Studies

Evaluation of the risk of bias in the individual study was based on Cochrane’s Collaboration’s tool (Green & Higgins, 2011). The following risk of biases was analyzed: (i) sampling bias (ii) performance bias resulting from allocated interventions during the study (iii) attrition bias due to handling of incomplete outcome data (vi) reporting bias resulting from selective outcome reporting (v) measurement bias due to inappropriate data measurement with non-validated criteria and (vi) bias in the analysis due to not reporting the necessary statistical coefficient related to the study.

Results

Study Selection

A total of 286 studies were identified for this review; 281 via electronic databases and five articles were retrieved from the reference lists. The screening phase involved the examination of research titles and abstracts of all studies identified. A total of 243 studies were excluded as they did not meet the inclusion criteria. This screening resulted in 43 studies selected for the eligibility phase. However, a total of 35 studies were excluded due to paper-based MCQs (N=5), non-PeerWise/other online quizzes (N=6), and non-physiology courses (N=24). Therefore, eight empirical studies fully met the eligibility criteria for inclusion in this systematic review.

Background Characteristics of Reviewed Articles

The background characteristics of the reviewed articles are presented in Table 1. The fields of study of the students involved are predominantly medicine (N = 3), followed by Pharmacy (N=2), Optometry (N=1), Biomedical Science (N=1), and Human Physiology (N=1). Due to limited published articles that specifically discussed the impact of PeerWise activity on physiology learning, this review includes, but not limited to, studies that were conducted among medical students that were known to include physiology subject during pre-clinical years (Kadir et al., 2014; Pathak & Aye, 2015; Walsh et al., 2017), studies that included physiology or pathophysiology topics (Tatachar et al., 2016; Poot et al., 2017; Tatachar & Kominski, 2017; Woods & Lotfus, 2018) and study related to physiology laboratory (Acosta et al., 2018). The selected articles were published between the years 2014 to 2018. Most of the study designs were cross-sectional quantitative and qualitative studies. The quantitative studies were looking at the pattern of PeerWise usage, exploring the association between PeerWise activity and students’ academic performance, and determining the quality of the questions created. Six studies involved qualitative research by utilizing a questionnaire or focus group discussions on evaluating students’ perceptions of PeerWise. Only one study thoroughly discussed the impact of PeerWise on the level of student engagement. Regarding the geographical characteristics of the studies included, two studies were from the United States, two were from Malaysia, two from the United Kingdom, one from Australia, and one from the Netherlands. Regarding the sample size, a total, 3353 participants were recorded. Some studies did not report the gender distribution of the participants. The three studies that disclosed the gender distribution of the participants showed more female participants (>70%) compared to males. Most of the participants were students from Year 1 and Year 2 enrolled in physiology and pathophysiology subjects from different academic programs. The mean age range of the participants was 20 years old.

The Pattern of PeerWise Usage

The overall usage of PeerWise by students was reported in five articles. Table 2 shows the total number of MCQs created, answered, and commented by the students. On average, medical students showed a high level of PeerWise usage by showing a high number of questions created as demonstrated by Pathak & Aye (2015) (N=258), Kadir et al. (2014) (N=460) and Walsh et al. (2017) (N=4671) in comparison to Biomedical students by Poot et al. (2017) (N=59) and Introductory Human Physiology students by Woods & Lotfus (2018) (mean of 8 questions per team). Regarding the number of questions answered by students, Walsh et al. (2017) reported the highest number with 606 658, which reflects on a large number of students who participated in this study, followed by Pathak & Aye (2015) with 2739 and Poot et al. (2017) with 1776. Woods & Lotfus (2018) found an increased mean number of MCQs answered among students that were introduced to PeerWise (N=75.99) in comparison to the control group (N=58.45). Three studies reported on the number of comments written by students. Walsh et al. (2017) recorded a total of 7735 comments among the two different cohorts, while Pathak
& Aye (2015) reported 861 comments. A study done by Kadir et al. (2014) showed that some students wrote more than 40 comments, but some students wrote less than ten comments during the running of the course. Moreover, two studies (Kadir et al., 2014; Pathak & Aye, 2015) found that the maximum number of questions were generated at the beginning of the course; second week and third week, respectively, and decreased towards the end of the course as students were busy preparing for their exams. In contrast, Walsh et al. (2017) reported that questions writing frequency increased around the examinations.

### Table 1. Background Characteristics of Reviewed Articles

<table>
<thead>
<tr>
<th>First Author (Year)</th>
<th>Country of Origin</th>
<th>Sample Size</th>
<th>Gender Distribution</th>
<th>Age Range</th>
<th>Sample Characteristics</th>
<th>Study Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kadir et al. (2014)</td>
<td>Malaysia</td>
<td>120</td>
<td>77% female</td>
<td>Mean age 20 years old</td>
<td>Year 2 Medical students, Nervous System Course</td>
<td>Cross-sectional quantitative and qualitative study</td>
</tr>
<tr>
<td>Pathak &amp; Aye (2015)</td>
<td>Malaysia</td>
<td>79</td>
<td>-</td>
<td>-</td>
<td>Year 1 Medical students, Physiology Course</td>
<td>Quantitative study</td>
</tr>
<tr>
<td>Tatachar et al. (2016)</td>
<td>United States</td>
<td>76</td>
<td>-</td>
<td>-</td>
<td>Year 2 Pharmacy students, Integrated Pharmacotherapy Module (End-Stage Renal Disease)</td>
<td>Qualitative study</td>
</tr>
<tr>
<td>Tatachar &amp; Kominski (2017)</td>
<td>United States</td>
<td>84</td>
<td>-</td>
<td>-</td>
<td>Year 2 Pharmacy students, Integrated Pharmacotherapy Renal Course (Chronic kidney disease-mineral bone disorder)</td>
<td>Quantitative &amp; qualitative study</td>
</tr>
<tr>
<td>Poot et al. (2017)</td>
<td>The Netherlands</td>
<td>109</td>
<td>37 males and 72 females</td>
<td>Mean age 20 years old</td>
<td>Year 2 Biomedical students, Physiology Course</td>
<td>Quantitative &amp; qualitative study</td>
</tr>
<tr>
<td>Walsh et al. (2017)</td>
<td>United Kingdom</td>
<td>603</td>
<td>-</td>
<td>-</td>
<td>Year 1 and Year 2 Medical students</td>
<td>Quantitative and qualitative study</td>
</tr>
<tr>
<td>Woods &amp; Lotfus (2018)</td>
<td>United Kingdom</td>
<td>2170</td>
<td>-</td>
<td>-</td>
<td>Year 1 and Year 2, Introductory Human Physiology Course</td>
<td>Cross-sectional quantitative study</td>
</tr>
<tr>
<td>Acosta et al. (2018)</td>
<td>Australia</td>
<td>112</td>
<td>70% female</td>
<td>Age 18–23 years old</td>
<td>Year 2 and Year 3, Ocular Anatomy and Physiology course of the Bptom degree</td>
<td>Cross-sectional quantitative and qualitative study</td>
</tr>
</tbody>
</table>

Among the eight articles reviewed, only three papers briefly discussed the quality of the questions created by the students (see Table 2). In general, these studies reported that students are capable of designing questions that met the minimum criteria. However, there were apparent differences in the findings on the level of the questions. Two studies evaluated the cognitive level of MCQs created via Bloom’s taxonomy framework. Kadir et al. (2014) and Poot et al. (2017) found that the majority of the MCQs created by the students were on the two lowest levels of Bloom’s taxonomy; Level 1 – Remembering and Level 2 – Understanding. One-third of the lower order thinking questions were a pure factual type of question (Poot et al., 2017) and students were unable to construct case-based scenario questions in addition to poor English language skills (Kadir et al., 2014). Walsh et al. (2017) evaluated the quality of the students’ questions via discriminative index and difficulty index. The findings reflected on the excellent quality questions with high difficulty levels created by the students in which the mean discriminative index for the top 100 most answered questions was between the range of 0.446 to 0.485 and the mean difficulty index was between the range of 37% to 43%. Discrimination index is a measure on the ability of a question item to differentiate between „good” students and „poor” students based on how well they know the question being tested and discrimination index of ≥ 0.40 is considered as very good items. With regard to the difficulty index, it is the percentage of students who answer an item correctly and it ranges from 0 to 100; the higher the value, the easier the question (Rahim, 2010).
Table 2. Pattern of PeerWise Usage and Quality of Questions

<table>
<thead>
<tr>
<th>Studies</th>
<th>MCQs Written (N)</th>
<th>MCQs Answered (N)</th>
<th>Number of Comments Made (N)</th>
<th>Quality of Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathak &amp; Aye (2015)</td>
<td>258</td>
<td>2739</td>
<td>861</td>
<td>-</td>
</tr>
</tbody>
</table>
| Kadir et al. (2014)      | 460              | <10 to >400 per student | <10 to >40 per student    | • The questions were generally rated highly by students.  
• The instructor rated more questions in the “fair” and “good” categories.  
• Lower levels of Bloom’s Taxonomy (Level 1 & 2).  
• Language: Poor  
• Unable to construct case-based scenario questions. |
| Poot et al. (2017)       | 59               | 1776              | -                          | • 11 out of 59 questions were labelled as higher order thinking questions based on Bloom’s taxonomy.  
• 1/3 of the lower order thinking questions were pure factual recall (in the format of a true–false question).  
• Majority of the lower order questions were classified in the understanding level of Bloom’s. |
| Woods & Lotfus (2018)    |                  |                   |                            | -                    |
| Control Year 1 & 2 = Mean 7.6 | Control Year 1 & 2 = Mean 58.45 | Exp Year 1 = Mean 8.2 | Exp Year 1 & 2 = Mean 75.99 | - |
| Walsh et al. (2017)      | 4671             | 606 658           | 7735                       | • The mean discriminative index for the top 100 most answered questions for each academic year were: 0.485 (2013–2014, year 1), 0.446 (2014–2015, year 2) and 0.480 (2014–2015, year 1).  
• The mean difficulty index in the three groups was 0.370 (2013–2014, year 1) 0.438 (2014–2015, year 2) and 0.362 (2014–2015, year 1). |

The Association between PeerWise and Academic Achievement

The relation between PeerWise activity and academic achievement was reported in six articles (see Table 3). In four studies (Kadir et al., 2014; Poot et al., 2017; Tatachar & Kominski, 2017; Walsh et al., 2017), the authors used final marks on the summative test as a measure for academic performance, while weekly assessments and total marks in laboratory test were used in studies done by Pathak & Aye (2015) and Acosta et al. (2018), respectively. Three studies (Kadir et al., 2014; Poot et al., 2017; Walsh et al., 2017) reported that student’s involvement in PeerWise leads to better performance on their examination scores. Moreover, Walsh et al. (2017) reported that the mean summative scores increased as question writing frequency increased. Similarly, Poot et al. (2017) revealed a significant difference in students’ achievement between motivated (engaged in PeerWise) and non-motivated students (not engaged in PeerWise) in the second summative exam after the introduction of PeerWise in comparison to results of the first summative exam. A negative relationship was
observed between students authoring questions and examination scores in three studies (Pathak & Aye, 2015; Tatachar & Kominski, 2017; Acosta et al., 2018). Tatachar & Kominski (2017) was the only study that experimented with intervention before and after the use of PeerWise. The researchers found no significant difference between pre and post-test scores.

The Impact of PeerWise on Student Engagement in Learning Physiology

A quantitative analysis on increase students’ engagement through game and competition via PeerWise was done by Woods & Lotfus (2018) (see Table 3). This study analyzed the total number of questions created, the total number of questions answered, and the total number of badges earned by the students in PeerWise between the control groups and competing groups (experimental years) in two different cohorts. The findings of this study revealed that students were more productive when they were competing and playing a game as shown by higher contribution in PeerWise among students in the competing groups with significantly higher badges compared to the control groups. Furthermore, this study highlighted that the students showed better attendance in tutorials with the incorporation of PeerWise into their learning activities.

Table 3. The Association Between PeerWise with Academic Achievement and Student Engagement in Learning Physiology

<table>
<thead>
<tr>
<th>Studies</th>
<th>Assessment Methods</th>
<th>Level of PeerWise Activity</th>
<th>Findings on Academic Achievement</th>
</tr>
</thead>
</table>
| Pathak & Aye (2015) | • This study evaluates academic performance in physiology among medical students.  
  • Students to author, explain their own MCQs, answer, evaluate, and discuss MCQs written by their peers via PeerWise for 6 months.  
  • Academic performance was measured via weekly assessment marks. | 64 students (81%) participation rate | No correlation between PeerWise activity and academic performance |
| Kadir et al. (2014) | • This study evaluates academic performance in Nervous System course that includes anatomy, physiology, pathology and pharmacology topics.  
  • Students were asked to create at least two MCQs for the duration of 5 weeks.  
  • Academic performance was measured via End-of-course marks. | A total number of 120 out of 124 (96.7%) participation rate | Significant correlations between PeerWise scores and the End of Course marks. |
| Poot et al. (2017) | • The study evaluates academic performance in physiology-oriented topics: respiratory, circulatory, and urinary organ systems (each part had a duration of three weeks).  
  • Students were free to develop questions on the 14 core topics of the course and they could practice with the questions constructed by their peers for the full three weeks before their small summative test.  
  • This study was conducted within 10 weeks.  
  • Academic performance after the intervention was measured via | 45 out of 109 (41%) participation rate | Significant difference on academic achievement in the motivated group (i.e. students who participated in PeerWise). |
students’ grades on the second summative test in the course. Students’ grades on the first summative test were used as a covariate.

**Tatachar & Kominski (2017)**

- This study evaluates students’ performance in pathophysiology topics, “Chronic Kidney Disease (CKD)-Mineral Bone Disease (MBD)”, with interventions via PeerWise and Case-Based Application.
- PeerWise team was tasked to create 2 MCQs related to CKD via PeerWise, answer all questions submitted by other teams, rated questions submitted by other teams & make constructive comments on at least one question written by their peers.
- Pre and post-test questions on CKD-MBD were conducted before and after the interventions.
- Academic performance was measured via total block exam scores and final exam performance.

**Walsh et al. (2017)**

- This study evaluates the usage of PeerWise among pre-clinical medical students for the period of 6 months.
- Mean raw scores over two summative assessments were converted to percentages and correlated with question writing, answering and commenting frequency by Spearman’s Rank correlation coefficient.

| Writing | Prolific users ≥50 |
| Answering | Prolific users ≥1000 |
| Commenting | Prolific users ≥50 |

- Significant correlations between writing, answering and commenting frequency with summative examination performance.
- For question writing, mean summative score increased as question writing frequency increased.

**Acosta et al. (2018)**

- This study compares students’ performance between the usage of static website & interactive website for anatomy and physiology laboratory classes in Optometry.
- Static website contained pictures and videos of laboratory activities, and interactive website contained interactive diagrams, videos, practise questions and PeerWise.
- Students were requested to use PeerWise to develop their own questions and answers for the laboratory topics.
- Academic performance was measured via the average total

- The interactive laboratory website was not a significant contributor to improvement in the total marks achieved in the laboratory test.
- Users of the static website had a significantly lower mark than those with access to the interactive website.
mark and distribution of marks in the laboratory test.

<table>
<thead>
<tr>
<th>Woods &amp; Lotfus (2018)</th>
<th>This study evaluates students’ engagement with the use of competition in a large Human Physiology Course.</th>
<th>100% participation rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>There were 4 assignments assigned using the website PeerWise. For each assignment, they created and posted 2 MCQs, answer 5 questions and provide feedback on those questions authored by their peers.</td>
<td>Higher total number of questions created &amp; questions answered in PeerWise among students of competing years compared to the control years.</td>
</tr>
<tr>
<td></td>
<td>Students had to write 8 questions, then answer and comment on 20 questions to earn full grades for the assignments.</td>
<td>Students of competing years earned more badges compared to the control years.</td>
</tr>
<tr>
<td></td>
<td>They were awarded a total of 4% for these assignments.</td>
<td>Students are more productive when they are playing a game.</td>
</tr>
<tr>
<td></td>
<td>The total number of questions posted, answered and total number of badges earned by the students in PeerWise were analysed between the control groups and competing groups in 2 different cohorts.</td>
<td>Students showed better attendance in tutorials</td>
</tr>
</tbody>
</table>

**Student Perceptions on the Usage of PeerWise**

The qualitative aspect of the studies was retrieved from six articles (see Table 4). Students’ perceptions of the value of PeerWise as a teaching-learning tool were evaluated. Two studies adopted a focus group survey (Walsh et al., 2017; Acosta et al., 2018) while the other four studies implemented open and close-ended questionnaires (Kadir et al., 2014; Tatachar et al., 2016; Poot et al., 2017; Tatachar & Kominski, 2017). Further analysis of students’ feedback on their experiences with PeerWise revealed four main themes; (i) learning competency, (ii) fun learning experience, (iii) engagement with peers, and (iv) motivation in learning. In all the articles reviewed, a clear majority of the students reacted positively to PeerWise activity.

Students qualitative responses suggested that PeerWise improved learning competency. Students considered developing new questions and answering other students’ questions in PeerWise helped them to reinforce knowledge and identify knowledge gaps (Kadir et al., 2014; Walsh et al., 2017), helped to identify the learning objectives of the course materials (Acosta et al., 2018) and useful for the revision of knowledge with a detailed exploration of learning contents (Poot et al., 2017; Tatachar & Kominski, 2017). PeerWise shows a significant relationship in terms of promoting student engagement. Perceptions from various studies revealed that question creation activity and reading other students’ comments fostered collaboration with peers (Kadir et al., 2014; Poot et al., 2017; Tatachar & Kominski, 2017). Most of them were motivated to develop their questions before practicing on their peers’ questions (Poot et al., 2017). Students were also interested in getting their questions rated by their peers and appreciated the ability to access peer-written questions for exam preparation (Tatachar et al., 2016).

The majority of the studies indicated that the PeerWise tool increased the level of interest in the subject as it is a fun and enjoyable learning tool. Students expressed earning virtual badges and comparing performance with their peers are motivational (Walsh et al., 2017). A focus group survey conducted by Acosta et al. (2018) reported that students preferred the blended-learning system as it provides a new engaging way of interacting and learning with other students. Overall, students reported that they want to use PeerWise again in future courses.
Table 4. Students’ Perceptions of PeerWise

<table>
<thead>
<tr>
<th>Studies</th>
<th>Assessment Methods</th>
<th>Findings on Student Perceptions</th>
</tr>
</thead>
</table>
| Kadir et al. (2014)    | Close-ended questionnaire                 | **Competency:**
|                        |                                          | • Developing new questions improved understanding of the topics. |
|                        |                                          | • Answering other students’ questions helped to learn better. |
|                        |                                          | **Fun learning:** PeerWise is something new and different. |
|                        |                                          | **Engagement with peers:**
|                        |                                          | • Reading other students’ comments improved existing knowledge of the topics. |
|                        |                                          | • Interested to see how other students rated their questions. |
|                        |                                          | **Motivation:** Will use PeerWise again in a future course. |
| Tatachar et al. (2016) | Pre-assessment and post-assessment surveys | **Fun learning:** Students enjoyed the collaboration with team members as well as the challenge and excitement of creating questions. |
|                        |                                          | **Engagement with peers:** Students appreciated having access to peer-written questions for exam preparation. |
|                        |                                          | **Motivation:** Students want to use PeerWise again in future courses. |
| Poot et al. (2017)     | Open questions and self-reported motivating and learning strategies questionnaire | **Competency:** Improved competency. |
|                        |                                          | **Engagement with peers:** Significant relationship in student engagement. |
|                        |                                          | **Motivation:** |
|                        |                                          | • Students wanted to answer the questions in order to increase their competency. |
|                        |                                          | • Most of the students felt strongly that before they could practice the questions of peers, they also had to develop some questions themselves that could help others. |
| Tatachar & Kominski (2017) | Survey questions on student learning, student satisfaction & interest in continuing participation | **Competency:**
|                        |                                          | • The PeerWise group reported higher levels of competence and understanding than the case-based group. |
|                        |                                          | • Student question creation resulted in detailed exploration of content. |
|                        |                                          | • Providing a rationale for alternatives and the correct answer improved understanding of material. |
|                        |                                          | **Fun learning:** Significant differences on enjoyment and level of interest in the subject matter in PeerWise group. |
|                        |                                          | **Engagement with peers:** Student question creation fostered collaboration with peers. |
|                        |                                          | **Motivation:** Greater likelihood of participation in the future. |
| Walsh et al. (2017)    | Focus group & thematic analysis          | **Competency:**
|                        |                                          | • Answering question reinforces knowledge. |
|                        |                                          | • Answering question identifies knowledge gaps. |
|                        |                                          | • Writing questions improves knowledge. |
|                        |                                          | **Fun learning:** Enjoyable and fun. |
|                        |                                          | **Motivation:**
|                        |                                          | • Virtual badges are motivational. |
|                        |                                          | • Comparing performance with peers are motivational. |
| Acosta et al. (2018)   | Survey and focus group                  | **Competency:**
|                        |                                          | • Online quizzes helped students identify learning objectives and self-assess knowledge. |
|                        |                                          | • Usefulness of the material to review knowledge before laboratories. |
|                        |                                          | **Fun learning:** In the focus group, students reported they preferred a blended learning over the website. |

Definition of the themes:

**Competency:** Monitor learning, identify gaps in knowledge, assessment on the level of achievement & obtain feedback on questions (Mckenzie & Roodenburg, 2017).

**Fun learning:** Enjoyable activity (Mckenzie & Roodenburg, 2017).

**Engagement with peers:** Interactive, valuable contributor, stimulate discussion ((Mckenzie & Roodenburg, 2017).

**Motivation:** Empowerment to achieve high levels of performance and overcoming the barriers in order to change (Tohidi & Jabbari, 2012).
Assessment of Risk of Bias in Individual Studies

A general overview of possible risks of bias across all reviewed studies is presented in Table 5. With regards to selection bias, all studies were rated as low risk due to their cross-sectional design with no likelihood of bias resulting from non-random allocation of participants. Performance bias was found to be of high risk across all eight studies as blinding of participants was not attempted in all these studies, thus contributing to potential sources of biases at the outcome levels. In terms of attrition bias, all studies were rated as low risk since no dropouts were reported in any studies, and the handling of incomplete data was done adequately. Reporting bias was rated high risk in one study (Poot et al., 2017) due to the poor reliability of the internal motivational scale used in the study while all other studies were rated as low risk as the reporting of methodology, research outcomes, negative results, and citations were done adequately. Assessment of measurement bias was rated as high risk in two studies (Kadir et al., 2014; Acosta et al., 2018) due to involving leading questions in their surveys on students’ perceptions, and the other six studies (Pathak & Aye, 2015; Tatachar et al., 2016; Walsh et al., 2017; Poot et al., 2017; Tatachar & Kominski, 2017; Woods & Lotfus, 2018) were rated as low risk. Finally, the bias in the analysis was rated as low risk in all the studies reviewed as all the essential statistical results were reported sufficiently.

<table>
<thead>
<tr>
<th>Studies</th>
<th>Sampling Bias</th>
<th>Performance Bias</th>
<th>Attrition Bias</th>
<th>Reporting Bias</th>
<th>Measurement Bias</th>
<th>Bias in Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kadir et al. (2014)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Pathak &amp; Aye (2015)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tatachar et al. (2016)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tatachar &amp; Kominski (2017)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Poot et al. (2017)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Walsh et al. (2017)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Woods &amp; Lotfus (2018)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Acosta et al. (2018)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes: High risk of bias = +; Low risk of bias = -

Discussion

Integrated learning in physiology is crucial to improve the quality of students and to produce healthcare professionals who can provide adequate diagnosis and better treatment to the patients. The traditional face-to-face lecture is a fully teacher-centered process with a one-way communication teaching method where no active learning is incorporated, and students are passive listeners. There are also no exposures to critical thinking and students facing difficulty in correlating and applying pre-clinical subjects with clinical practice. In some cases, experts are not always good teachers, and unnecessary repetition and confusion might occur (Vijaya, 2010). Blended learning is widely defined as a way of learning that combines traditional classroom lessons with computer-mediated instruction and delivered over the internet (Cambridge, 2013). It has been proven to potentially enhance both the efficiency and effectiveness of the learning experience (Garrison & Kanuka, 2004).

The present systematic review aimed to identify relevant empirical evidence on the impact of blended learning via PeerWise on students learning experience in the subject of physiology. The primary data extracted from the studies reviewed encompassed the geographical distributions of the studies and key characteristics of the participants, methodological features and risk of bias in individual studies, the pattern of PeerWise usage, the effects of constructing, answering, and evaluating MCQs on academic achievement, students engagement and their perceptions on the value of the tool. As the studies varied significantly in terms of methodologies, a direct
comparison of the data was not possible. However, consistent findings were reported across the studies that PeerWise has significantly improved learning experience, engagement with peers, and examination scores.

In terms of the background characteristics of the study samples, this review demonstrated that most of the studies were conducted among the developed countries which imply that the application of blended learning could be closely associated with the technological advancement of the country. Ololube (2013) suggested that the transformation from the traditional face-to-face teaching to a blended learning approach is heavily dependent on the use of technologies and the lack of economic support in developing countries has hindered its educational growth. From the review, it was also noted that the majority of the participants were medical students which proves that the blended learning teaching method has gained its popularity as increasing numbers of medical colleges use the internet as a forum for teaching and learning (Thakore & McMahon, 2006). Given the present limited findings, future research on PeerWise particularly in physiology must include (i) more representative samples from developing countries such as South-East Asia and Middle-Eastern countries, (ii) comparison results of participants from different years of study and (iii) participants from different health sciences programs that include physiology subject. This may help increase the study’s validity of findings from different populations.

Regarding the methodological features of the studies reviewed, it was evident that most of the studies adopted cross-sectional designs rather than a more complex and costly design, such as experimental research. The way researchers measure the association between level of PeerWise usage and academic achievement in their research was either via students’ grades in weekly assessments or final exam scores. Attention should be given on standardizing the assessment criteria by adopting an experimental study design with intervention before and after the use of PeerWise. The use of extensive inconsistent and non-validated criteria may render additional methodological difficulties potentially compromising the advancement to produce more reliable cross-cultural research. In terms of the survey questions on student perceptions, majority of the studies did not explain on the reliability and validity of their survey questions of which may attribute to poor reliability of responses. It was also reported that most of the studies used paper-and-pencil technique rather than online questionnaires as they found it challenging to get students to respond to the online survey. A focus group survey was also adopted in some studies which to its advantage can encourage good discussions with more in-depth answers, but it can be time-consuming, costly, and can generate qualitative data that is harder to analyze.

One of the main objectives of the present review was to explore the usage of PeerWise in the current physiology curriculum. Results demonstrated a high level of participation among students; even no mandatory participation was imposed on them. The driving factor is probably due to the coursework marks that the students earned according to their level of participation, as suggested by Paterson et al. (2011). Besides, the interactivity element of PeerWise (virtual rewards such as leaderboards, badges, level, and points) (Nicholson, 2012) managed to attract students’ interest to participate and stay engaged in their learning (Denny, 2013) and changed the tedious learning method of physiology into a fun and enjoyable experience. Enjoyment has been linked to engagement in study and learning, to say the least (Goetz et al., 2006). This review highlighted that students were encouraged to be more productive in learning when they were competing with their peers through PeerWise social gaming elements as reflected by better attendance in tutorial sessions with higher motivation to achieve their goal to win the year-end prize. Other studies have shown similar findings that engagement in learning is a positive predictor of the quality of learning, grades, and results in exams, a positive predictor of regular school attendance, and successful school graduation (Fredricks, Blumenfeld & Paris, 2004; Appleton, Christenson & Furlong, 2008; Salmela-Aro & Upadyaya, 2013). High student engagement in learning was also reported when students were actively involved in constructing MCQs compared with passively answering MCQs (Bottomley & Denny, 2011; Singh, 2014).

The articles reviewed in this paper showed that students preferred to contribute more questions than was expected of them. Creating questions was perceived to help remember and apply knowledge particularly in a physiology subject that contains a lot of difficult concepts. A deeper understanding of the structures and functions of the physiological content can be achieved through question creation as it enables and enhances critical thinking and reasoning skills of students. Various studies have shown the positive impact of interactive approaches to teaching and learning which helps in understanding difficult concepts (Rodriguez-Barbero & Lopez-Novoa 2008).

It was also observed in this review that students answered more questions than they constructed. Students reported that they enjoyed answering questions as they felt they were learning more through question-answering in comparison to question creation. The use of answering questions for learning is supported by a various body of evidence suggesting that repeated practice is effective for enhancing learning (Larsen 2015; Pan 2016),...
helpful for revision and knowledge consolidation (Rhodes 2013; McQueen et al. 2014), and facilitate a higher level of cognitive skills and promotes deeper engagement in learning.

Collaboration through making comments and evaluating peers’ contributions was perceived to be valuable and equally helpful. Students reported that they obtained a considerable amount of knowledge by sharing ideas and discussing it with their peers (Oakley et al., 2004). This is particularly important in learning physiology as the online discussion of questions through PeerWise may help students with serious misconceptions about physiological phenomena. It has been proven that peer-level teaching and learning can have a powerful impact on student performance and clinical knowledge (Han, Chung & Nam, 2015; Seifert et al., 2015). Students who are not high performers may benefit from it via engagement with their peers.

The frequency patterns of PeerWise usage among the participants demonstrated in this review were reported to be high at the beginning of the course as well as during the examination period. This could be due to the multiple deadlines, reminders, and email notifications sent by the instructors/teachers, which may lead students to revisit the system more often (Hakulinen & Korhonen, 2010). In addition, students perceived it as a reliable and an efficient revision tool as answering questions is considered to be helpful for knowledge consolidation and highly associated with memory retention (Sykes, Denny & Nicolson, 2011; Rhodes, 2013; Mcqueen et al., 2014).

The present findings regarding the association between PeerWise activity and academic achievement found in this review are suggestive that it helps to improve examination scores. This is consistent with other studies that reported a positive correlation between student activities on PeerWise and examination performance (Denny, Luxton-Reilly & Hamer 2008; McQueen et al. 2014). It has been shown that questions writing frequency have the strongest correlation with summative performance. This could be explained by the fact that writing questions especially targeting the higher order of Bloom’s Taxonomy is challenging which requires a deeper approach to learning as the author needs to appraise the question critically. In relation to physiology, it can promote a deeper understanding of the structures and functions of the physiological content as this subject requires critical thinking and reasoning skill of students.

Although authoring and answering questions theoretically contribute to knowledge consolidation, this review observed two studies that did not show a positive association with academic performance. These less positive outcomes could be explained by the fact that MCQ writing could be a complex task for students especially when the requirements by the instructors are high as such to include questions with high level of content knowledge, problem-solving and content integration where they need greater instructional support from the teachers (Leppink & Duvivier, 2016). Students might also feel overburdened with the activities of authoring and answering questions together with giving feedback to their peers particularly during the running of a course with heavy content such as physiology and pathology (Smith et al., 2020). The time required to generate a high-quality question could be another discouraging factor for the students as they might take three to four times longer than the time required for reading a text (Hoogerheide et al., 2019). It is possible that with fewer requirements from the instructors, students might show more positive responses towards the usage of PeerWise.

**Limitations and Recommendations**

Several types of limitations were identified across the reviewed studies. The limitations found can be clustered within three major categories: (i) operationalization and measurement issues, (ii) question quality issues, and (iii) language issues. Operationalization and measurement issues found within the reviewed articles involved problems related to (a) short duration of the studies, (b) no mandatory requirement for participation, (c) poor reliability of the survey scale, (d) insignificant differences between two methods in measuring student perceptions, (e) large differences in the number of questions developed and answered, and (f) students did not respond to the online survey. Questions quality issues involved (a) poor quality of questions and (b) no guidance from lecturers or facilitators on constructing questions. Poor English language skills limiting the ability to construct high-quality questions and questions based on clinical scenarios.

Different methodological features and measurement tools to assess the association between PeerWise activity and students’ academic performance adopted by the researchers across all the studies reviewed may constitute further complication, therefore future research should carry out research using a more unified and consensual measure to assess the relationship between the two factors, to allow replication of the study design and produce reliable cross-cultural research findings. Application of PeerWise as part of a course can be improved by guiding the students on creating MCQs with reference to different levels of Bloom’s Taxonomy. Teachers can
also make PeerWise participation mandatory to all students and encourage them to use it regularly to familiarize students with the system. It is also recommended to add teacher-guided discussion at the end of PeerWise activity to clear any confusion that may arise.

Conclusion

In conclusion, the present findings elucidate the usefulness of including PeerWise as part of the blended learning methodology in Physiology rather than relying solely on the traditional didactic teaching method. By shifting focus towards the development of more student-centered and engaging learning experiences, students can overcome the difficulties they have in learning physiology via collaborating with peers in creating, answering, and reviewing questions. Apart from being a valuable learning tool for students, PeerWise offers additional benefits for instructors as it can be an online repository of MCQs, that are readily available to instructors to use or modify for future examinations. A large body of question banks can be collected with a long-term continuation of such student MCQs creation exercise. In advantage, this can reduce the time and effort required by instructors to construct new exam questions. Future research may improve the understanding of PeerWise usage and benefits by evaluating the impact of creating and answering questions on memory retention and students’ motivation in learning, and take into account the present limitations and recommendations discussed in this review to enhance the quality of published studies in the field of PeerWise.

References


Khashaba


**Author Information**

**Ahmed Sayed Khashaba**
Associate Professor of Medical Physiology
Department of Basic Sciences
Riyadh Elm University
Riyadh 12734
Kingdom of Saudi Arabia
Contact e-mail: ahmedkhashaba@riyadh.edu.sa