

---

# Firecracker Course Integration Case Study

---



**Alan Rozet**  
Data Scientist, Firecracker  
September 16, 2016



---

## Introduction

Firecracker is an adaptive learning platform that uses proven learning and memory science to help medical students simultaneously prepare for their classes and board exams. We work with schools to identify struggling students and difficult concepts early, strengthening weak areas over time (instead of cramming just before an exam). This is made possible through a spaced repetition system that adaptively recommends focused daily review assignments to prepare students for their courses and licensing exams. Finally, our practice tests pinpoint precise areas of weakness, and our content can be aligned with medical school courses to allow for flexible targeting of subject material.

## Summary

- Firecracker can identify lower-performing students and topics, demonstrated in this study by measuring the difference in average flashcard review score from a national aggregate.
- The number of flashcards reviewed on Firecracker and final flashcard review score are both significantly correlated with increased performance on curriculum-aligned weekly quizzes and cumulative exams.
- Student performance on curriculum-aligned Firecracker-authored weekly quizzes are correlated with school-authored summative exams.

## Background

Firecracker provides topic notes on key medical concepts, reinforced using flashcard questions, and tested using clinical case questions and practice examinations. The chart on the following page shows the sequence a student will follow as they advance towards mastery of a concept. The diagrams afterward visualize student flashcard review scores, weekly quiz scores, and cumulative final exam scores from a curriculum-aligned partner school with a customized mix of Firecracker-authored and school-authored questions. Questions presented in earlier weeks recurred in later ones to solidify concepts. Students had full access to 16,000 Firecracker flashcard review questions, and could flexibly take each weekly quiz from Friday to Monday, at a time that suited their schedules.



#### Topic Concept



**Cardiac Anatomy**

The most Posterior chamber of the heart is the left atrium, which is anterior to the esophagus. Thus, enlargement of the left atrium can cause:

- Dysphagia by compressing the esophagus
- Hoarseness by compressing the left recurrent laryngeal nerve

The card also features a diagram of the heart with anatomical labels and a list of 10 empty circles for notes on the left side.

#### Flashcard Question



Q: What is the most posterior chamber of the heart?

Below the question are five colored buttons labeled 1 through 5, representing a Likert scale for rating knowledge.

#### Test Style Question



A 67-year-old man presents to his primary care physician complaining of weight loss. Upon further questioning, he notes that he has experienced moderate dysphagia (difficulty swallowing) over the last year. The patient has a history of congestive heart failure and his last echocardiogram revealed an ejection fraction of 35% (normal > 55%). The physician suspects that the patient's weight loss may be a result of enlargement of the heart secondary to congestive heart failure.

Which of the following changes is the most likely cause of this patient's dysphagia?

- A. Left ventricle enlargement
- B. Left atrial enlargement
- C. Right ventricle enlargement
- D. Ascending aortic aneurysm
- E. Right atrial enlargement

## Student and Topic Analysis

It is better to identify and assist struggling students early on. Waiting until exams to identify struggling students necessitates helping them catch up through lengthy and costly summer or school-year remediation programs, or via one-on-one tutoring.

Through Firecracker schools can identify difficult topics, using student performance on active flashcard review questions. After viewing a question or fact presented on a flashcard, students have the opportunity to rate their knowledge on a 1-5 Likert scale. By aggregating students' first-time flashcard review scores, we can gain insight on their comfort with topics compared to students' at other schools. The set of graphs below depicts a school's first-time flashcard reviews for their top three reviewed topics; the graph in the center shows a lot of students rating their understanding of Toxicology as a "1," or "low understanding," indicating a need for content revision.





## Curriculum Alignment and Firecracker Assessments

Firecracker’s content is rich with citations to popular medical textbooks such as Costanzo’s Physiology and review books such as Goljan’s Pathoma and McGraw Hill’s First Aid. The content includes useful images, diagrams, and mnemonics to help students learn. Once a school’s learning objectives are aligned with Firecracker’s content, schools can measure student performance on weekly quizzes, while examining a host of item-level statistics to hone in on complex topics or poorly formulated questions. The table below demonstrates some of the metrics that Firecracker generates on a per-item basis<sup>1</sup>.

Topic Name	Vignette body	Question body	Correct Answer	Position In Exam	Point Biserial Correlation	Discrimination Index	Difficulty Index	Questions In Topic
Carbohydrates	SGLT1 channels in intestinal epithelial cells ...	Glucose entering intestinal epithelial cells t...	Indirect active transport	23	0.458	0.5000	0.74	1
Epithelial Cell Junctions	Tight junctions are present in skin and epithe...	Which type of membrane protein is essential fo...	Claudins	13	0.280	0.1250	0.96	3
Epithelial Cell Junctions		Which of the following molecules are involved ...	Cadherins	15	0.338	0.0625	0.98	3

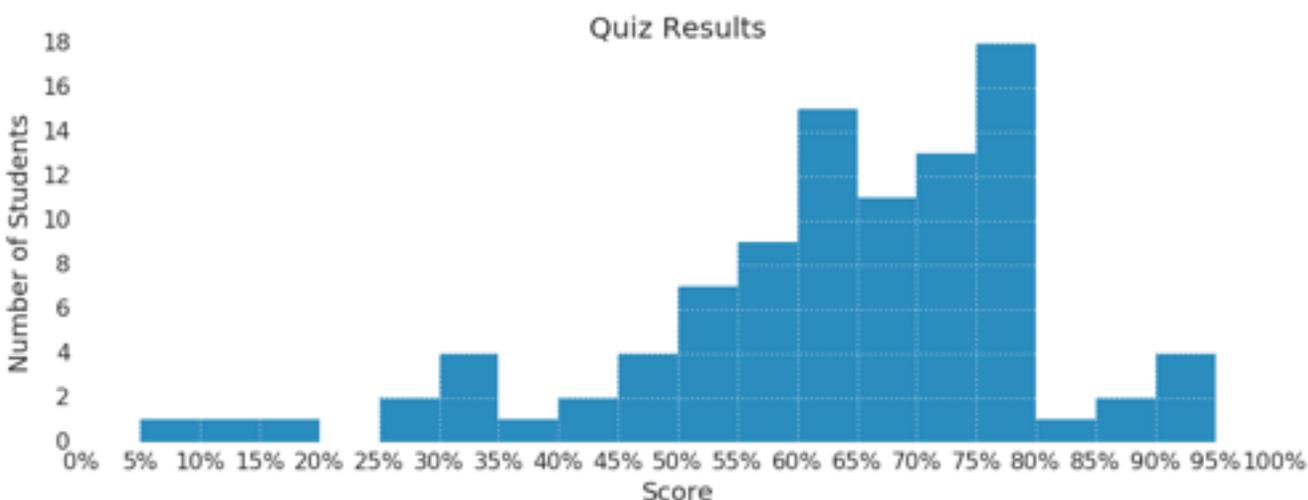
<sup>1</sup> Vignettes are descriptive bodies of text associated with questions. The point biserial correlation is the correlation between student scores on a given question with their overall test score, and is used as a marker of item reliability. The discrimination index is a measure of question difficulty; it is calculated by creating groups of highest and lowest performing students, and dividing the difference in the two groups’ performance on a given question by the size of each group. Difficulty index is the number of students who got a question correct divided by the total number of students.



Using Firecracker, faculty can examine student effort week by week. The table below shows a subset of useful metrics that faculty members at our partner school used to uncover student-specific trends in performance, study habits, test-taking behaviors, engagement with the material, etc.

Number Of Flashcard Reviews	Score	Normalized Score	Quiz Started At	Quiz Duration	Number Of Pauses	Total Pause Duration
1164.0	88.889	89.918	2016-07-31 18:51:20.010969	00:33:35.580889	1	00:00:06.053000
1557.0	81.481	80.905	2016-07-30 10:27:47.188654	00:32:02.516719	1	00:04:51.641000
378.0	100.000	103.438	2016-07-31 12:48:31.423918	00:38:16.464170	15	00:18:29.563000

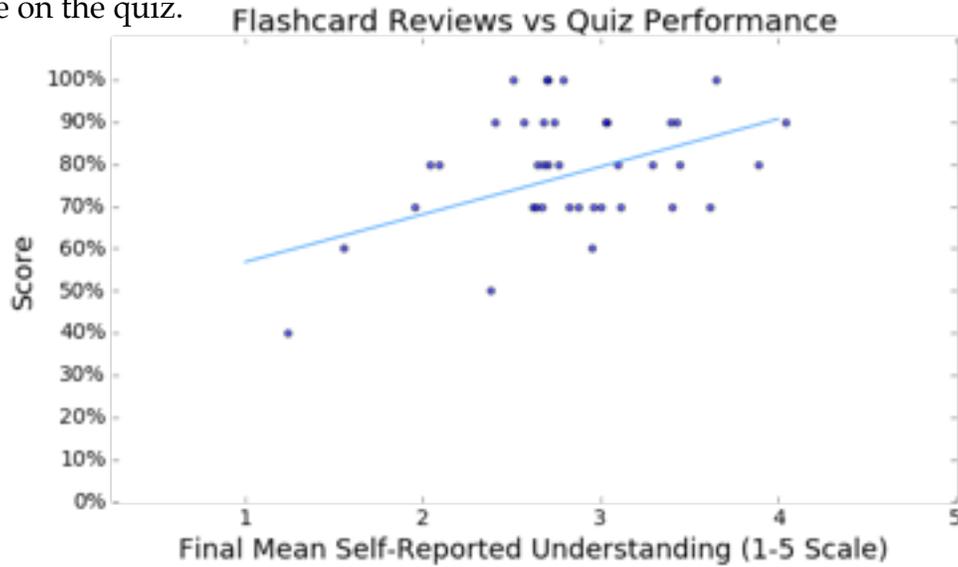
Firecracker includes aggregate statistics and charts as concise visualizations of student progress and achievement. The following histogram of student performance on a quiz at our partner school includes a clearly visible set of lower-performing students on the left side. This data allows faculty members to begin remediation efforts earlier than would have been possible without Firecracker alignment and weekly quizzes. For example, the students could be sent to a learning specialist long before midterms or finals.



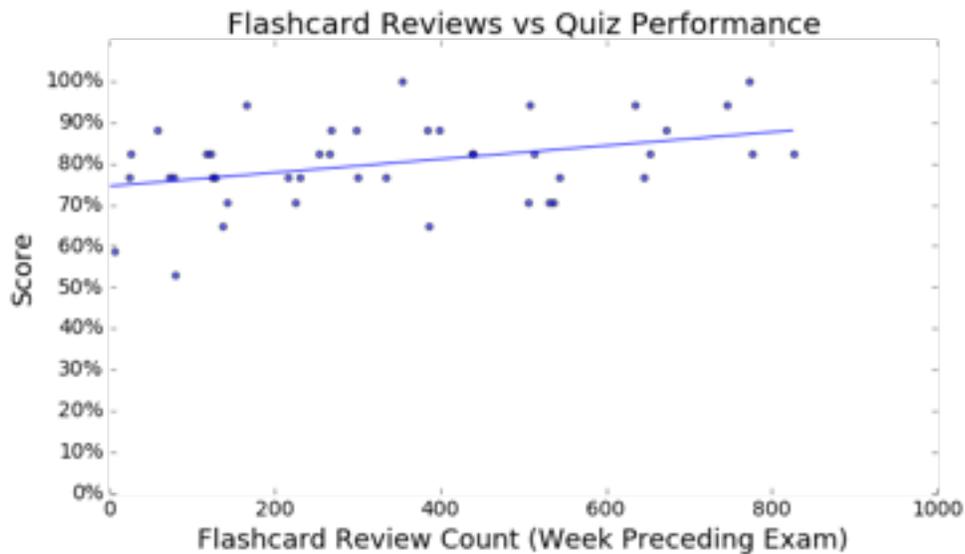
Students can optimally prepare for regular weekly exams, while simultaneously mastering concepts on their licensing exams, through adaptively spaced review on Firecracker. The graph below, employing data from the second week of a course at our partner school, demonstrates the correlation ( $R=0.4415$ ,  $p < .05$ ,  $n = 39$ ) between students'



final self-report flashcard review ratings on topics that that would be on their quiz, and their performance on the quiz.



Reviewing more flashcards relevant to a weekly exam, regardless of final self-reported score, is also correlated to better performance ( $R=0.3835$ ,  $p < .05$ ,  $n = 43$ ).



Finally, the graph below depicts the consistency with which prior quiz performance in a curriculum-aligned course correlates with a school-authored summative exam in that course ( $R=0.8456$ ,  $p < .05$ ,  $n = 50$ ). Alignment with Firecracker allows for continuous assessment that is predictive of overall student performance in the course, and that permits early intervention for struggling students.



