

Questizzle: A Crowdsourcing Tool for Questions and Assessments

Danny Grullon

Georgia Institute of Technology
North Ave NW, Atlanta, GA 30332
dgrullon3@gatech.edu

Abstract

Generating fair yet accurate assessments can be a tedious task due to the many different hidden characteristics a question must have for it to be considered valid and reasonable [1]. Writing such content requires a deep level of cognitive understanding of subject areas, which is why traditionally the task of creating quizzes and exams is usually handled by professors, teachers, and knowledge experts. Yet despite their professional expertise in education, assessments can still suffer from a lack of fairness due to human errors, biases, and recycled content from their creators [2]. Crowdsourcing the writing of questions and exams allows these potential issues to be better filtered out through a more democratic process while encouraging students to provide valuable feedback. *Questizzle* is a web application developed to harnesses the power of massive open online courses in rapidly generating high-quality questions and assessments on behalf of a course instructor. Content is created via the perspective of a test taker, while students are encouraged in continuously practicing the subject material.

Keywords

Educational Technology, Crowdsourced Content, Question Banks

1 Introduction

Periodically creating exam questions can be very time-consuming for teachers. Many tools already exist that helps aid the organization, evaluation, and distribution of test materials. However, despite the use of such resources, issues often still occur within the assessments themselves, such as lack of clear wording, inaccurately judging how long a test takes, or not having the proper test coverage. This can lead to an increase in frustration and stress for both students and professors. For example, University of Sheffield students once demanded their institution launch an investigation of an unfair economics exam, which students claimed contained advanced math that was never taught in class [2]. Much time and resources were spent by both parties in disputing such allegations. To make matters worse, some professors feel adding uncovered surprise topics to an exam as an ethical and accurate means of assessment [3]. The following sections will cover an online tool that helps counteract these issues through crowd-driven questions and assessments.

2 The Problem

One must know the intricate challenges posed in authoring any kind of assessment before thoroughly understanding the benefits of the proposed tool. Crafting an exam can be difficult because of the many factors that a good question comprises of. A posed problem should be consistently fair, measures knowledge reliably, and covers topics students should reasonably know to the point where the question itself can be legally defensible [1]. This is often easier said than done due to the many common mishaps that can occur in writing an assessment. For example, questions can be written with improper grammar or negative wording which can lead to increased confusion [4]. Tests can be written in a way that may unintentionally cause one minority group to misunderstand the possible answers [5]. Authors also have a hard time gauging the

difficulty and length of an exam. Tests that have too much of a time constraint can cause an increase in stress and panic which leads to students performing worse in remembering material [6].

Another important factor to consider is the format of the question itself, which largely depends on what kind of understanding will be measured. For example, multiple-choice, true/false, and matching questions are great at measuring a student's knowledge retention. Essays and short answers are good at gauging high cognitive understanding of the subject material. Every question type can have subtle variations in execution which can impact its level of difficulty. For instance, multiple-choice questions can be written with a single correct answer or multiple correct answers along with the option of granting no partial credit. Possible choices in answers can have "distractors" that work very well in persuading test takers. Writing these distractors is often a difficult task due to the need in exploring common misunderstandings about a concept [4].

3 Existing Solutions

Even though plenty of tools exist that aid in the gathering and evaluation of questions, few take into consideration the harnessing of online communities to outsource the creation of questions while providing better student engagement. None of the existing options allow communities to piece together the questions that comprise of an assessment, especially through a democratic survey-like process. Per the European Joint Research Commission, *TCEXam* is regarded as one of the most popular open source options in question bank software, yet the tool only provides basic housing and administration of questions [7]. Commercial solutions, like *Speedwell eSystems*, offer advanced analytics but doesn't offer any crowdsourcing capabilities [8]. Social networks, such as *Quizlet*, incorporate gamification and flashcard techniques but doesn't integrate well with existing courses [9]. The only option that comes close to the proposed tool is *PeerWise*, which is a site that allows instructors to delegate the creation of study questions to their students, but only allows using the accumulated content for practice [10].

Fig. 1. SWOT analysis of existing solutions and proposed tool.

	Strength	Weakness	Opportunities	Threats
Questizzle	<ul style="list-style-type: none"> • Crowdsourced questions. • Can promote questions to be part of exam. 	<ul style="list-style-type: none"> • Lack of mature features. • Abuse is possible. 	<ul style="list-style-type: none"> • Open source community can extend. • Can lead to more research. 	<ul style="list-style-type: none"> • Lack of adoption. • Other vendors can provide same features one day.
TCEXam [7]	<ul style="list-style-type: none"> • Open source and popular. • Random question selection. 	<ul style="list-style-type: none"> • Lacks advance analytics. • No student-generated content. 	<ul style="list-style-type: none"> • Popularity encourages adoption. 	<ul style="list-style-type: none"> • Dated user interface. • Instructor-oriented content.
Speedwell eSystem [8]	<ul style="list-style-type: none"> • Single author. • Performance statistics. • Multiple formats. 	<ul style="list-style-type: none"> • No student-driven content. • Not meant for studying material. 	<ul style="list-style-type: none"> • Commercial support encourages adoption. 	<ul style="list-style-type: none"> • Not open source. • Proprietary data. • Instructor-oriented content.
Quizlet [9]	<ul style="list-style-type: none"> • Supports mobile. • Vibrant community. • Gamification. 	<ul style="list-style-type: none"> • Geared towards flashcards. • Lack analytics. 	<ul style="list-style-type: none"> • Apply machine learning. • Allow voting. 	<ul style="list-style-type: none"> • Not open source. • No on-premise options.
PeerWise [10]	<ul style="list-style-type: none"> • Crowdsourced questions. • Vibrant community. • Q&A Metrics. 	<ul style="list-style-type: none"> • No feedback after deadlines. • Limited exporting. • Used for only practice. 	<ul style="list-style-type: none"> • Support more question types. • Lack of competitors. • Expose data via API. 	<ul style="list-style-type: none"> • Source code not available. • Only available to schools and institutions.

4 Solution

The proposed solution is a web application called *Questizzle*. The app allows online communities to accumulate questions via dedicated portals. A course instructor would register a portal for their classroom

while also providing a list of authorized students. These users would then be allowed to contribute questions to the classroom's bank of questions. Students can view each question individually, provide feedback, or use them as practice for an exam. The feedback consists of a comment system and a set of ratings: one for judging quality and another for difficulty.

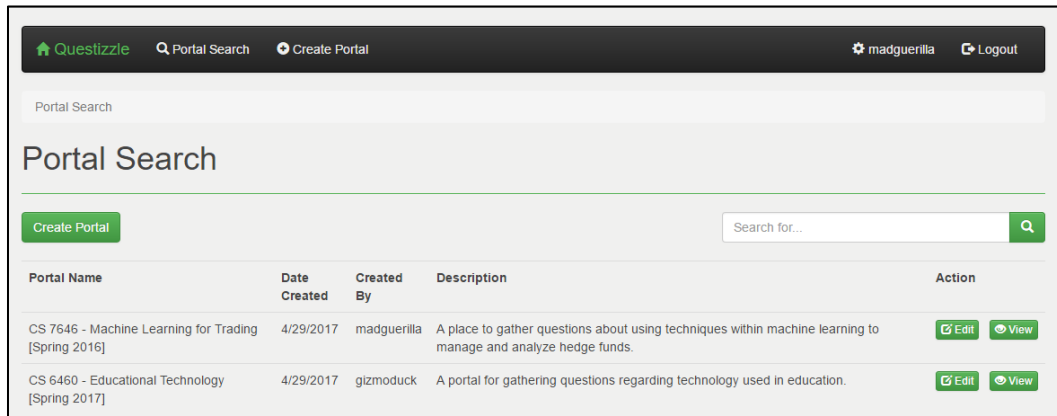


Fig. 2. Portal search screen allows managing dedicated banks of questions and assessments.

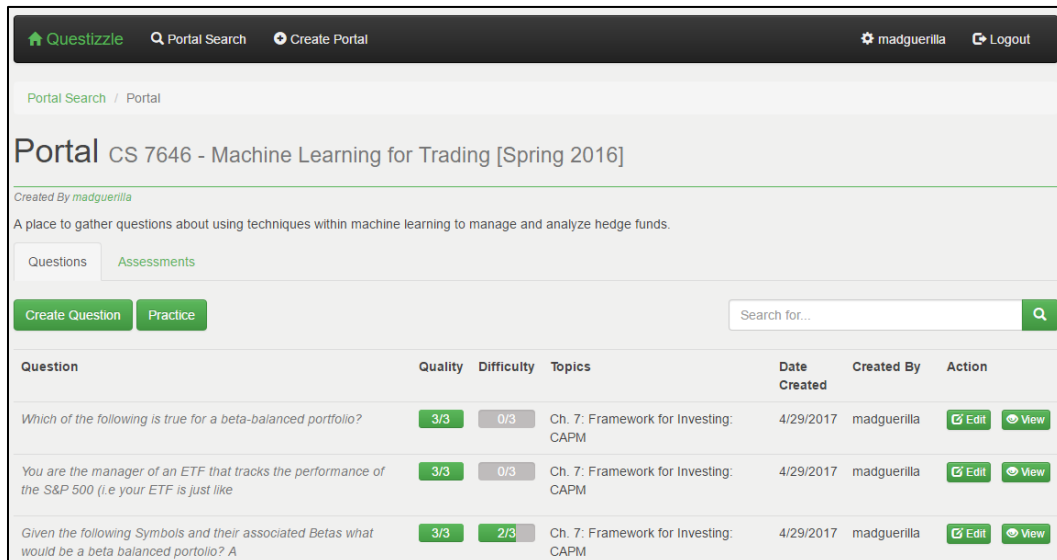


Fig. 3. Portal view with quality and difficulty metrics summarized per question and assessment.

The quality rating keeps track of how many users voted bad, poor, average, good, and excellent on a question, while the difficulty rating tracks how many users voted a question as easy, average, or hard. These metrics can be quickly glanced at via a portal's search area, allowing students to quickly find highest quality content. The best content can then be included to be part of an assessment, where another set of voting controls allow to provide feedback on the set of questions that comprise of a test. A user can vote on each assessment's question if they like the addition of that question as part of the exam. Assessments can then be exported through a printer friendly view once they reach a level of overall quality and stability.

Question

Created By gizmoduck

Fund A has outperformed SPY by 2% over the last year. The standard deviation of the difference between Fund A and SPY has been 8%. Consider these two questions:

- 1) What is the information ratio (IR) of the fund?
- 2) Assuming the fund can maintain it's IR over more stocks or trades what change to its trading could enable the fund to triple its information ratio?

[Choices](#)
[Hints](#)
[Feedback](#)

Quality

Rank the quality by selecting one of the below options:

[Bad 0](#)
[Poor 1](#)
[Average 2](#)
[Good 0](#)
[Excellent 0](#)

Difficulty

Rank the difficulty by selecting one of the below options:

[Easy 0](#)
[Average 1](#)
[Hard 2](#)

Comment

Type a comment here

[Post](#)

Fig. 4. View of a question, the voting, and comment system.

Assessment Midterm [Spring 2016]

Created By madguerilla [Printer Friendly](#)

Name Midterm [Spring 2016]

Instructions Answer the following multiple choice questions within the time allotted.

Topics

- Ch. 3: An Illustrative Hedge Fund Strategy: Arbitrage
- Ch. 4: Market-Making Mechanics
- Ch. 5: Introduction to Company Valuation
- Ch. 6: How Valuation Is Used by Hedge Funds
- Ch. 7: Framework for Investing: CAPM
- Ch. 8: The Efficient Market Hypothesis (EMH)
- Ch. 9: The Fundamental Law of Active Portfolio Management

Questions	Question	Topics	Date Created	Created By	Likes	Action
	When talking to your friend, he claims that certai	Ch. 5: Introduction to Company Valuation	4/29/2017	darkwingduck	1	Like View
	You are the manager of an ETF that tracks the perf	Ch. 7: Framework for Investing: CAPM	4/29/2017	madguerilla	3	Like View
	In the Capital Asset Pricing Model, on average wha	Ch. 7: Framework for Investing: CAPM	4/29/2017	madguerilla	1	Like View
	Barren Wuffett's portfolio has returned 20% over t	Ch. 9: The Fundamental Law of Active Portfolio Management	4/29/2017	gizmoduck	2	Like View

Feedback

Quality

Rank the quality by selecting one of the below options:

[Bad 0](#)
[Poor 0](#)
[Average 2](#)
[Good 1](#)
[Excellent 0](#)

Fig. 5. Assessment view allows to vote on each individual question and the overall assessment itself.

The web application comprises of two modules: a single page application and a web service. The user interface is built using *Angular 2*, *CSS 3*, *HTML 5*, *Twitter Bootstrap*, and *Typescript*. The web service is created using *Java 8*, the *Spring Boot* framework, and served via an embedded *Tomcat* application container. With some configuration, the RESTful web service can act as an API for third-party clients, leaving the door open for possible consumers of accumulated question bank content, such as native mobile apps or future research projects. The JSON payloads follow the latest in HATEOS and HAL standards allowing the API to be more discoverable and self-documenting. Authentication is handled via the exchange of JSON web tokens (JWT), which means the backend micro service is truly stateless and can scale horizontally as needed. *MongoDB* is used for persisting high-volumes of metrics and questions while providing flexibility with its schema-less document stores.

5 Methodology

In the Spring of 2016, the professor behind Georgia Tech's online course in *Machine Learning for Trading* asked each student to generate just one multiple-choice question per assigned chapter [11]. Students were encouraged to complete the assignment for participation credit if the question met high enough standards. The process was very manual, involved little-to-no feedback, and frequently resulted in duplicated or trivial content, yet the teacher assistants manage to gather between 200-300 exam questions easily. The class was then able to use the generated question bank as a giant study guide. The professor stitched together the course's mid-term by pulling content from submitted student questions and previously used exams. The proposed tool is an attempt at streamlining this process through online collaborative features.

Questizzle reduces the burden of generating high-quality assessments by distributing the problem towards students of massive open online courses (MOOC). MOOCs are a natural extension of online social communities, which means they have the same capabilities of generating content through the individual contribution of each volunteer [12]. The success of the experiment within the *Machine Learning for Trading* course was largely due to the class having hundreds of students. MOOCs usually comprise of a high number of students which implies that most online courses have the same potential in harnessing useful content from the classroom. Based on a recent study of *PeerWise*, a class comprised of 500 students generated more than 150 questions per day [13]. *Questizzle* extends this crowdsourced workflow to not just ask the community to submit questions, but also participate in the construction of future assessments themselves.



Fig. 6. The amount of contribution can be seen via the user profile view.

Assigning students to write exam questions forces students to repeatedly review the course material until they have a deeper cognitive understanding. This often involves students scanning covered topics in hopes of finding bits of information worthy of assessment. The accumulated question bank then provides them an almost endless supply of practice material. Continuous review and practice of course content have been shown to increase the overall academic performance of a class [14]. Incorporating feedback into the process

allows students to spot potential problems early on and can guide the revision of such questions accordingly. The “wisdom of the crowd” will make sure bad content is filtered down, while the best content surfaces to the top. Over time, the best content can be incorporated into future assessments. This creates a win-win situation for both the student and the teacher. Students learn more, while instructors save time.

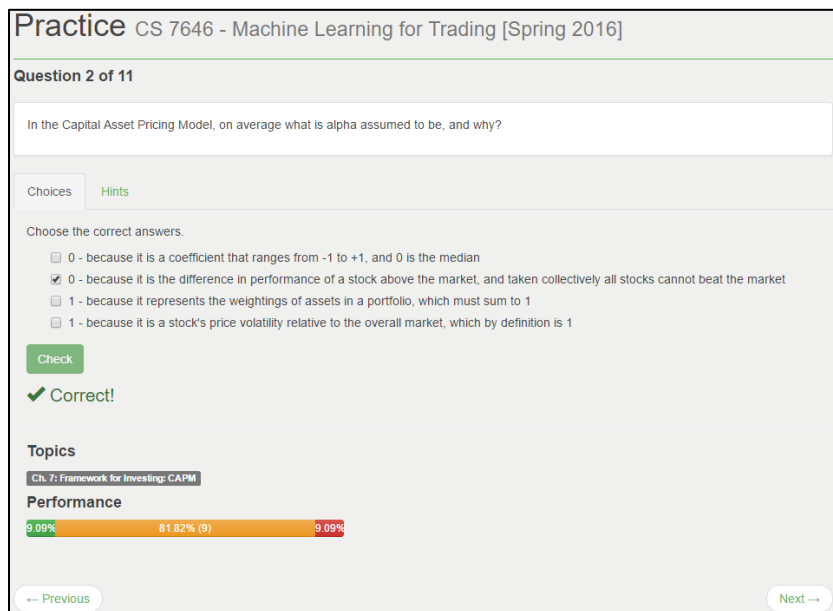


Fig. 7. View of a randomly generated practice test.

6 Future Work

The hope of developing this tool is to demonstrate the concepts behind a more social student-driven process of generating course content has practical potential; thus, the application is just a proof-of-concept and lacks matured features. Furthermore, no real-world data exist yet of its effectiveness within the classroom, so the obvious next step is to polish up existing features, offer the tool to online courses, and record its efficacy. The feedback can be used in evaluating alternate user interfaces, as well as other useful metrics, such as measuring the associated p-value of each assessment item as a more objective means of gauging question performance. Exposing the application’s source code to the open source community can also lead to more valuable enhancements, such as integration with learning management systems, gamification techniques, or intelligent tutoring software. Standardizing the underlying API can allow universities to share the accumulated classroom data and open the doors for further exploration of academic research or business products.

7 Conclusion

The writing of exam content can be a difficult task due to the many factors that go into making sure a question is deemed understandable, fair, and unbiased. Even though many tools already assist in the administration of tests, most tend to be designed on having one central authority. This means tests can suffer from potential mistakes a single author can make due to the natural limitations that exist in any human writer. The proposed proof-of-concept is an attempt at improving the traditional methods of exam creation by opening the doors to multiple contributors; therefore, embracing a more democratic process of quality selection and evaluation. Online courses are a natural fit for this process because they comprise of social communities that help promote human participation and knowledge sharing. *Questizzle* aims at harnessing

the untapped potential in these online communities by providing instructors the tools necessary for granting students a voice during the initial development of an assessment.

Acknowledgements

I would like to thank Greg Weber for mentoring this project, Tucker Balch for exposing me to the crowdsourced methodology, and my classmates who provided valuable feedback.

References

- [1] C. J. Brame, "Writing Good Multiple Choice Test Questions," in Center for Teaching, 2013. [Online]. Available: <https://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/>. Accessed: Feb. 26, 2017.
- [2] K. Gander, "University of Sheffield Students Demand Investigation Over 'Highly Unfair' Exam," in The Independent, The Independent, 2015. [Online]. Available: <http://www.independent.co.uk/student/news/university-of-sheffield-students-demand-investigation-over-highly-unfair-exam-10014553.html>. Accessed: Feb. 26, 2017.
- [3] S. K. Green, R. L. Johnson, D.-H. Kim, and N. A. Pope, "Ethics in Classroom Assessment Practices: Issues and Attitudes," *Teaching and Teacher Education*, vol. 23, pp. 999–1011, 2007.
- [4] M. E. Piontek, "Best Practices for Designing and Grading Exams," Center for Research on Learning and Teach Occasional Papers, no. 24, 2008.
- [5] Great Schools Partnership and L. Concepts, "Test Bias," in The Glossary of Education Reform, Great Schools Partnership, 2014. [Online]. Available: <http://edglossary.org/test-bias/>. Accessed: Feb. 26, 2017.
- [6] Corpus Christi College Cambridge, "What are my issues about exams," 2012. [Online]. Available: <http://www.corpus.cam.ac.uk/what-are-my-issues-about-exams/>. Accessed: Feb. 26, 2017.
- [7] N. Asuni, "TCEXAM," 2004. [Online]. Available: <https://tcexam.org/>. Accessed: Feb. 12, 2017.
- [8] Speedwell, "Question Banking for Digital Exam Software," in Speedwell Software. [Online]. Available: <https://www.speedwellsoftware.com/exam-software/question-banking/>. Accessed: Feb. 12, 2017.
- [9] Quizlet Inc., "Learning Tools and Flashcards for Free," in Quizlet. [Online]. Available: <https://quizlet.com/>. Accessed: Feb. 12, 2017.
- [10] The University of Auckland, "PeerWise," 2016. [Online]. Available: <https://peerwise.cs.auckland.ac.nz/>. Accessed: Feb. 12, 2017.
- [11] Udacity, "Machine Learning for Trading," in Udacity. [Online]. Available: <https://www.udacity.com/course/machine-learning-for-trading--ud501>. Accessed: Feb. 26, 2017.
- [12] J. Prpic, J. Melton, A. Taeihagh, and T. Anderson, "MOOCs and Crowdsourcing: Massive Courses and Massive Resources," *First Monday*, vol. 20, no. 12, 2015.
- [13] P. Denny, A. Luxton-Reilly, and J. Hamer, "The PeerWise System of Student Contributed Assessment Questions," University of Auckland.
- [14] K. A. Ericsson, R. T. Krampe, and C. Tesch-Romer, "The Role of Deliberate Practice in the Acquisition of Expert Performance," *Psychological Review*, vol. 100, no. 3, pp. 363–406, 1993.