# QuizCram: A Question-Directed Video Studying Interface

**Geza Kovacs**
Department of Computer
Science, Stanford University
geza@cs.stanford.edu

## Abstract

We present QuizCram, a question-focused format for navigating and reviewing lecture videos. QuizCram shows users a question to answer, with an associated video segment. Users navigate through the video segments by answering questions. We encourage users to review questions, by keeping track of their question-answering and video-watching history and scheduling users to review questions they have not fully mastered. We also allow users to review using a timeline of previously answered questions and videos. QuizCram-format courses can be generated automatically from lectures with in-video quizzes, though the format is flexible enough to accommodate multiple questions per video segment. Our user study comparing QuizCram to in-video quizzes finds that users are better able to remember answers to questions that they encountered when using QuizCram.

## Author Keywords

video flashcards, lecture reviewing, in-video questions

## ACM Classification Keywords

H.5.2. [Information Interfaces and Presentation (e.g. HCI)]: Graphical User Interfaces

## Introduction

Online lectures focus heavily on viewing video content. However, a phenomenon known as the *testing effect* shows that actively quizzing learners on the content is more effective for retention than simply having them passively watch videos [8]. Platforms such as Coursera have in-video quizzes which bring the benefits of testing into the video context by asking the user a multiple-choice question at key points in the video about the content that they have just watched. However, in-video quizzes are still given relatively little focus: they are shown only after the relevant segment has been viewed, and are often skipped by users.

Our system, Quiz-driven Video Cramming (QuizCram), attempts to make users focus more on questions that test the user, and helps direct their review process. To do so, we include the following features:

- Our interface encourages users to look at the in-video quiz before watching the video, so that it serves as an advance organizer to prime them towards the key concepts they should focus on while watching the video.

- Our system provides useful feedback in response to an incorrect answer, encourages the user to review the relevant portion of the video, and enforces that users can answer the question on their own before advancing them to the next portion of the video.

- To encourage people to review videos, our system keeps track of which video portions users need to review (using a score based on question scores on associated segments, percentage of the segment reviewed, and recency of reviewing), and gives them suggestions of questions and video portions to

review once they have watched all the video segments.

- We enable the more flexible addition of questions into the video, by allowing questions depend on video segments other than just the immediately preceding one. This allows for there to be a higher density of questions in the QuizCram format.

To evaluate the effectiveness of QuizCram for helping users study, we used a within-subjects study design comparing it to an in-video quiz format. Our contributions are:

- QuizCram, a format for viewing lectures in an interactive, question-centric manner, that can be automatically generated from existing videos with in-video quizzes.

- Users using QuizCram better remember answers to questions presented in the video better than when viewing videos in the in-video quiz format.

- Users are satisfied with QuizCram, and find the interface features for answering questions and reviewing videos to be helpful.

## Related Work

We designed our system features around a set of phenomenon from the education literature, which are also exploited by many other systems.

*Testing and Pre-Testing Effects*
The testing effect finds that repeated testing combined with fast, informative feedback helps students remember material [8]. Guiding questions and in-video quiz systems

are based on this principle: by testing the user on the video contents that were just viewed, they help students remember the material [4].

The Pre-Testing Effect finds that asking users to try answering a question before they actually study the material enhances long-term retention [7]. Our system encourages the use of the pre-testing effect by showing a question simultaneously with its associated video, encouraging users to preview the question first before studying the video.

*Advance Organizers: Video Transcript Summaries*
Advance organizers are information presented prior to learning, that helps the learner process the material that is about to be presented [10]. An example of an advance organizer for lecture video content would a summary of the video content that is to be watched. Video Digests is a system that creates such summaries about videos, and uses them as an advance organizer and navigational guide for video lectures [6]. Our system follows a similar strategy of breaking the video into segments associated with an advance organizer, but we instead use a question as an advance organizer that summarizes the clip to the user before they start watching it.

*Spaced repetition: Flashcards*
Spaced repetition is a technique designed to help learners retain information by having them review items at regular intervals [2]. A class of applications that exploit this are flashcards, where information is split into independent chunks that are scheduled for review based on factors such as mastery and recency of review. Flashcards can also have associated multimedia, such as video clips.

Similar to flashcards, our system also schedules items for review according to mastery and recency of review. One

key difference is that lecture videos build on each other, so this is an additional constraint for scheduling: the user needs to have covered the previous videos. Another key difference is the cost of review: a user memorizing vocabulary using flashcards only needs to spend a few seconds on each flashcard, while answering a question or reviewing a video clip takes an order of magnitude more time. Hence, the user will make fewer review passes through the video content than they would with vocabulary flashcards.
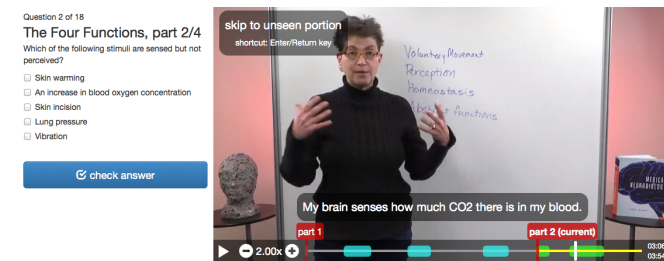
## System Design



**Figure 1:** The QuizCram interface, showing the current video. The focus question is on left, and the associated video is on the right. The progressbar highlights the relevant portion of the video in yellow. Already-watched segments of previous sections is in blue, already-watched segments of the current part are in green. Because we are currently watching a section we have already viewed, an option to skip to the unseen portion is shown.

QuizCram's interface shows users a question to review, with an associated video segment, as shown in Figure 1. It also shows a scrollable timeline of previously answered questions and associated video segments below the current question. Questions are first scheduled in order, then once the user has made an initial pass, questions are selected for review algorithmically, based on historic

correctness of responses, percentage of associated video that has been watched, and the recency of review. We also use the video progressbar to indicate the section of the video that is relevant to the current question, and portions of the video that the user has previously seen.

An existing course with in-video quizzes, such as MOOCs on Coursera, can be automatically transformed into the QuizCram format. This results in each video segment having one associated question. However, unlike in-video quizzes, the QuizCram format is also suitable for having multiple questions associated with a single video segment.

*Question-Focused Video Viewing*
For each section of the video in the course, we have one or more associated questions. We can get these question-video pairs automatically from existing videos with in-video quizzes, by associating the in-video quiz section with the immediately preceding video segment. For video segments that did not have an associated in-video quiz, we either automatically insert a generic "How well did you understand this video" question, or manually write a new question.

Whenever the user advances to a new section, we show the question and video concurrently, with the question to the left of the video, as shown in Figure 1. The video does not autoplay, so that the user has time to read the question before they start watching the video. If the user already knows the answer, they can answer the question and move on to the next section. Even if the user does not already know the answer, reading the question before they watch the video serves as an advance organizer which summarizes the key points they should pay attention to when watching the video.

Unlike in-video quizzes, which users are freely able to skip

over, in QuizCram the user must correctly answer the question before they can move on to the next question and associated video segment. This is designed to ensure that users learn the material before advancing onwards, as opposed to simply passively watching the videos without testing themselves.

Forcing users to answer the question may lead to frustration if the user is unable to determine the correct answer even after watching the video. Hence, whenever the user answers the question incorrectly, we provide them with immediate, informative feedback by showing the answers and providing an explanation, as opposed to the model used by Coursera where it states that the answer is incorrect, and only shows the explanation and correct answer after 3 tries. We made this design choice based on literature that finds that specific feedback that explains the correct answer to learners is more helpful and motivates them more than simply stating that their answer is incorrect [9].
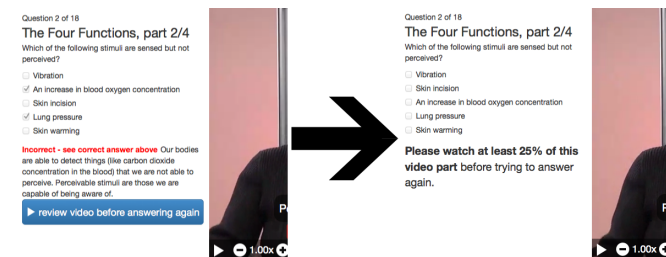


**Figure 2:** In response to an incorrect response, the user is asked to view an additional 10% of the video, the answer options are shuffled, and the user needs to re-answer the question correctly before moving on.

Of course, immediately showing the answer in response to an incorrect answer leads to the risk that learners may

choose to immediately reveal the answer without attempting to answer the question themselves. To discourage such behavior, even though we show the user the answer and explanation in response to an incorrect response, we do not advance them automatically. Rather, we shuffle the answer options and require them to view an additional 10% of the video, which is roughly 20 seconds, before attempting to answer it again, as shown in Figure 2. We do not enforce the 10% viewing requirement if the user has already watched over 75% of the video. This viewing task encourages users to view unseen portions of the video, incentivizes users to answer questions correctly, and ensures they aren't simply storing the answers in short-term memory and reproducing them. Requiring users to view the video and then retesting them after an incorrect response creates an additional retrieval opportunity, which should improve retention of the material [8].

While shuffling the answers and requiring video watching in response to an incorrect response discourages users from simply submitting the incorrect response and memorizing the answers, it does not entirely eliminate the risk. We can further discourage memorization of answers by having multiple questions for each video segment, which we alternate between. For example, in a algebra context we could simply ask the question again with different variable values whenever the user responds incorrectly. However, we did not use this option in our user studies since it would require us to write additional questions.

*Scheduling Questions and Video Sections for Review*
We want users to spend their study time focusing on material that they have not yet mastered. Hence, we assign each question a *mastery score*, which represents how well the user currently knows the material, and show users the questions for which they have low mastery score. The question's mastery score is based on the following 3 factors:

- *Past performance on question*: This element of the score encourages users to review questions they answered incorrectly. Each time a user tries answering the question, we give them a score between 0 to 1 based on the percent of checkboxes they correctly checked (the questions used in our study were all multiple-check questions). We then do a weighted-mean of all historic scores, with each newer score assigned 2 times more weight than the previous score (so more recent performance is weighted more heavily). For those video segments that have no associated question, we obtain this score by asking users to rate "How well did you understand this video?". If the user has never answered the question before, this has a default score of 0.

- *Fraction of associated video segment watched*: This element of the score encourages users to view video segments they have not seen. For each section of video, we keep track of whether the user has ever watched it. This score is the number of seconds watched in the question's video segment, divided by the total length of that video segment.

- *Recency of review*: This element of the score encourages spaced repetition for the questions. It also ensures that users are not shown the same questions repeatedly, which would make users bored. It is equal to 1 / number of questions elapsed since this question was last seen by the user.

If the question has never been seen, this has a default score of 0.

The mastery score is a weighted sum of these factors, where question correctness is 4/7 of the score, fraction of the video watched is 2/7 of the score, and recency of review is 1/7 of the score. We assign question correctness the highest priority because users should all be able to answer the questions correctly, but some users may choose to not watch portions of video they consider irrelevant or already know.

Sorting by the mastery score alone does not enforce that users have met the prerequisites for understanding the video and answering the question, before we show them the video and question. Unlike flashcards, lecture videos are meant to be watched in order and build on each other, so each video segment has a set of prerequisite videos which need to be watched before students can understand them. In our implementation, we enforce prerequisites by requiring that the user has correctly answered the questions for preceding video segments, before we show them the next video segment and associated question.

Sorting questions by mastery score and enforcing the prerequisites effectively results in users first being shown questions that work them through the videos in the order the course covers them, then asking them to review the questions they got low scores for and videos did not finish watching.

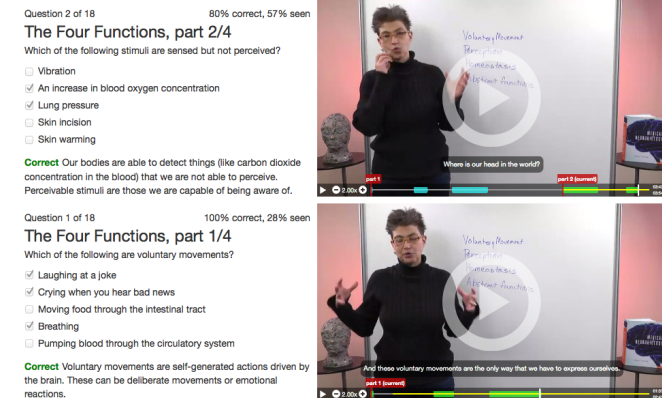*Timeline of Previous Questions and Videos*



**Figure 3:** The scrollable timeline, shown immediately below the current question, displays the past set of questions the user has answered. We list the correctness score and video progress scores to help users locate the questions they had difficulty with, and videos they have not yet fully watched.

Although QuizCram focuses the user's attention towards the current question and associated video segment, we also wish to make it easy to refer back to the previously answered questions and video segments. Whenever a question is correctly answered, we insert the next question and associated video segment at the top of the interface, and push the existing questions down. This results in a scrollable visual history of the previously answered questions and videos which we call the *timeline*, shown in Figure 3. The timeline displays the question and its answer and a miniaturized version of the video which can be clicked to enlarge it to full size and play it. The miniaturized video displays the frame the user left off at, so it serves both as a visual summary, and also allows users to easily resume viewing progress of previous videos. We also show the historic correctness of the user's

answers to that question, and percentage of the video they have watched, to help users identify questions they had trouble with and videos they did not fully watch.

The timeline gives users the option to use a more traditional, self-directed reviewing strategy, in contrast to the flashcard-style reviewing that our question scheduling algorithm encourages. By organizing the list of previous video segments according to the associated question that users answered, this allows users to scan video segments with a more salient summary than just the title. Question-based video navigation also allows users to search at a higher granularity, as questions refer to a specific subsection of the video, while the title refers only to the entire video contents. Furthermore, re-reading the previously answered questions helps trigger the users' memory of the associated clip, giving learners another retrieval opportunity to solidify their memory of the video contents.

*Directing Attention to Parts of Video Relevant to Question*
Standard in-video quiz viewers show the entire video at once. However, QuizCram shows only the part of the video relevant to answering the question, specifically, the start of the video up until the point where the question would be located (in an in-video context). We additionally highlight in yellow the section of the progressbar where question answer is located. This is designed to focus the user's attention to the portion of the video that will help them answer the question.

For questions generated from in-video quizzes, we highlight the segment of the video that immediately precedes the in-video quiz to indicate that it is where the answer is found, as shown in Figure 1. However, because we can highlight any preceding portion of the video to indicate that it is relevant to the current question, this

also allows us to have more flexibility in question writing and placement: we can place questions where they would fit most naturally, rather than immediately following the section where the answer is covered. This also enables us to have multiple questions that cover a single segment of video, without confusing users about where the answers to the questions are located.

*Directing Attention to Unseen Parts of Videos*
In addition, because QuizCram encourages reviewing videos, we wish to make it easy for users to keep track of what parts they have already watched. Hence, we highlight on the progressbar the already-seen parts in green (if it is from the current part of the video), or blue (if it is from a previous part of the video). If the user is viewing a section that has already been watched, we show a button at the top-left of the video that allows them to skip to the unseen portion. Similar techniques for visualizing the user's video viewing history have been presented in the literature [5] [3], though our system adds the novel feature of allowing users to skip to the next unseen portion.
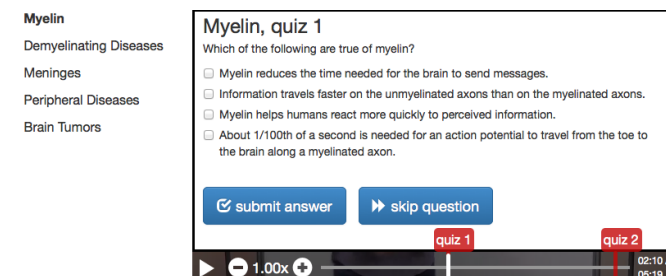


**Figure 4:** The in-video quiz format that served as our baseline. Left side lists videos, right side is a video viewer that shows the in-video quiz when reached. Locations of quizzes are indicated in red on the progressbar.

## Evaluation

Our user study was an within-subjects study that compared users' studying behavior with QuizCram to an in-video quiz interface that imitates the format used on Coursera, as shown in Figure 4. We took the videos, in-video quizzes, and unit exam from an existing Neurobiology course on Coursera. The QuizCram condition was generated from the original in-video quizzes, but also included additional questions that we inserted. We wished to answer the questions:

- Do users using QuizCram better remember answers to the original in-video questions?

- Do users using QuizCram perform better on the unit quiz?

- Can we improve recall of particular facts from the video by inserting additional questions with QuizCram?

- Do users find QuizCram helpful for studying videos?

*Study Design*
The study was a within-subjects design, where each learner used QuizCram and an in-video quiz viewer interface to study a set of videos. They were asked to provide qualitative feedback immediately after viewing, and were tested on the material they studied a day later.

*Participants*
We recruited 18 students by posting on university mailing lists and job boards. 12 were female, 6 male, their average age was 21.7 (stddev=4.91, min=18, max=37), and all had native-level English proficiency. We asked specifically that they have no experience with neuroscience, to ensure that they did not know the material beforehand. 9 participants reported having previous experience with MOOCs, and of these 6 had experience with Coursera. Participants received $60 for participating in the 2-hour online study.

*Materials*
The course materials – videos, in-video quizzes, and unit exams — were the first and second halves of Unit 1 of an existing Neurobiology course on Coursera. We generated the initial QuizCram materials directly from the course. Because we felt the question-to-video ratio in the original videos (9 questions for each 25-minute segment) was lower-than-optimal for QuizCram, we wrote additional questions for the QuizCram condition to double the total number of questions. We wrote questions in the same multiple-checkbox format as the original questions, and made sure that they did not ask the same facts that were tested on the unit exam or in-video questions.

We also wrote a set of free-response questions, one corresponding to each of our extra multiple-checkbox questions. We used these free-response questions to test whether users had learned the material tested by well enough to recall it.

*Procedure*
The study was conducted online over 2 days, with a 90-minute study section on the first day, and a 30-minute test section on the second day. Before users started the study, we informed them that they would be given 2 sets of videos, they should study them for 40 minutes apiece, and they would be given an exam on their contents in 24 hours. We did not tell them about the content of the exams.

On the first day, users used one tool to watch the first

half of Unit 1 (5 videos of length 23 minutes total). They were told after 40 minutes to fill out a survey about the tool. Then, they used the other tool to watch the second half of Unit 1 (5 videos of length 25 minutes total), and filled out the survey after 40 minutes of watching.

On the second day, users filled out a set of exams in the order listed below:

1. Extra free-response questions (as described in the Materials section), both halves

2. Original in-video questions from Coursera, both halves

3. Original unit exam from Coursera, both halves

4. Extra multiple-checkbox questions (as described in the Materials section), both halves

Parts 2-4 of the exam were automatically graded, giving each question a score equal to the fraction of checkboxes correctly checked. The free-response questions, which were of the general form "List N examples of X" or "State N facts about X", were graded by first marking each example provided by students as correct or incorrect. Then, we scored each response via the formula:

$$\frac{\#correct\ examples\ given}{Maximum(\#examples\ requested,\ \#examples\ given)}$$

*Exam Results*
Users were better able to answer the original in-video questions when using QuizCram. They averaged 85.4% with QuizCram, compared to 81.3% with the in-video quiz format. This difference was statistically significant (t=2.24, p=0.0391).

Unit quiz scores were similar when using QuizCram compared to the in-video condition. Average scores on the portion of questions covered by QuizCram was 65.1%, while scores for the questions from the portion viewed using the in-video interface was 63.4%. A t-test indicated no significant difference (t=0.44, p=0.669). Users performed slightly better on the questions from the second half of the unit exam: the average score was 59.2% for questions covering the first half of the videos, and 69.3% for questions covering the second half of videos (t=-1.98, p=0.064).

Users were better able to answer the extra questions we inserted in the QuizCram condition when viewing the section with QuizCram. They averaged 85.5% with QuizCram, compared to 76.0% with the in-video interface. This difference was statistically significant (t=2.44, p=0.0260). This is expected: users had previously seen these questions if they were using QuizCram, but they were not shown to users in the in-video quiz condition.

Users were also better able to answer the free-response questions when using QuizCram. They averaged 67.6% correctness with QuizCram, compared to 49.0% correctness with the in-video quiz format. A t-test showed this difference was statistically significant (t=3.95, p=0.0010).

*Survey Results*
Users reported their preferences as follows:

- In response to "Which tool would you rather use for studying?", 9 (50%) preferred QuizCram.

- In response to "Which tool would you rather use for studying if you were preparing for an exam and were short on time?", 11 (61%) preferred QuizCram.

- In response to "Which tool would you rather use for studying if wanted to remember the material long-term?", 11 (61%) preferred QuizCram.

Users generally liked QuizCram's question-focused format:

*I liked that it picked out the key information I should retain by asking me questions. It helped me decide what to focus on as I watched the video. The chunks were very manageable as well. I liked how it was broken up.*

## Discussion

The design goal behind QuizCram is to increase users' focus on questions, utilizing questions as a means to navigate and review the video material. Our user study focused on a short-term study task, modeling an exam-cramming scenario. In reality, however, we want to remember the contents of entire courses rather than single units, and need to study it across the period of months rather than hours. We believe the QuizCram format is well-suited for such use cases.

When reviewing lectures with traditional interfaces, the user needs to keep track of what they remember and what they need to review. This may be an easy task if they are reviewing only an hour of video. However, when studying entire courses over the course of a month, a user can easily lose track of what their study progress was. Instead, QuizCram keeps track of users' historic performance on questions and video progress, and makes suggestions for questions and associated segments of video to review. Thus, it relieves the user of the mental burden of needing to keep track of their study progress and determine what they need to review.

Another interesting finding was that with QuizCram, we were able to increase the number of questions associated with a video segment without adversely effecting the user experience or exam performance. In fact, we find that users remember the material covered by these additional questions well enough to answer them in free-response format. This paves a way for further increasing the amount of testing that occurs within video content.

Current online courses have external problem sets and quizzes outside of the ones in videos, because they cannot test the content in sufficient depth using in-video quizzes. However, if we consider the engagement patterns of users with MOOCs, the majority of users are interacting only with the videos and never doing the problem sets or quizzes [1]. Thus, moving more of the course content out of external quizzes and making the video more interactive and question-oriented provides a way to benefit these viewers' learning by testing their knowledge, without removing them from the scaffolding of videos. By gradually moving along this trajectory of making videos more question-focused and recommending review material to users, online courses of the future could entirely eliminate their need for external problem sets and quizzes, and transform into video and question-based intelligent tutoring systems.

## Conclusion

We have presented QuizCram, a system that uses questions to direct users' video viewing. QuizCram breaks the video into segments associated with questions, and always shows a focus question alongside the video. The focus question serves as an advance organizer that directs the user's attention towards the key points in the video. QuizCram also encourages reviewing based on questions: it displays a timeline of questions previously answered and their associated videos. It keeps track of users' progress

through questions and videos, and suggests users to review questions that they have not fully mastered. Courses in the QuizCram format can be automatically generated from existing video content with in-video quizzes, though it also has the flexibility to accommodate additional questions.

Our user study finds that QuizCram indeed does increase focus on questions – when the questions presented during viewing were tested again a day later, users using QuizCram performed better at answering the questions than users who encountered the questions as in-video quizzes. We also found that increasing the amount of questions presented with QuizCram results in users remembering the material tested by the additional questions better, even when answering based on recall not recognition.

Users' qualitative feedback indicates that they felt questions were an important part of the system. Users were divided in preferences between the QuizCram format and the standard in-video quiz format currently predominant in MOOCs. Some users liked the question-directed viewing format and thought it was more engaging, though others thought that displaying the questions were distracting. We believe the QuizCram format is a logical step from the in-video quiz format towards more interactive, question-focused intelligent video viewing platforms.

## References

[1] Anderson, A., Huttenlocher, D., Kleinberg, J., and Leskovec, J. Engaging with massive online courses. In *Proceedings of the 23rd international conference on World wide web*, International World Wide Web Conferences Steering Committee (2014), 687–698.

[2] Karpicke, J. D., and Bauernschmidt, A. Spaced retrieval: absolute spacing enhances learning regardless of relative spacing. *Journal of Experimental Psychology: Learning, Memory, and Cognition 37*, 5 (2011), 1250.

[3] Kim, J., Guo, P. J., Cai, C. J., Li, S.-W. D., Gajos, K. Z., and Miller, R. C. Data-driven interaction techniques for improving navigation of educational videos. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, ACM (2014).

[4] Lawson, T. J., Bodle, J. H., Houlette, M. A., and Haubner, R. R. Guiding questions enhance student learning from educational videos. *Teaching of Psychology 33*, 1 (2006), 31–33.

[5] Mertens, R., Farzan, R., and Brusilovsky, P. Social navigation in web lectures. In *Proceedings of the seventeenth conference on Hypertext and hypermedia*, ACM (2006), 41–44.

[6] Pavel, A., Reed, C., Hartmann, B., and Agrawala, M. Video digests: A browsable, skimmable format for informational lecture videos. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, ACM (2014).

[7] Richland, L. E., Kornell, N., and Kao, L. S. The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied 15*, 3 (2009), 243.

[8] Roediger III, H. L., and Butler, A. C. The critical role of retrieval practice in long-term retention. *Trends in cognitive sciences 15*, 1 (2011), 20–27.

[9] Shute, V. J. Focus on formative feedback. *Review of educational research 78*, 1 (2008), 153–189.

[10] Stone, C. L. A meta-analysis of advance organizer studies. *The Journal of Experimental Educational* (1983), 194–199.