

# How Learners Engage with In-Context Retrieval Exercises in Online Informational Videos

**Rimika Chaudhury**  
School of Computing Science  
Simon Fraser University  
Burnaby BC, Canada  
rchaudhu@sfu.ca

**Parmit K. Chilana**  
School of Computing Science  
Simon Fraser University  
Burnaby BC, Canada  
pchilana@cs.sfu.ca

## ABSTRACT

Learners increasingly refer to online videos for learning new technical concepts, but often overlook or forget key details. We investigated how retrieval practice, a learning strategy commonly used in education, could be designed to reinforce key concepts in online videos. We began with a formative study to understand users' perceptions of cued and free-recall retrieval techniques. We next designed a new in-context flashcard-based technique that provides expert-curated retrieval exercises in context of a video's playback. We evaluated this technique with 14 learners and investigated how learners engage with flashcards that are prompted automatically at pre-defined intervals or flashcards that appear on-demand. Our results overall showed that learners perceived automatically prompted flashcards to be less effortful and made the learners feel more confident about grasping key concepts in the video. However, learners found that on-demand flashcards gave them more control over their learning and allowed them to personalize their review of content. We discuss the implications of these findings for designing hybrid automatic and on-demand in-context retrieval exercises for online videos.

## ACM Classification Keywords

H.5.m Information interfaces and presentation (e.g., HCI): Miscellaneous

## Author Keywords

retrieval practice; video based learning; in-context exercises;

## INTRODUCTION

When trying to look up unfamiliar concepts or develop new skills, people are increasingly turning to online reference and multimedia resources. Among these resources, online informational videos have become one of the most popular avenues for informal learning [6]. In fact, many learners who sign up for learning platforms, such as MOOCs, often do so only to access the relevant videos [2].

Despite the convenience of accessing informational content through online videos, many learners often find these videos to be complex and overwhelming [25, 35], making it difficult to retain key concepts. Many of the lectures, demonstrations, and other instructional materials that learners access online rarely offer any formal structure for reviewing key concepts. Furthermore, informational videos in expert domains, such as information technology, can contain a lot of jargon and new concepts that learners may overlook or even forget by the end of the video [5, 32]. In some cases, a learner may end up playing the same video several times in a row or even consult other resources to reinforce the content they just viewed.

Given the pace at which new informational videos are being created and shared in online communities every day, how can we help learners better engage with these videos and help them retain key concepts? In particular, how do we offer this help to the growing population of informal learners who may not sign up for formal online courses [25], but are still interested in getting the most out of the videos that they do watch for developing specific skills?

In our research, we are exploring the design of interactive techniques to help learners retain key concepts in online informational videos by applying the high utility learning strategy of *retrieval practice* [15, 28]. The core idea of the retrieval practice strategy is that learners explicitly recall details of the content that they have previously viewed. This strategy has been shown to be effective in many scholastic learning situations [26, 27] where students are asked to reinforce their learning through intermittent quizzes, assessments, and other techniques. The other common usage of retrieval exercises is in areas which need memorization, such as language-learning [14]. More recently, retrieval exercises have also been used in tasks that not only improve recall through rote-memory, but also improve a learner's ability to infer [34]. Still, little is known about how retrieval practice could be used outside of the classroom environment in more informal learning contexts.

In this paper, we investigate the design and evaluation of in-context retrieval exercises for online informational videos from a user-centered design perspective. We first carried out a formative study with 14 participants to investigate their perceptions of common cued and free-recall retrieval exercise techniques (writing summaries, concept mapping, flashcards, multiple choice questions) in the context of watching videos related to information technology. Based on insights from this

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

L@S '19, June 24–25, 2019, Chicago, IL, USA

© 2019 ACM. ISBN 978-1-4503-6804-9/19/06... 15.00

DOI: <https://doi.org/10.1145/3330430.3333621>

study, we designed a novel in-context flashcard-based retrieval exercise technique that can be embedded in a video (Fig. 1). We evaluated two versions of this technique through another user study with a different set of 14 participants and assessed learners' perceptions of flashcards that were prompted automatically at predefined intervals in the video or flashcards that were accessed on-demand. Our findings revealed that learners preferred the automatic prompts over on-demand access as the automatic approach made the learners feel more confident and required less effort in determining which key concepts were worthy of review. However, many learners did find on-demand access to be useful as it gave them more control over their learning and allowed them to personalize their review by focusing on concepts that required further reinforcement.

Our main contribution in this paper is in providing empirical insights into how learners perceive and engage with different retrieval exercises when watching technical videos and in highlighting the strengths and weaknesses of offering these exercises automatically or on-demand during playback. Given the recent advances in automatic assessments and learning approaches [13, 38], our findings provide useful insights into what users actually find useful and also highlight the importance of giving learners some control over their learning. We discuss several implications of our findings, including the need to explore more hybrid automatic and on-demand personalized in-context retrieval exercises for informational videos.

## RELATED WORK

Our work builds upon research in learning sciences and human-computer interaction (HCI) with a particular focus on the theory and applications of retrieval practice.

### The Theory of Retrieval Practice

*Retrieval practice* is identified as one of the high utility learning techniques in education [28, 29] that makes few demands from the learner and promises high learning gains. Often referred to as the "testing effect", cognitive science explains retrieval practice as a way to promote learning by using low-stake "tests" or prompts to retrieve recently viewed concepts or facts from memory [29]. Studies have shown that retrieval exercises help improve short-term retention [30], and also play a role in long-term retention [27] in helping make inferences [34]. Other studies have further explored the idea of spacing retrieval practice and found that exercises that are spaced out over a span of time have been shown to be more effective than the ones that are close together [18, 27].

Most of these studies have been carried out in formal education settings and have focused on testing wide gaps of intervals, ranging from days and weeks over months. Still, given the learning gains possible through retrieval exercises, we explore this idea to see if learners would find such exercises useful in the context of learning from a single informational video.

### Innovations in Implementing Retrieval Exercises

Several innovations in learning have explored how retrieval exercises can be designed in different formats (e.g., free versus cued recall) [34] or in terms of spacing of the recall intervals [27, 18]. Investigations have also tried to identify how suitable

retrieval exercises can be adapted and integrated into courses [3, 5, 13, 15, 35]. However, much of the related work for utilizing or exploring the different ways of reaping the benefits of retrieval exercises have been done in a formal scholastic settings, mostly involving students enrolled in courses or a more mixed audience for MOOCs [1, 5] or in language learning settings [9, 32, 39]. In contrast, we investigate learners' perceptions and explore designs for retrieval exercises in non-scholastic situations, such as with the use of informational videos online.

Studies suggest that the act of attempting to recall itself provides benefits regardless of the success of the exercise [4]. Retrieval exercises can broadly be categorized into: 1) free recall, where the learners do not receive any hints while attempting the retrieval, and 2) cued recall, where some form of support is provided to help the learners recall better [34]. Some popular formats include brain dumps, quizzes and flashcards [15, 18, 21]. Some works have also found that the effect of retrieval practice could be attenuated depending on the material and its style of presentation [10]. Although the assistance provided during cued recall does not increase the benefits from the retrieval act itself, it does allow the learner a chance to correct their understanding. Taking this finding into account, we designed our formative study to initially assess learners' perceptions of retrieval exercises in multiple formats.

### Innovations in Video Based Learning

Recent innovations in HCI are also tackling the space of improving learning with online videos through various interactive strategies [7, 16, 19, 20, 23]. Some examples include the integration of interactive components within the context of a video, including comments, threads, assessments [23, 36], prompts with reflective questions [19, 37], and interactive note-taking [7]. These works suggest that integrating interactive elements to augment the video watching activity positively enhances not only learning outcomes, but also the overall learner engagement [8]. Our work complements these existing works by studying and designing interactive designs for retrieval exercises that can be integrated within videos. We focus on learners' perceptions and how they engage with different techniques to derive design insights rather than measuring learning outcomes resulting from the use of these techniques.

Integrating interactive components within the context of videos introduces the possibility of breaking the flow for the learner. Studies suggest solutions to the problem of inserting these break-points in the video in a suitable manner [31] and some innovations [7] have tried to predict suitable points. We have taken inspiration from spaced-retrieval using flashcards [18] and use expert-curated intervals, and also encourage learners to practice older items to strengthen their weak-points [19].

Finally, some recent innovations have tried to enrich the video-based learning experience by exploring collaborative designs, including crowdsourcing [20], teacher-student oriented feedback [16] as well as peer-to-peer [7] based techniques. Although our current investigation focuses on assessing the perceptions of individual learners, we offer insights for future work to explore collaborative designs for designing curated retrieval exercises at scale.

## HOW LEARNERS PERCEIVE RETRIEVAL EXERCISES: A FORMATIVE STUDY

As discussed above, retrieval exercises have been shown to be beneficial in improving learning outcomes in formal education settings where practice can be controlled and enforced. However, there are few insights into how such retrieval exercise techniques should be designed in more informal learning contexts, such as watching informational videos.

The goal of our formative study was to compare between different styles of retrieval exercises and assess how well learners engage with these techniques when watching informational videos. In particular, we focused on two main categories [34] of retrieval exercises: free recall and cued recall. During free recall, the learner is tasked to reproduce the material entirely on her own (e.g., writing summaries, concept mapping). During cued recall, the learner receives a cue that serves as an aid for memory retrieval (e.g., multiple choice questions (MCQ), flashcards). Our overarching research question was: *What are learners' perception of cued and free-recall retrieval exercises when watching informational videos?*

We designed a lab-based observational study with 14 learners and also collected user feedback through survey responses and follow-up interviews, as described below.

### Study Design and Procedure

During the study, we asked participants to watch short video lectures on popular technical concepts (Table 1) selected from *Lynda.com*. The content of each video mostly targeted learners on-the-job and we made efforts to maintain consistency by selecting videos that were similar to each other in terms of duration, style of presentation and in terms of number of concepts discussed. Following each video segment there was an associated learning activity (5 minutes), using one of the aforementioned retrieval techniques: free recall (writing summaries, concept mapping) and cued recall (MCQ and Flashcards). The study was conducted in a lab setting using the think-aloud protocol where participants were encouraged to share their thoughts throughout the study.

The retrieval exercises were generated using *Microsoft OneNote* as it offers various features for writing text, drawing, and adding forms. The free recall tasks required some typing interaction (writing summaries) and some drawing (concept mapping). The cued recall tasks only required click interactions (flashcards, MCQ). The participants were not provided with any review content or feedback (correctness of responses) after the free recall tasks. However, for the cued recall condition, the tasks were created using a list for MCQ and some text based flashcards where the cues themselves acted as review content. Optionally, participants could check the correctness of their response for the cued recall exercises if they wanted.

Following each condition, we asked participants to complete a survey related to the learning activity and rate their experience on a 5-point Likert scale and we also conducted a final short interview where we probed them about their experience and perceptions. The survey and interview questions focused on selected proxy engagement measures [24, 33], such as whether learners found these exercises to be helpful, enjoyable, dis-

Video Chapter	Test Conditions	
	Technique	Condition
IoT Foundations	Flashcards	C1
Data Science	Summaries	C2
Programming Foundations	MCQ	C3
Cloud Architecture	Concept-Map	C4

Table 1. Four retrieval exercise techniques were mapped to four videos.

tracting and to what extent the exercises encouraged reflection, offered a sense of control, and boosted learner confidence.

### Participants

There were 14 participants (8 male, 6 female) and all were university students within an age range of 19-31 and came from Computer Science or Information Technology backgrounds. All of the participants had experience in consulting technical videos outside the classroom. They received \$15 Amazon Gift Cards for an hour-long study. The retrieval techniques and the videos were paired as shown in Table 1. We followed the Latin Square arrangement to expose the participants to the treatments to counter any order effects. Conducting one-on-one user study allowed us to draw out the deeper reasons behind the pros and cons of the various retrieval techniques from the participants' perspective.

### Results

Based on our analysis of the observational, survey, and interview data, we synthesized key insights into how learners perceived and engaged with different retrieval exercises.

*Retrieval Exercises Promote Thinking while Watching Videos*  
Overall, participants indicated that *"all of the activities [exercises] were engaging"* [P02] as these exercises gave them a chance to pause and think more deeply about the concepts shown in the video. Participants noted that the exercises made the ideas concrete or made them realize immediately what they could not recall: *"without the activities [exercises], I would simply end up re-watching the videos again and again"* [P02].

Although, most of the participants (11/14) indicated that the exercises were overall not distracting, some participants did mention that free recall exercises, such as summaries, *"should not appear in the middle [of the video] when I am already thinking of something...preferable to have something simple there.."* [P03]. Many participants (5/14) also indicated that they did not enjoy the free recall exercise of concept mapping.

### Less Effort Made Cued Recall More Enjoyable

Participants in general were more positive about cued recall exercises compared to free recall. The overall consensus was that cued recall required less cognitive effort to remember key details: *"It is hard to recall when a topic is new"* [P13] and participants also felt having cues as a *"confirmation [of their understanding] would be useful"* [P06]. Although participants felt that free recall would be more beneficial for their learning in the long run, free recall exercises in the context of watching videos seemed like *"too much work"* [P12] and most participants did not feel satisfied with their recall. In contrast, the

presence of hints in the cued condition and the option to check the correctness of their responses served as useful feedback for the participants:

*"...Writing summary was my least favorite... least useful. No feedback does not help me."* [P02]

#### *Learners Like Control with Flashcards*

With cued recall emerging as the favored retrieval exercise technique, we looked closer to compare MCQ and flashcard styles. We found that 9 out of 14 participants reported enjoying the MCQ exercises because the "MCQ [format] was familiar" [P10] and "there are the other options that give you more hints" [P13]. However, many participants rated MCQs as being more complicated than flashcards due to the level of detail in the question and sometimes confusing or tricky options. They preferred a simpler representation of the content. Of the participants who agreed that cued retrieval techniques encouraged reflection of key concepts, 2 participants mentioned that MCQ did not give them "much control" and that these exercises did not always provide "enough practice".

In contrast, participants perceived flashcards to be more favorable. In fact, 10 out of 14 participants agreed to some extent that they felt more "in control" of their study as they could review cards several times before testing themselves. They found the material on the flashcards to be less distracting and overall found them to be useful. For example, one participant explained: "[flashcards] are useful in case you forget some material and self-test helps confirm if you get it" [P05]. Participants also mentioned that they would prefer flashcards when they have a "[low] attention span or... time is a constraint" [P01]. Participants perceived flashcards to provide a comprehensive review and thus were more useful in helping them recognize their pain-points, as opposed to MCQ which offered little benefit for repeated practice.

#### **Design Implications for Retrieval Exercises**

Based on the formative study insights, we identified the following implications for designing engaging retrieval exercises for informational videos:

##### *D1. Minimize recall effort*

A key takeaway from the formative study was that learners perceived cued recall, with flashcards in particular, to be more enjoyable and less effortful than free-recall. Cued recall exercises should be designed for informational videos to be easy to perform, require minimum interaction, and require minimal knowledge construction from scratch.

##### *D2. Provide curated exercises*

We found that content that is curated by experts was a trusted source of feedback and actively sought by learners. Hence, retrieval exercises in informational videos that provide curated feedback would be more useful for learners.

##### *D3. Allow Learners to have more control*

While learners find curated exercises useful, they also want some flexibility in selecting the content and the frequency of their review based on their evolving understanding. Retrieval exercises in informational videos should offer learners control

over their pace of learning and opportunities to reflect on what they do or do not understand.

#### **DESIGNING CURATED IN-CONTEXT FLASHCARDS**

Informed by the design implications above, we designed a novel web-based retrieval exercise technique consisting of interactive curated flashcards that can be retrieved in-context of video playback within a browser (Fig. 1.1).

#### **Interactive Curated Flashcards**

The key idea motivating our flashcard retrieval exercise technique was that a learner should be able to interactively reflect on key concepts at different intervals within the video. In line with D2, we offered flashcards that were curated by subject matter experts.

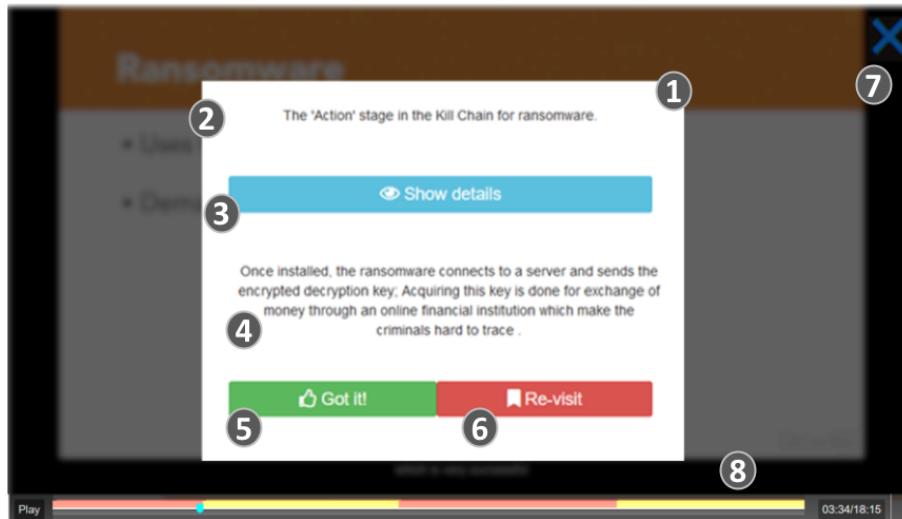
Each flashcard consists of two statements: Statement 1 (S1) presents a concept discussed in the video (Fig. 1.2) and provides the learner a general statement about the concept as a cue (D1). Statement 2 (S2) (Fig. 1.4) reveals more minute details relating to the concept in S1 (D2). By default, S2 remains hidden and is only revealed when the learner clicks on the "Show details" (Fig. 1.3) button. S2 may contain specific keywords, jargon, terminologies or precise examples and instances of the concept in S1. The choice of viewing S2 for details allows the learner to control the depth of the retrieval exercise (D3) and provides additional confirmation if desired by the learner.

If a learner is satisfied by her understanding of the the concept on the flashcard, she can select the "Got it!" button (Fig. 1.5), and move on to the next flashcard. Otherwise, she can bookmark the concept by clicking on "Re-visit" (Fig. 1.6) and the card will be re-retrieved in the next set of flashcards. There is also a dismiss button (Fig. 1.7) available to allow learners to dismiss the exercises at any stage of completion and continue watching the video. In line with D3, the learner has control in assessing her comprehension, reflecting on her weaknesses, and doing no reviews or additional reviews when necessary. Lastly, a comprehensive practice at the end of the video is also available to allow the learner to run through all of the flashcards again for one final review.

#### **In-Context Access to Relevant Flashcards**

With our interactive flashcard technique, a new flashcard becomes available after a concept has been discussed in the video. A time-stamp of the moment that particular concept was covered in the video gets attached to the flashcard. The segmented seek-bar, in Fig. 1.8, shows different segments in the video where flashcards would be available. For example, in Fig. 1.1, the learner has completed watching a quarter of the video and has arrived at the end of the first segment. At this point, the learner gets a set of flashcards that have a smaller time-stamp than the the current moment in the video. At the next stopping point in the video, a new set of flashcards become available for practice along with a set of cards from the previous segment(s) that the learner had bookmarked (if any). This design pertains to the goal of providing curated exercises (D2) with timely availability for engaging the learner (D1).

In designing our flashcard-based in-context retrieval technique, a key question we faced was when and how often would learners actually want to interact with retrieval exercises when



**Figure 1.** This figure shows the interface of a flashcard in an exercise set, that appears on an overlay on top of the video. 1) The flashcard. 2) Statement 1 (S1) is the cue pertaining to the concept. 3) Button that reveals Statement 2 (S2). 4) S2 pertaining to more minute details relating to S1. 5) "Got it!" button to indicate that the learner does not need to practice the card again. 6) "Re-visit" button for bookmarking a card for practice at the next interval. 7) Dismiss button to cancel the prompt. 8) Segmented seekbar that provides visual cues for when the next set of exercises will appear.

watching informational videos? To investigate this, we explored two designs for our flashcard retrieval exercise technique: 1) *Automatic Prompts*: flashcards that appear automatically in the video at predetermined intervals decided by experts; or, 2) *On-Demand Prompts*: flashcards that can only be retrieved on-demand by learners when they feel the need to review or reflect on a concept.

### USER EVALUATION OF IN-CONTEXT FLASHCARDS

To assess the design of our flashcard retrieval technique and to compare the two possibilities for offering the flashcards in context of informational videos, we designed a comparative observational user study. We sought to answer two main research questions:

1. *What are learners' perceptions of the utility of interactive in-context flashcard-based retrieval exercise embedded in informational videos?*
2. *What are learners' perceptions of flashcard exercises offered automatically, at intervals prescribed by experts, versus flashcard exercises that are accessed on-demand by learners on their own?*

### Study Design

We used a within-subjects design for our study, where each learner used both versions (automatic and on-demand prompts) of the flashcard-based retrieval technique embedded in two different informational videos (Table 2). We worked with experts to first identify key concepts in each of the videos and accordingly created a flashcard for each of these concepts (total 12 for each video). Each concept on a flashcard (Fig. 1) was associated with a time-stamp to facilitate the retrieval of relevant flashcards at different intervals during playback.

Video Chapter	Technique	Test Conditions
		Condition
Cloud Security	Automatic Prompts	C1
Cyber Security	On-Demand	C2

**Table 2.** Automatic and on-demand flashcard techniques were mapped to two videos

### *Automatic Prompts Condition*

In the automatic condition, a set of flashcards automatically appeared as an overlay in the video at intervals curated by experts. These intervals were natural topic breaks in the video and experts created flashcards relevant to content that would appear within 4 roughly equally-spaced intervals. The intervals were indicated on the seek-bar with alternating colors to provide a visual cue (Fig. 1.8). The video automatically paused at the end of these intervals and a precursory prompt with a link to access flashcard exercises was shown to the learner. The whole set of qualifying cards (based on the time-stamp) was retrieved where there were two to three flashcards in each set, relevant to the concepts discussed in the immediately preceding interval.

### *On-demand Prompts Condition*

In this condition, no flashcards or intervals were shown to the learner automatically- the learner could access flashcards by explicitly pausing the video at any point. The pause would then trigger the same overlay to extend over the video and display the precursory prompt with a link to access flashcard exercises. Since each flashcard was associated with a specific time-stamp, it would only become available for review after the learner had viewed the relevant portion of the video. The visual cue using colors to indicate topic breaks was removed in this condition. Other features of the interactive flashcards described earlier, such as bookmarking and skipping, worked the same way in both conditions.

## Procedure

We recruited a new set of 14 participants (7M, 7F), all between the ages of 19-40 who either had a computer science related education, or relevant technical experience. All participants were given \$15 Amazon Gift Cards for their participation.

Each participant first began the study by filling out a demographic questionnaire. Next, participants interacted with the two different flashcard techniques described above. To counter order effects, all participants were randomly assigned to either the automatic or the on-demand condition. Before starting either condition, participants were given the following scenario: they were newly hired employees of a technology company which expected them to come up to speed with some new technical concepts for their upcoming project. The participants were then given a tutorial of the flashcard-based exercises and asked to use it just as they would if they were in the aforementioned scenario while watching two 18-minute long video lectures - one for each condition.

The participants were asked not to take notes and were instead encouraged to make the best use of the hints on the flashcard to recall the concept-related details. We introduced the bookmarking feature and encouraged participants to use it until they were comfortable with the concept. There was no time limit for the exercise sessions or for completing the surveys, but we maintained an upper-bound of 1.5 hours for the entire study. The videos were sourced from Lynda.com (see Table 2) and were similar to each other in terms of duration, style of presentation and the number of concepts discussed.

For the automatic condition, the participants were informed about the visual cue on the video seek-bar that indicated the segments in the video and the points where a set of flashcards would appear automatically. They were free to either take a suggested review at the end of each segment or dismiss it if they felt they did not need it but were asked to take at least one exercise at any point. For the on-demand condition, the participants were asked to pause the video on their own and access the flashcards for practice whenever they felt the need for it. In this condition also, they were asked to take at least one exercise at any point.

The study was conducted in a lab setting and the Think-Aloud protocol was followed where the participants were encouraged to share their thoughts during the study.

## Data Collection and Analysis

We collected data by directly observing participants' behavior during the study, recording the screens, collecting their feedback through survey responses and drew further insights from interviews which were recorded and later transcribed.

In this evaluation, we allowed learners to have more freedom than our formative study by allowing the participants to choose how many times they wanted to engage with the flashcard exercises. We also made observations to find out whether the learners were cognitively engaging with the exercises by recording their screen and noting their bookmarking behavior, how many times they dismissed an automatically prompted exercise or voluntarily stopped to view an exercise set, how often they saw the details in each card, and also the total time

spent in engaging with exercises. Similar to our formative study, we also used selected proxy engagement measures [24] which were relevant to our study. We were mainly interested in finding out if our in-context flashcard exercises were useful, enjoyable, offered a sense of control over their study, boosted confidence about the material, seemed time consuming or distracting and promoted focused attention.

## Results

We present our main findings about how flashcard-based retrieval exercises were perceived when they appeared in context of an informational video and how the automatic delivery of the flashcard exercises compared with on-demand access.

### Utility of In-Context Flashcard Exercises

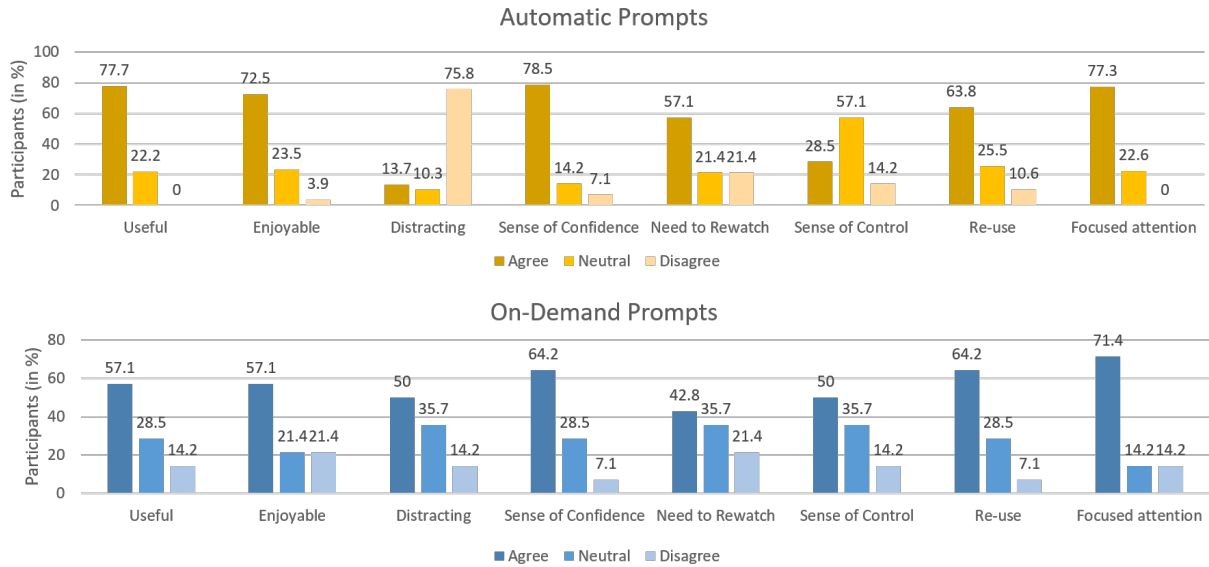
We found that most participants (10/14) agreed to some extent that they found in-context flashcards to be useful. They cited many reasons, such as being able to quickly remind themselves of the key concepts before progressing on to the next sections in the video. The flashcard exercises also helped participants confirm their understanding: *"The flashcards gave me a more solid understanding of what the topic was [all] about"* [P03] Flashcards showed potential for use in conceptual or informational videos, especially those that contained new information for the learner:

*"...watching information based [videos] where they are trying to transfer some kind of knowledge, these kinds of [flashcard] reviews help. When you are less familiar with a topic, you need to retain more and need to grasp more"* [P11]

Participants mentioned that in situations where they experienced difficulty in grasping the concept, they tended to re-watch the video several times. However, after taking the flashcard-based retrieval exercises, over half of the participants (8/14) disagreed or were neutral about feeling the need to re-watch the video. One participant explained: *"You could actually do away with the video after just one watch."* [P12]. Overall, flashcards provided a useful alternative to re-watching the videos, *"If I am in a hurry"* or *"if I am looking at a conceptual material"* and *"do not want to go through the whole material"* [P12]. The participants displayed a high level of engagement by staying committed to the flashcard exercises once they had decided to take them.

Many participants (8/14) made use of the bookmark feature in the automatic condition (34 bookmarks) and 6/14 participants used bookmarks in the on-demand condition (26 bookmarks), generating 60 bookmarking events in both conditions combined. This shows that the learners were not only actively engaging in the exercises, but also regulating their understanding and planning for repeated practice on their own (without being prompted by the experimenter). The participants confirmed in the interview that they found bookmarking to be a useful feature: *"If you are just doing the final flashcard review in the end, and if you are doubtful about it, you can bookmark it, go through the whole thing and have it pop up again."* [P04]

Additionally, most participants (10/14) in either condition attempted the optional final review at the end of the video,



**Figure 2. Study results for user-engagement measures for Condition 1: Automatic Prompts and Condition 2: On-Demand Prompts. Participants rated the review exercises on a 5-point Likert scale ranging from Strongly Disagree (Rating 1) to Strongly Agree (Rating 5). In the above figure, Strongly Agree and Agree responses are added together and labelled as Agree. Similarly, Strongly Disagree and Disagree are clubbed and labelled as Disagree.**

suggesting that they were engaged with the final retrieval exercise and did not rush to finish the study.

#### *Accessing Flashcards Automatically vs. On-Demand*

To compare both the automatic and on-demand conditions for accessing flashcards, we analyzed the data collected through direct observations, survey responses and by interviews. We focused on user interaction and behavior in both conditions as well as perceptions of effort, confidence, and control.

**User Interaction and Behavior:** In the automatic condition, we observed that participants’ engagement with the exercises was consistently high across all of our key metrics (Fig. 2). Although participants were required to attempt only a minimum of one exercise and could skip the rest, it was encouraging to see that all 14 participants took 3 or more exercises when automatically prompted (in fact, 12/14 participants took all 4 exercises). Survey feedback indicated that only one participant found these exercises to be distracting when watching the informational video and the distraction mainly seemed to stem from the colors in the video player seek-bar (Fig. 1.8).

The rationale behind the on-demand access was to provide the learners with more control over their study session - allow them to pause for reflection more naturally, when they felt the need for it. We observed that most participants (10/14) proactively paused to take the retrieval exercises during the on-demand condition, suggesting that retrieval exercises were helping the participants to grasp key concepts in a progressive manner. We also observed that the participants varied greatly in how frequently or regularly they paused - mostly ranging between 2 to 5 times, but in extreme cases - never pausing in between [P3, P6, P8] or pausing up-to 12 times [P10]. However, unlike the expert-curated intervals in the automatic condition, the user-selected pauses rarely aligned with the natural topic-breaks in the video. This behavior suggests that

the participants were indeed relying on their natural sense of flow to determine when to stop, which may not necessarily align with topic-breaks in the video.

**Perceptions of confidence in understanding content:** In the automatic condition, most participants (11/14) agreed that they felt more confident about the content after taking the flashcard exercises. The participants felt that the retrieval exercises at pre-defined intervals provided a beneficial guidance. Qualitative insights confirmed that the automatic prompts helped the participants gain more confidence about the material and automatic prompts at regular intervals “*made the ideas stick*” [P04] as they progressed through mounting details. This also helped them gain better value out of the comprehensive exercise in the end. “[*Automatically prompted flashcards*] helped me recollect things I forgot from the earlier segments.” [P02]. “*I was able to have small recaps before I could take the final one.. so I was more confident about the particular sections made the comprehensive review at the end more useful*” [P04]. In contrast, during the on-demand access condition, fewer participants (9/14) agreed to feeling confident about the material after taking the exercises. Our interviews revealed that participants were more uncertain in the on-demand condition as to whether they had actually stopped at an appropriate point to seek review in the video.

**Perceptions of effort:** As we noted, half of the participants (7/14) in the on-demand condition only took one or two of the four possible exercises. We found that although the participants liked having the freedom to pause at any point for a retrieval exercise, there were cases where the participants felt that having to pause on their own added to their efforts of trying to learn from the video. The on-demand access appeared to add more effort because it was difficult for learners to identify suitable places to stop: “*I have to tell myself I need to stop and*

take some exercise. Before watching the video I was planning to pause 3 times but later I found that some important things are going on and I cannot pause here.” [P14].

In fact, some participants were unable to identify natural topic breaks at all due to the absence of visual cues: “*The second video had more concepts... but no [topic] segments.*” [P03]. The automatic prompts also made the decision to stop for exercises easier: “*better if it [retrieval exercise] is forced on me*” [P08] and “*I am lazy... so I wouldn’t pause on my own if I know [the flashcards] will appear automatically.*” [P03].

**Perceptions of control over learning:** Despite the overall merits of the automatic condition, a key drawback noted by participants was a lower sense of control over their study. Although a learner could dismiss or postpone a prompted exercise and bookmark cards for more practice, only a few participants (4/14) agreed to some extent, feeling in control. A key reason for this was that the participants felt that the predefined timing for the prompts did not always align well with their need for a review: “*It was not the right timing for me... too many pauses slow me down.*” [P07]

The on-demand access of flashcards in the second condition made up for the shortcomings of the automatic prompts and control. Participants overall appreciated having control over their learning by being able to access exercises whenever they felt the need to reinforce the concepts they just learned: “*I prefer the [on-demand design] where I have control because I don’t want to get interrupted. I like the flow.*” [P04] Furthermore, with the ability to access the exercises on demand, participants said that when they “*...found [my] attention wandering...it helped [me] slow down and pay more attention.*” [P07].

We also observed some instances where participants wanted even more control: “*I prefer to be able to choose what [I want] to check*” [P06]. This participant preferred to see only the most immediate flashcard, instead of revisiting all cards.

Participants overall indicated that although they had a slight preference for the automatically prompted flashcards, the ideal case would be one in which they have “*best of both worlds*” [P12], meaning on-demand access was a desirable feature but best when augmented with automatic prompts.

## DISCUSSION

This paper contributes insights into how learners perceive and engage with retrieval exercises in informational videos. Our goal was not to assess the effectiveness of these techniques in promoting learning outcomes, but rather to take a user-centered approach to explore the design of retrieval exercises and how they can be made useful. In particular, our findings have shed light on several benefits and drawbacks of designing in-context flashcard-based retrieval exercises and how they should be offered within informational videos.

We now reflect on some limitations of our study and highlight several opportunities for future research in learning sciences and HCI to explore further innovations and empirical results.

## Study Limitations

One limitation of the current study is that we only assessed retrieval exercises in the context of technical informational videos. Although our interviews with participants suggest that the design of retrieval exercises would be relevant to other subject areas, future work should investigate the relevance of our findings for other types of informational videos. In addition, while all of our participants were adult learners who had experience in learning from online videos, learners in the workplace or other settings where informal learning may take place could exhibit different behaviors and perceptions. We acknowledge that this study used an experimental prototype with basic video features and there are several other ways of designing video-watching interfaces and presenting curated content that should be further explored. Lastly, in this study, we used proxy measures to gauge engagement. Future work should explore the use of more direct measures for assessing engagement. Nonetheless, our study is a first step towards designing engaging user-centered retrieval-based interventions for enhancing retention in the context of watching informational videos and opens up several promising research directions for further enhancing retrieval-based interventions.

## Supporting automatic interleaved in-context exercises

The design and evaluation of our in-context flashcard retrieval technique showed that overall learners found the flashcard-based retrieval exercises to be useful and engaging. In particular, the automatically prompted flashcards were most useful as they required less effort and allowed learners to feel more confident about the video content. The automatic in-context format helped learners maintain focused attention and prevented them from steering away to find other resources when they encountered difficulties in comprehension [11]. However, despite the strengths of the automatic approach, a key challenge in designing these exercises is to determine how to appropriately offer them to learners. We observed that although participants rarely missed an automatically prompted exercise, they did like having the choice to dismiss it.

One implication here is to design automatic prompts to serve as a gentle reminder at suitable intervals rather than pausing the video. Since earlier studies involving retrieval exercises have established the benefits of spaced out retrieval practice [27], future research could potentially explore variations in spacing these intervals, and predicting an appropriate moment to prompt the learner for taking an exercise over different lengths of duration in informational videos. In addition, future research could also explore how flashcard content itself could be generated automatically to complement recent efforts in automatic generation of questions in various styles of presentation, such as MCQ [13] and fill in the gaps [3].

## Scaling curating efforts

Another key insight from our findings was that the learners found expert-curated intervals to be helpful in aligning their thoughts rather than figuring out where they should pause-and-reflect entirely on their own. This presents several opportunities to design interventions that provide curated content within the context of a video. For example, future work can consider crowdsourcing-based techniques that have already been



leveraged in several studies in HCI in the context of learning [12, 22], such as for aiding video navigation and generating study material for learners [7, 20]. Prior research provides insights into peaks in learner's viewership in the time-line of a video [17], which could be used to aid potential contributors [38], such as video-authors or domain-experts, to locate suitable points for inserting exercises. Potentially, peers can also contribute content which could then be moderated by experts. There is already an active community of domain experts who engage in contributing content online, such as in forums. Currently, in-context exercises are integrated within videos only before production. However, using automated web technologies there is a chance to augment the videos with simple reflective exercises based on crowdsourced locations where learners actually could benefit from reflection. Future work could look into ways of tapping into the potential of the learner community to aid them to extend their ways of contributing content for informational videos.

### Personalizing retrieval exercises with hybrid designs

Our study revealed that learners like to be able to attempt an exercise when they start feeling overwhelmed or find their attention waning. Several learners also pointed out that they like their study to be more targeted and wanted to be able to choose which of the curated exercises to attempt or practice. Although they liked to be able to bookmark a flashcard to revisit it later, they felt it was unnecessary to practice other items every time. Learners also described pain-points with a concept that they may experience in between two automatically curated exercises, highlighting one of the strengths of the control offered by the on-demand design.

These findings point toward a retrieval exercise technique that is a hybrid between an automatically guided routine and a self-regulated on-demand approach. The automatic approach offers learners the chance to refine their understanding of the material, but a hybrid design, with on-demand exercises, could enable learners to make their study more selective and directed. Pain-points may differ among individuals and a hybrid of these two designs provides a chance to personalize their reflection and self-assessment. Another direction for personalizing reviews in the future could be to design approaches for learners to even create their own in-context flashcards, or to customize expert-curated flashcards to facilitate and complement their own note-taking processes.

### CONCLUSION

Thousands of informational videos are being created and shared in online communities every day, but learners often overlook or forget key concepts introduced in these videos. We investigated how the theory of retrieval practice that has been shown to be successful in scholastic learning can be adapted and designed for informational videos. Informed by findings from our formative study comparing different retrieval exercise formats, we designed a new in-context flashcard-based technique that provides expert-curated retrieval exercises in context of a video's playback. We compared automatic and on-demand techniques for offering these flashcard exercises, revealing several insights into how these exercises should be designed from the perspective of learners in the future. We

believe that these insights will be useful for future designs for incorporating retrieval exercises in a range of online informational videos.

### ACKNOWLEDGEMENTS

We thank Laton Vermette, Bernhard Riecke, and the Natural Science and Engineering Research Council of Canada.

### REFERENCES

1. 2016. Retrieval practice and study planning in MOOCs: Exploring classroom-based self-regulated learning strategies at scale. *Lecture Notes in Computer Science* (2016), 57–71.
2. Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. 2014. Engaging with massive online courses. In *Proc. international conference on World wide web*. ACM, 687–698.
3. Lee Becker, Sumit Basu, and Lucy Vanderwende. 2012. Mind the gap: learning to choose gaps for question generation. In *Proc. Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 742–751.
4. Janell R. Blunt and Jeffrey D. Karpicke. 2014. Learning with retrieval-based concept mapping. *Journal of Educational Psychology* 106, 3 (2014), 849–858.
5. Dan Davis and Claudia Hauff. 2018. The Half-Life of MOOC Knowledge A Randomized Trial Evaluating Knowledge Retention and Retrieval Practice in MOOCs. *Proc. Learning Analytics and Knowledge* (2018), 1–10.
6. Degreed.com. 2016. How the Workforce Learns in 2016. (2016).
7. Vania Dimitrova, Antonija Mitrovic, Alicja Piotrkowicz, Lydia Lau, and Amali Weerasinghe. 2017. Using learning analytics to devise interactive personalised nudges for active video watching. In *Proc. User Modeling, Adaptation and Personalization*. ACM, 22–31.
8. Samuel Dodson, Ido Roll, Matthew Fong, Dongwook Yoon, Negar M Harandi, and Sidney Fels. 2018. An active viewing framework for video-based learning. In *Proc. L@S*. ACM, Article 24.
9. Darren Edge, Stephen Fitchett, Michael Whitney, and James Landay. 2012. MemReflex: Adaptive Flashcards for Mobile Microlearning. In *Proc. MobileHCI '12*. ACM, 193–194.
10. Luke G. Eglington and Sean H.K. Kang. 2018. Retrieval Practice Benefits Deductive Inference. *Educational Psychology Review* 30, 1 (2018), 215–228.
11. John Frens, Erin Walker, and Gary Hsieh. 2018. Supporting answerers with feedback in social Q&A. In *Proc. L@S*. ACM, Article 10.
12. Neil T Heffernan, Korinn S Ostrow, Kim Kelly, Douglas Selent, Eric G Van Inwegen, Xiaolu Xiong, and Joseph Jay Williams. 2016. The Future of Adaptive Learning: Does the Crowd Hold the Key? *International*

- Journal of Artificial Intelligence in Education* (2016), 615–644.
13. Yi-Ting Huang, Ya-Min Tseng, Yeali S Sun, and Meng Chang Chen. 2014. TEDQuiz: automatic quiz generation for TED talks video clips to assess listening comprehension. In *IEEE Conference on Advanced Learning Technologies (ICALT)*. 350–354.
  14. Sean HK Kang, Tamar H Gollan, and Harold Pashler. 2013. Don't just repeat after me: Retrieval practice is better than imitation for foreign vocabulary learning. *Psychonomic bulletin & review* 20, 6 (2013), 1259–1265.
  15. Jeffrey D Karpicke. 2009. Metacognitive control and strategy selection: deciding to practice retrieval during learning. *Journal of Experimental Psychology: General* (2009), 469–486.
  16. Juho Kim, Elena L Glassman, Andrés Monroy-hernández Meredith, and Ringel Morris. 2015. RIMES : Embedding Interactive Multimedia Exercises in Lecture Videos. *CHI* (2015), 1535–1544.
  17. Juho Kim, Philip J. Guo, Daniel T. Seaton, Piotr Mitros, Krzysztof Z. Gajos, and Robert C. Miller. 2014. Understanding in-video dropouts and interaction peaks in online lecture videos. *Proc. L@S* (2014), 31–40.
  18. Nate Kornell and Kalif E Vaughn. 2016. How retrieval attempts affect learning: A review and synthesis. In *Psychology of learning and motivation*. Vol. 65. Elsevier, 183–215.
  19. Geza Kovacs. 2015. QuizCram: A Question-Driven Video Studying Interface. *Extended Abstracts CHI'15 2* (2015), 133–138.
  20. Ching Liu, Juho Kim, and Hao-chuan Wang. 2018. ConceptScape : Collaborative Concept Mapping for Video Learning. *CHI* (2018), 1–12.
  21. Mark A McDaniel, Pooja K Agarwal, Barbie J Huelser, Kathleen B McDermott, and Henry L Roediger III. 2011. Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology* (2011), 399–414.
  22. Piotr Mitros. 2015. Learnersourcing of complex assessments. In *Proc. L@S*. ACM, 317–320.
  23. Toni-Jan Keith Palma Monserrat, Yawen Li, Shengdong Zhao, and Xiang Cao. 2014. L.IVE: an integrated interactive video-based learning environment. *Proc. CHI* (2014), 3399–3402.
  24. Heather L O'Brien, Paul Cairns, and Mark Hall. 2018. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies* 112 (2018), 28–39.
  25. Yeonjeong Park, Insung Jung, and Thomas C Reeves. 2015. Learning from MOOCs: A qualitative case study from the learners' perspectives. *Educational Media International* 52, 2 (2015), 72–87.
  26. Bernhard Pastötter and Karl Heinz T. Bäuml. 2014. Retrieval practice enhances new learning: The forward effect of testing. *Frontiers in Psychology* (2014), 1–5.
  27. Henry L. Roediger and Andrew C. Butler. 2011. The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences* 15, 1 (2011), 20–27.
  28. Henry L Roediger III and Jeffrey D Karpicke. 2006a. The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science* (2006), 181–210.
  29. Henry L Roediger III and Jeffrey D Karpicke. 2006b. Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological science* (2006), 249–255.
  30. Ralf Schmidmaier, Rene Ebersbach, Miriam Schiller, Inga Hege, Matthias Holzer, and Martin R Fischer. 2011. Using electronic flashcards to promote learning in medical students: retesting versus restudying. *Medical education* (2011), 1101–1110.
  31. Niels Seidel. 2016. Video Assessment Patterns: Designing Video Player for integrated Assessment. *EuroPlop '16: Proc.* (2016), 1–20.
  32. Burr Settles and Brendan Meeder. 2016. A Trainable Spaced Repetition Model for Language Learning. *Proc. Association for Computational Linguistics* (2016), 1848–1858.
  33. Vivek Singh, Balaji Padmanabhan, Triparna de Vreede, Gert-Jan de Vreede, Stephanie Andel, Paul E Spector, Steve Benfield, and Ahmad Aslami. 2018. A content engagement score for online learning platforms. In *Proc. L@S*. ACM.
  34. Megan A Smith and Jeffrey D Karpicke. 2014. Retrieval practice with short-answer, multiple-choice, and hybrid tests. *Memory* (2014), 784–802.
  35. Tim Van Der Zee, Dan Davis, Nadira Saab, Bas Giesbers, Jasper Ginn, Frans Van Der Sluis, Fred Paas, and Wilfried Admiraal. 2018. Evaluating Retrieval Practice in a MOOC: How writing and reading summaries of videos affects student learning. *Proc. Learning Analytics and Knowledge Article* 10 (2018), 216–225.
  36. Sarah Weir, Juho Kim, Krzysztof Z Gajos, and Robert C Miller. 2015. Learnersourcing subgoal labels for how-to videos. In *Proc. CSCW*. ACM, 405–416.
  37. Joseph Jay Williams and Juho Kim. 2018. Understanding the Effect of In-Video Prompting on Learners and Instructors. *Proc. CHI* (2018), 1–12.
  38. Joseph Jay Williams, Juho Kim, Anna Rafferty, Samuel Maldonado, Krzysztof Z Gajos, Walter S Lasecki, and Neil Heffernan. 2016. Axis: Generating explanations at scale with learnersourcing and machine learning. In *Proc. L@S*. ACM, 379–388.
  39. Kathryn T Wissman, Katherine A Rawson, and Mary A Pyc. 2012. How and when do students use flashcards? *Memory* 20, Article 6 (2012), 568–579.