

Scripted Vicarious Dialogues: Educational Video Augmentation Method for Increasing Isolated Students' Engagement

Thitaree Tanprasert
 tt1996@cs.ubc.ca
 University of British Columbia
 Vancouver, British Columbia, Canada

Sidney Fels
 ssfels@ece.ubc.ca
 University of British Columbia
 Vancouver, British Columbia, Canada

Luanne Sinnamon
 luanne.sinnamon@ubc.ca
 University of British Columbia
 Vancouver, British Columbia, Canada

Dongwook Yoon
 yoon@cs.ubc.ca
 University of British Columbia
 Vancouver, British Columbia, Canada

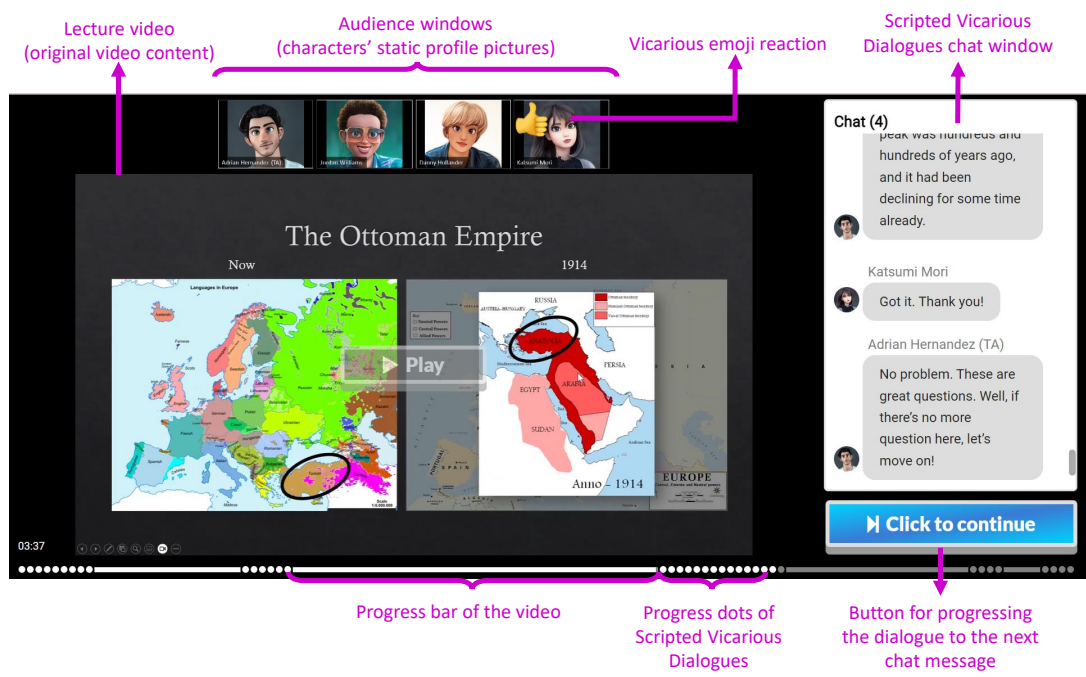


Figure 1: A screenshot of Scripted Vicarious Dialogues (SVD) interface. A lecture video is framed in the metaphor of Zoom-like online classroom that features a suite of vicarious social interaction components (e.g., audience pictures, emoji reactions, chats). The learner can navigate the content using the progress bars (for the lecture video), progress dots (for SVD chats), and the "Click to continue" button (for both).

ABSTRACT

Videos are convenient resources for asynchronous learning, but they lack interpersonal interactions found in synchronous classrooms. Due to missed social connectedness, the isolated video-based

learners experience low emotional, behavioral, and cognitive engagement. This work presents "Scripted Vicarious Dialogues" (SVD), a technique for engaging students in a pseudo-social experience of witnessing scripted dialogues between virtual characters (teaching assistants and students) around a video. We conducted a participatory design study to derive design guidelines for SVD. The findings indicate the need to distinguish the virtual components and to give students control of the dialogue's pace. We then implemented an interactive prototype of SVD and evaluated it (N=40) against a non-social, direct-learning baseline. The results show that the preference for SVD versus the baseline is polarized (25 of 40 preferred SVD; no neutral preferences), and those who preferred SVD had

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 the author(s) 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.
 CHI '23, April 23–28, 2023, Hamburg, Germany
 © 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
 ACM ISBN 978-1-4503-9421-5/23/04...\$15.00
<https://doi.org/10.1145/3544548.3581153>

significantly higher emotional and behavioral engagement with SVD compared to the baseline.

CCS CONCEPTS

• **Applied computing** → **E-learning**; *Distance learning*; • **Human-centered computing** → **Empirical studies in HCI**.

KEYWORDS

video-based learning, augmented learning environment, vicarious learning, learner engagement

ACM Reference Format:

Thitree Tanprasert, Sidney Fels, Luanne Sinnamon, and Dongwook Yoon. 2023. Scripted Vicarious Dialogues: Educational Video Augmentation Method for Increasing Isolated Students' Engagement. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, April 23–28, 2023, Hamburg, Germany. ACM, New York, NY, USA, 25 pages. <https://doi.org/10.1145/3544548.3581153>

1 INTRODUCTION

Educational videos are useful for isolated learning environment, which has become more prevalent with the surging popularity of learner-centered pedagogy, the availability of educational videos platforms like MOOCs, and the shift to remote learning during the COVID-19 endemic [16]. However, there are pedagogical limitations of lecture videos. Firstly, they feature the instructor alone and do not have any interpersonal interactions, such as questions from other students or peer discussion [7]. Research shows that learners may feel isolated, which affects their motivation to learn, leading to declining "in-the-moment" behavioral, emotional, and cognitive engagement [44].¹ Lowered engagement is problematic, as it can diminish the learners' learning outcome [17]. The lack of social interaction and means of engagement is an opportunity for enriching the learning environment around educational videos.

There are several methods for adding social components to video-watching experiences, but they have shortcomings that undercut the advantages of video-based learning. For instance, a synchronous co-watching session between students and instructors or TAs would sacrifice the convenience for students to learn at a time and place of own choosing [69]. A recording of a synchronous classroom may allow students to learn from others' interactions vicariously, but there may be a privacy issue for reuse of recorded class [5]. Moreover, the quality and quantity of questions are not guaranteed in live lectures – what students say might not always be conducive to the vicarious learners and the number of questions might be too low to produce positive effects [24]. Going beyond synchronous approaches, the visible and timestamped comment features, such as Danmaku [71], YouTube comments [20], or TrACE [18], can facilitate some interactions, but the community-driven comments tend to be noisy and require hands-on moderation to guarantee the relevance and quality [32]. In contrast, our goal is to devise and test a method for learner engagements where the augmented content is well-curated by the instructors and conducive to the learner's convenience of access.

In this paper, we propose Scripted Vicarious Dialogues (SVD) for increasing learners' engagement. SVD is a mode of video augmentation, where learners witness the augmented content as scripted instructional dialogues between virtual characters (e.g., teaching assistants, peer students) as if they were co-watching the video in a Zoom classroom. We draw upon Social Constructivism in using interactions to increase the engagement of participators (in this case, the virtual characters). And the vicarious interaction theory supports that the benefits can be transferred to the vicarious learners, who witness the dialogues². As shown in Fig. 1, the interactions between the virtual characters are scripted as though they are watching the lecture video for the first time together on a video conferencing platform (e.g., Zoom) and engage in activities and discussions related to the video along the way.

Although SVD utilizes virtual characters, it should be noted that SVD is different from agent-based approaches, such as tutoring/pedagogical agents, teachable agents, and peer agents [8, 39]. With SVD, the learners do not interact with any virtual characters or influence their dialogues in any way; all their engagements will happen vicariously through observing the virtual characters' interactions. The scripted nature of SVD gives the instructor full control to make the dialogue's content productive and comprehensive, e.g., they can include frequently asked and naive questions and common mistakes that learners can learn from. In contrast, virtual agents simulate dialogues based on learners' behaviors and performances. In other words, the agent-based approaches are about dialogue generation, while SVD is about dialogue presentation.

We aim to design the SVD experience – how to present the dialogues and the characters for the best effect – and examine its efficacy in invoking students' engagement. Since SVD is a novel concept in its combination of multiple-character settings and vicarious settings, we cannot derive its design requirements and predict its effects based on the literature on existing approaches. We need instructors' inputs on their strategies to engage students and students' requirements of the SVD concept. The research question that drives our investigation of SVD is: how can we design SVD in a way that increases isolated learners' emotional, behavioral, and cognitive engagement?

To answer the research question, we conducted a two-phase study. In the first phase, we derived design guidelines for SVD through a participatory design study with six university instructors and three teaching assistants experienced in using lecture videos in their classes. We identified three types of interpersonal interactions that constitute effective SVD and derived three design guidelines for SVD. Based on these findings, we designed a preliminary prototype of SVD to get feedback from students, then finalized the design. In the second phase, we implemented an interactive prototype and conducted an evaluative study. We ran an evaluative study with 40 undergraduate students, comparing SVD to the baseline (non-social, in-video, pop-up quizzes). We found that the preference for SVD versus the baseline is polarized (no neutral preference) and that the potential factors affecting the preference are the learner's individual need for social connectedness, perception of virtuality, and learning strategies. We also found that those who preferred SVD

¹We will provide a full definition of *engagement* and brief discussion of its taxonomy in Section 2.1.

²See Section 3.1 for a detailed explanation of the conceptual and theoretical underpinnings.

had a significantly higher emotional and behavioral engagement with SVD compared to the baseline.

The primary contribution of this paper is the conceptual contribution of SVD as a novel mode of augmenting educational video. Within this contribution, there are three secondary contributions:

- (1) the design guidelines for SVD: we present three design guidelines for the presentation, the virtual characters' personas, and the learners' control of SVD (Section 3.3).
- (2) SVD's design artifact: the design as shown in Fig. 1 grounded on the theoretical literature and the design guidelines (Section 3.4).
- (3) the empirical evaluation of SVD (with World History as an example subject domain): our results show that SVD increases learners' emotional and behavioral engagement without sacrificing cognitive engagement, if the learner values interpersonal interactions and are comfortable with virtual dialogues (Section 5). This result is also an empirical evidence that the vicarious interaction theory can be applied to observing dialogues between purely virtual characters.

2 RELATED WORK

2.1 Definition of Engagement and relevant theories

In this section, we will define the notion of *engagement* as used in this research and summarize two theories related to engagement that are the theoretical underpinning for the SVD concept: Social Constructivism and vicarious interaction theory.

Engagement is a term with multiple definitions and is used in many contexts. In this study, we employ the definition of academic engagement proposed by Fredricks et al, which consists of three constructs: behavioral, emotional, and cognitive engagement [22]. We chose this definition for two reasons: (1) this is a widely used definition in educational research and is associated with instruments used for measuring engagement; and (2) because it encompasses social engagement with peers, which aligns with our research goal [1].

- Emotional engagement includes satisfaction, enjoyment, excitement, interest, and curiosity when interacting with class materials, instructors, and peers.³
- Behavioral engagement includes students' attention, concentration, persistence, and in synchronous class environments, participation in class activities.
- Cognitive engagement consists of shallow engagement, such as identifying important topics and memorizing, and deep engagement, such as constructing concepts, applying theories, and evaluating the usefulness of the knowledge.

This definition can be applied to a wide range of educational contexts, however, in this paper, we are particularly interested in situational or "in-the-moment" engagement. This refers to situation-specific engagement over a short period, which contrasts with long-term engagement that looks at students' engagement over a course, a semester, or a school year [64].

³It should be noted that, for emotional engagement, there is the reverse disengagement called *disaffection*, which looks at students' negative emotions such as frustration or boredom. This is different from a lack of engagement and is important for understanding the engagement phenomenon as well [31].

The first theory, Social Constructivism, states that "knowledge is co-constructed in the environment (interpsychologically) with others. [...] Individual learning is actually a product of knowledge creation, and this happens when collaboration takes place and knowledge itself gets co-created in the environment" [67]. The theory has been applied in many pedagogical approaches and interfaces for increasing students' attention and learning gain through interaction with real and virtual peers or More Knowledgeable Others (MKOs) [21, 55, 61]. (MKO refers to anyone who is "more capable" or "possesses higher cognitive capabilities" than the students [70].) In social science or arts subjects, such interaction could also facilitate emotional engagement through shared experiences [30]. However, since this emotional effect has not been shown to generalize to other subjects, we will only apply this theory to motivate the increase in behavioral and cognitive engagement with SVD.

The second theory is vicarious interaction. The theory is derived from Social Cognitive Theory, which explains how humans learn from observation [4, 41]. It states that "when an otherwise passive student (*vicarious interactor*) actively observes, absorbs, and processes the ongoing interactions between other students and between other students and their instructor... [they] can enjoy benefits that are essentially equivalent to those achieved by *direct interactors*" [59]. The approach increases learning gain via transfer of knowledge, though the gain is likely to be less than with direct learning. Vicarious learning has also been shown to increase emotional engagement, making learning experiences more enjoyable by transferring enthusiasm and enjoyment of the direct learners [23]. In this study, we use this theory to motivate the conceptual design and hypothesize the benefits of Social Constructivism for real students.

This paper presents a practical instance of these theories, manifested in the design of interface features for enhancing engagement. (The theoretical underpinning of SVD's design will be explained in detail in Section 3.1.) It also provides an empirical case that vicarious interactors may derived benefits from observing the interactions of purely virtual characters.

2.2 Increasing social interactions and engagement in video-watching experiences

There are many approaches to adding social components to make a video-watching experience more enjoyable. In educational settings, there have been attempts to use empathetic agents and timestamped comments. They range from an interface that motivates students through text messages and animated facial expressions when prompted [13] to a robot that interprets students' actions and adaptively encourages their participation in group learning [2]. The problem with empathetic agents is they are often designed specifically for a certain purpose or subject and therefore can be difficult for instructors to use and costly for students. Moreover, there are ethical concerns surrounding empathetic pedagogical agents regarding data privacy, affective privacy, and other implications of virtual relationships between humans and AI [51]. On the other hand, timestamped, textual comments, such as Danmaku, YouTube comments [20], or TrACE [18], are collected from all previous viewers of the video. Back-and-forth comments help facilitate

instructor-student and student-student interactions for utilitarian purposes [9], while hedonic and social comments help decrease students' loneliness [38]. However, the community-driven comments tend to be noisy and require hands-on moderation to guarantee relevance and quality [32]. And specifically, in the case of Danmaku, where the comments fly across the screen, they are shown to distract students from the video and require higher cognitive loads [71, 72].

In non-educational settings, reaction videos and co-watching sessions are common ways to add social presence and interactions to video watching. Reaction videos (including memes and GIF reactions) work through affective metawitnessing, where seeing other people's affective reactions can induce an impression of proximity for spatially distant audiences [53]. Reaction videos can also be costly to make since it requires real people to watch the content and record themselves in advance. Moreover, there is a risk of affective deviance, where the incongruence between the reactor's and the audience's reaction to the content can cause a feeling of aversion to the experience [60]. Finally, in co-watching sessions, chatting during intermissions has been shown to make the experience more enjoyable and engaging [69]. Danmaku comments can also be used in co-watching sessions, although they may present visual clutter [14]. For online learning, the limitation of co-watching sessions is that they require multiple audiences to attend a session synchronously, which reduces the convenience benefit of lecture videos that can be viewed asynchronously.

Additionally, researchers have explored video-making strategies and add-on systems to increase in-the-moment engagement for isolated students without utilizing social interactions. Different types of video visuals and voice-overs have been shown to affect behavioral engagement [12]. Attention cues, such as acoustic [28] and visual cues [29], have been employed to increase students' engagement at specific points in a video. Building on this, Kasperuniene et al's work on avatar-based lectures shows that complex, multi-modal cues create more engagement than single-modal ones [34]. However, the cues only evoke behavioral engagement and occasionally emotional engagement, if they are aesthetically pleasing, but they do not engage students cognitively. Many studies employ the *ICAP (Interactive, Constructive, Active, and Passive)* framework and adapt contents or learning activities based on students' learning behaviors to improve their cognitive engagement [15, 63, 68]. Agent-based approaches, such as tutoring and teachable agents, are also prevalent for improving cognitive engagement as a means for improving learning gain [8, 39]. Both ICAP-based and agent-based approaches have similar issues to empathetic agents and Danmaku comments regarding information privacy, lack of instructors' control, and potentially noisy behavioral signals.

In contrast to these solutions, SVD allows the instructors to have control over all the interactions and can make sure that they are appropriate and helpful to the students. On the student's side, the enhanced lecture videos remain accessible at any time. Moreover, since the virtual peers do not interact with the actual students, the solution circumvents the ethical concerns of intelligent agents.

2.3 Online vicarious learning

Existing vicarious learning environments for isolated online learners can be classified into four categories based on the characteristics of the direct interactors. First, a vicarious learning setting where the target students watch a dialogue between one real instructor and multiple real students. The most prominent example is a recording of a Zoom classroom. A large-scale study shows that the dialogues in such recordings significantly increase the cognitive engagement of language learners [47]. Second, a setting where the target student watches one intelligent virtual instructor interact with one real student has been explored with a tool called GuruTutor, which has been shown to reduce mind-wandering and improve behavioral engagement [42]. Third, a setting between one instructor and one instructor-authored virtual student has been explored by Nugraha et al [43]. It allows instructors to create vicarious learning scenarios by themselves, circumventing the need to schedule students and giving instructors better control of the dialogue content than with the virtual agents. Fourth, a setting where the target student learns from crowdsourced questions and answers of other online students. This option is used in a tool called YouTute, which produces a vicarious learning experience for existing educational videos based on other students' responses [49]. Despite the evident benefits of vicarious learning on learning gains, large educational platforms, such as MOOCs, do not have any features that adopt this concept [41].

SVD is a combination of these different types. It gives instructors complete control over dialogue content and expands the interaction to include multiple virtual characters to increase the similarity to synchronous classroom settings. We also expand the scope of videos from the instructor's own lectures to any online videos, so that SVD can be applied within educational video platforms, to enhance and customize pre-existing content.

3 DESIGNING SCRIPTED VICARIOUS DIALOGUES

In this section, we describe the first phase of the study. The goal of this phase is to derive the design guidelines for SVD. We based our design on theoretical investigation and participatory design. Our study focuses on four aspects of SVD:

- The types of classroom interpersonal interactions to be simulated,
- The presentation (e.g., degree of realism and interaction metaphor) of the simulated interactions,
- The level of controls the learners have over the simulated interactions, and,
- The persona and behavior of the virtual characters.

3.1 Theoretical Underpinnings and Conceptual Design

In this section, we will explain how Social Constructivism and vicarious interaction theory (summarized in Section 2.1) motivate our design of SVD and why we expect the design to be effective in increasing the learners' engagement.

The conceptual design of SVD consists of at least one MKO character (a teaching assistant, a more capable peer, or an instructor)

and multiple virtual students. Since the character's actions are completely controlled by the real instructor, who is an MKO, we can assume that the actions of this character are that of an MKO for the virtual student characters. The characters would watch the original video content together and contribute to a discussion around the subject of the video. The learner would learn both from watching the original video and observing the interaction around it.

In the first step, we apply Social Constructivism, which says that social interactions with MKOs lead to higher behavioral and cognitive engagement [67, 70]. According to Social Constructivism, we expect that the virtual students will have higher behavioral and cognitive pseudo-engagement with the subject matter.

In the second step, we apply vicarious interaction theory to model how the pseudo-experience of the virtual students will be transferred to the learner [59]. We assume that the learners will exhibit higher behavioral and cognitive engagement will be the benefits transferred from virtual students via vicarious learning. It should be noted that the vicarious interaction theory does not explicitly specify that behavioral and cognitive engagements are among the transferable benefits and whether it holds in scenarios where the direct interactors are purely virtual. Hence, our study contributes to filling this gap. The implication of our study on the theory will be further discussed in Section 6.1.

In addition to cognitive and behavioral aspects of learning, the theory also has been applied to emotional transmission between virtual entities and humans [23, 36]. We expect that observing the enjoyment and enthusiasm of the virtual characters may also have the potential to increase the learner's emotional engagement.

3.2 Participatory Design

After developing a conceptual design of SVD based on the theoretical underpinnings in the previous section, we ran a participatory design study to gather the instructor's impression and input on the concept of SVD. We conducted the study with university instructors and TAs to gather information about interpersonal interactions in actual classrooms. We based the participatory design study on the methodology of Scheepmaker et al's paper [54]. Each session consisted of an interview about their teaching and video-making experience, a brainstorming activity to elicit interactions they have seen in class, and a collaborative design activity (between one participant and one researcher). In the brainstorming activity, the researcher prompted instructors to list as many interactions they had seen in classrooms (both online and offline) as possible. Then, the instructors explained the importance of each interaction to the researcher and brainstormed how they thought the interaction could be replicated for students who watched the lecture video of the class. In the design activity, we prepared low-fidelity design sketches of SVD in advance to prompt the participants to critique and build upon them.

We recruited participants by snowball sampling within the university through lab and departmental connections to recruit instructors and TAs from across multiple disciplines and to gauge the cross-disciplinary potential of this tool. The participants are six university instructors and three teaching assistants (P1-P9) from 7

different departments across science, social science, and humanities fields. Every participant has taught 1-3 courses asynchronously with lecture videos and has offered the same courses synchronously, so they could compare how the missing social components from lecture videos affect their teaching strategies and students' engagement. Each participant was compensated for their time and effort with USD \$65.

Based on the design guidelines from the literature review and participatory design, we created a preliminary prototype of SVD. We recruited 10 undergraduate students (P10-P19) via Prolific⁴. Participants are required to be undergraduate students who are fluent in English. During the study, they watched a video on World War 1 or the mechanism of pseudorandom number generators, augmented with SVD, and provided feedback on the design and contexts in which SVD could potentially benefit. All participants were compensated USD \$9.50.

3.2.1 Data analysis. From the participatory design study, we transcribed and analyzed the data including the interview transcripts, brainstorming sheets, and session notes, to derive potential design guidelines. From the preliminary evaluative study, we calculated the average Likert scale ratings on the perceived efficacy of each component of SVD (e.g., chat messages, TTS voice, emoji reactions, and the three types of interactions as detailed in Section 3.3.1) to determine the areas of problems in the design, then we analyzed their qualitative comments to understand the nature of the problems. Finally, we refined and added to the initial design guidelines to account for the students' needs and perception of the design, before finalizing the design of SVD.

3.3 Findings

3.3.1 Classroom interactions that constitute an effective SVD. Our SVD design incorporates three types of interactions repeatedly mentioned by the participants as conducive to enhancing students' engagement and replicable in online classroom settings. The interaction types are:

- (1) *Q&A.* This is an informative interaction that helps students gain more insights into the course content. It benefits all students, including those who are less active or confident in raising questions can benefit from the instructor's response through vicariously observing the Q&A between the student who asked the question and the instructor (P1, P2, P3, P5). The instructor can also initiate this interaction to test the student's knowledge with a question or a prompt for the students to brainstorm for ideas or examples based on a topic or concept (P4, P7). Instructor's questions are valuable because it pushes the students to verify and deepen their understanding of the class content. Moreover, students can get a sense of their standing in class and "feel reassured if they know their mistake is common" (P2).
- (2) *Social chit-chat.* This is a social and emotional interaction to "engage the student's attention and bring up the energy of the class" (P9) and to build rapport within the class, between instructor and students and also between students themselves (P8). It usually consist of greetings, non-academic

⁴<https://www.prolific.co/>

conversations, and other phatic interactions. It often happens at the beginning and the end of the lecture. We found from the preliminary evaluation that, in SVD, social chit-chat SVD could feel tedious and irrelevant (P11, 15, 18, 19), since the students did not perceive the characters as real people they can socialize with. Hence, we kept the social chit-chat short and related to the learning content of the video.

- (3) *Nonverbal reactions.* the nonverbal reaction for social and emotional purposes (e.g., stickers and emojis) has become prevalent during the COVID-19 pandemic (P5). It is valuable because it can stand in for emotional expressions (e.g., smiling, laughing, rolling eyes), which are important cues for instructors and peers to notice and empathize with "students who look like they are flagging" or "some concept or detail has not been understood" (P9).

3.3.2 Design guidelines. From the participatory design study, we derived three design guidelines (DGs). Here we elaborate the guidelines and how they supported our final design of SVD experience presented in the Section 3.4.

DG1. Balancing virtuality and realism Too much realism in the SVD's appearance is problematic. This is because the script of SVD is written by the instructor, in advance, and therefore contains well-thought out questions, which are atypical and more advanced than what real students are likely to ask in synchronous classrooms. Paired with a realistic appearance of the characters who would ask such questions, students could be deceived into comparing their performance to that of the characters' (P7). Hence, it is recommended that characters' visual representations in SVDs be obviously virtual, to reinforce that the SVD is a script that is intended only to simulate a real class. We should note, however, for voice narration of the dialog, a synthetic virtual quality can be problematic. In particular, our initial design implemented voice narration of the virtual characters using text-to-speech (TTS), but many participants (P12, P15, P18, P19) disliked the synthesized voice because they found it "robotic", "annoying", and "uncanny". This aversion to the unnaturalness of the voice disengages the participants emotionally and therefore affects their emotional engagement.

DG2. Create relatable characters with individuality and diverse identities It is important for the learner to be able to "relate to the characters so that they can be more engaged [with the characters' interactions]" (P5). To engage a wide range of learners, the characters' identities, such as race and gender, must be diverse and inclusive (P4). P8 also noted that, beyond their personalities, it may be interesting to give the characters distinguishable characteristics, such as hobbies or relationship with each other, since these are the topics that allow students to discuss and bond over in real classrooms.

We should caution that these three aspects can conflict with each other, and achieving a balance between them is a challenging design problem. For instance, P4 said that that we had to be careful to avoid stereotypes when including other traits, such as proficiency in the course. Moreover, making the characters representative of diverse racial, gender-based and other profiles means that a learner may not be able to relate to all characters equally.

DG3. Allow learners to self-pace through the dialogue. The initial design of SVD was a video playback, as if it were a recording of a Zoom classroom. During the preliminary evaluation, participants found two issues with this design. On one side, some participants reported that the interaction is too slow since "humans can take forever to explain to you an idea." (P13) However, some other participants found the experience to be passive — they did not feel the need or have enough time to think actively on the questions before seeing the answers, as P11 commented: "The Vicarious Dialogues doesn't require you to think [...] There's a pause, but even then the [virtual] students were really fast [at answering]. A lot faster than me. So, I just read their answers." This shows that learners have different reading speed and different ways of absorbing the dialogues (just reading vs. thinking actively on it.) So, the interface of SVD should provide a convenient way for learners to navigate the content at the pace they prefer.

3.4 Design

In this section, we will describe the final design of SVD after incorporating the design guidelines. The conceptual model is a Zoom classroom, which is reflected in the layout and style of the audience window, video, and chat window (see Fig. 1.)

Following DG1, all interactions between the virtual characters happen as text messages and emojis when the video is paused, so that they are temporally and spatially distinguished from the real component, i.e., the original video. The temporal separation is highlighted in the progress bar, where the continuous video and discrete SVD components are represented with lines and dots, respectively. We also use a cartoon effect on the characters' profile pictures to highlight the virtualness of the interaction. Finally, we stress that the learners are informed prior to using the interface that the dialogues are completely scripted and not reflective or based on any real classes.

DG2. informs our script authoring process. There are three distinguishable characteristics that make up the characters' personas: their genders, races, and interaction frequency. We represent the characters' various gender expressions and racial identities via their profile pictures and names. We chose the frequency of texting a response versus using an emoji because it was easily recognizable in a texting scenario than personality traits. Some virtual students will type longer answers and frequently ask follow-up questions in the chat, but some will type only phrases or words, not ask follow-questions, and use emojis when possible. We are also careful not to script a virtual student to perform obviously better or worse than others, so we do not create or enforce any stereotypes related to academic performance.

Finally, according to DG3, we made SVD an interactive interface, where the learner has to click the "Click to continue" button for each message to appear one-by-one. We also give the learner control to navigate back-and-forth within and across dialogues with the dots in the progress bar. Each dot is clickable and navigates the learner to a specific text bubble or emoji reaction. Through these features, the learner can self-pace in the way that is most conducive to their learning.

4 EVALUATION METHODS

To evaluate the design from Section 3, we ran a comparative evaluation of SVD against the baseline (a direct-learning, non-social, video augmentation method). According to the definition of engagement presented in Section 2.1, we measured three types of engagement in this study: emotional, behavioral, and cognitive. Additionally, we also measured the learning gain, which is an expected aftereffect of all three types of engagement.

4.1 Baseline system and experimental materials

We selected in-video quizzes as the baseline since they are a familiar activity for students and can be controlled for fairness against SVD in three ways. First, it is a direct-learning and non-social set-up, which contrasts with the vicarious and social set-up of SVD. Second, it does not differ from SVD regarding the amount of learning content it conveys and the timing of its augmented content. Lastly, SVD and in-video quizzes require similar amount of input and effort from instructors. (We discuss the feasibility of automatically converting quiz prompts to SVD scripts in Section 6.5.) For implementing the baseline condition, we adopted EdPuzzle⁵ because its learning interface was minimal and intuitive, while other similar tools come with many extra features that may distract the learner and interfere with the objective of the study. Moreover, it is a stable and easily accessible web application. Example screenshots of EdPuzzle are shown in Figure 2

To choose the videos for the prototype, we followed four criteria: (1) the videos must be on the same subject for topical homogeneity; (2) the videos must be the same length, and they should be 6-9 minutes long after augmentation with EdPuzzle and SVD for maximum engagement [27]; (3) the videos must not require other videos as prerequisites; and (4) the videos must not be on a widely known topic. We searched the collection of videos on Khan Academy⁶, a popular online learning channel, to find videos that matched the criteria. In the end, we chose a video on World War I⁷ and a video on the Napoleonic War⁸.

To script the augmented content, we trimmed three sections from each video and replaced them with augmentation content with the exact same content. We wrote the script according to DG2 (3.3.2). We promoted fairness between two videos by structuring the dialogues the same (e.g., the content of the greetings and farewell dialogues, the number of times a virtual student answers incorrectly, the length of instructor's video recap before starting a discussion). We also scripted each dialogue to contain no more than 15 chat bubbles or emojis to ensure that the dialogues do not disengage the learners from the video. Then, we created the pre- and post-test questions (the pre- and post-tests were the same). Each test consisted of five memorization questions and five application questions for capturing shallow and deep cognitive activities, respectively. We controlled for the tests of both videos to draw the same number of questions from the augmented content versus the original video, as well as the same number of multiple-choice and short-answer questions.

⁵<https://edpuzzle.com/>

⁶<https://www.khanacademy.org/>

⁷<https://youtu.be/IyoUWRARhQ>

⁸<https://youtu.be/ufymsKi9S3U>

4.2 Experiment Setup and Procedure

We designed a 2 x 2 factorial experiment (Condition: Edpuzzle and Vicarious Dialogues; and Videos: World War I and the Napoleonic War).⁹ Condition is within-group variables. The pairing Condition x Video is between-subjects. We controlled the content and timing where the Q&A augmentation will appear in the video across conditions. We also controlled the SVD setup (one TA and three students) across the videos. The participants watched one video with SVD, and watched another video with EdPuzzle. The pairing of video and augmentation methods is counterbalanced. The order of presenting the augmentation methods is also counterbalanced.

We measured 4 dependent variables. Before and after watching each augmented video, the participants completed the same *test* (Appendix A) to measure their learning gain. After completing the post-test, the participants filled out a *questionnaire* (Appendix B) that measured their behavioral, emotional, and cognitive engagement on a 7-point Likert scale. Participants were informed in the beginning that the pre- and post-tests for each video contained the same questions. At the end of the experiment, the participants filled out a post-task *survey* (Appendix C). Within 24 hours after survey completion, we sent unique follow-up questions to expound on their survey responses. Each participant was compensated USD \$11.50.

4.3 Engagement questionnaire

The questionnaire for measuring engagement was derived from four existing questionnaires: Situational Interest scale [11], Original Cognitive Engagement scale [26], Engagement vs. Disaffection scale [56], and Situational engagement scale [66]. We categorized the questions based on the definition each type of engagement as described in Section 2.1. Then, we removed the questions that are not relevant to the context of the study, including:

- Questions that address student's self-efficacy (e.g., "I understood the lesson well.")
- Questions about self-regulation strategies (e.g., "I used rewards to help myself study.")
- Questions about a course (e.g., "I compare and contrast different concepts") and school ("I try hard to do well in school.")
- Questions about the student's active participation in class (e.g., "When I'm in class, I just act like I'm working.")

After this, we removed repeated questions (e.g., "I enjoy the lesson" and "the lesson was enjoyable"). For similar questions (e.g., "I concentrated during the lesson", "I focused during the lesson", and "I was attentive during the lesson"), we picked either (1) the statement with the highest factor loading; or, if the original survey did not provide that information, (2) the statement that was repeated across multiple surveys. Finally, we adapted some questions to match the context of the study as follow:

- "Activity" in Situational Interest scale was replaced by "Lesson"
- "Class" in Situational Engagement scale was replaced by "Lesson"

⁹Link to the working demo: <https://ttanprasert.github.io/SI-interfaces/interface-SVD-final.html>

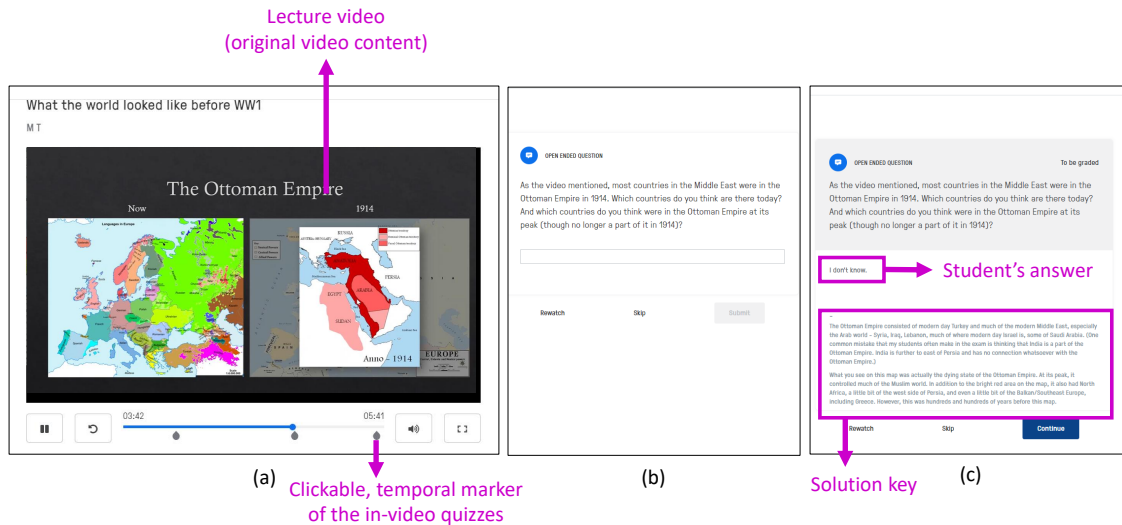


Figure 2: Screenshots of EdPuzzle’s learning interface. (a) the progress bar of the lecture video has markers to indicate when the in-video quiz will occur. The student can click on a marker to skip to the corresponding quiz. The lecture video is paused at the marker, and (b) the quiz interface would appear on the right of the video. Student has to submit an answer to see (c) the solution key to the question.

- “Underline the reading material” in Original Cognitive Engagement scale was replaced by “take notes when watching the video”
- The emotional effect from interpersonal interaction from Engagement vs. Disaffection scale was adapted to be the “real student’s” emotional effect from witnessing the “virtual character’s” interpersonal interaction.

4.4 Participants

We recruited and ran the study with 40 participants (P20-P59) via Prolific. Prolific has more than 130,000 users worldwide¹⁰ and has been used for participant recruitment in human-computer behavioral studies in recent years [33, 46, 52, 62]. The participants were undergraduate students and were fluent in English. 18 participants identified as men, 21 as women, and 1 as nonbinary. The average age of participants is 25.3 (S.D.=6.88). Participants were from 30 different programs/fields/departments and had various experience with Zoom classrooms (6 had never taken a class on Zoom, whereas 22 had taken more than 6 classes via Zoom.)

4.5 Data analysis

We calculated learning gain using the formula for normalized change proposed by Marx and Cummings [40], so the resulting learning gain ranged from -1 to 1. Then, to account for the possible effects of other variables, we fitted mixed effects models to analyze the four dependent variables.

A standard linear regression model for the learning gain (LG) outcome has the form:

$$LG = \beta_0 + \beta_1 \times SVD + \beta_2 \times WWI + \gamma_j \times participant_j + \epsilon_i \quad (1)$$

¹⁰<https://www.prolific.co/>

where SVD and WWI are binary variables representing the intervention and video, respectively, that a participant watched (e.g., SVD=0 represents EdPuzzle condition and WWI = 0 represents the Napoleonic War video) and their coefficients represent the fixed effects. Each unique participant j is given their own coefficient γ_j to account for the random effect arising from dependency of observations (since Condition is a within-group variable). The model also includes an error term ϵ_i that represents variation in the values unexplained by the other variables included in the model. The β_1 is the slope from EdPuzzle to SVD, so in order to verify if participants’ has significantly higher learning gain with SVD than with EdPuzzle, we checked whether the β_1 parameter is positive and significantly different from 0 ($p < .05$). On the flip side, a significant and negative β_1 would mean that participants has significantly higher learning gain with EdPuzzle.

As for the engagement score, each of the three engagement scores is an aggregated 7-point Likert scale score: behavioral engagement score range from 4 to 28; emotional engagement score range from 7 to 49; and cognitive engagement range from 10 to 70. Despite the raw responses being on Likert scales, aggregated Likert scale scores can be treated as continuous [58], and thus, we can use a linear mixed effects model for them, too. The interpretation of the model is the same as for learning gain.

Finally, we triangulated the participant’s response to post-task surveys and follow-up interview questions with their engagement scores and learning gain to verify the effect of our design guidelines, specify the benefits of SVD, and extrapolate how the design of SVD interacts with participants’ learning styles and preferences.

5 FINDINGS

During data analysis, we observed that the preference for SVD vs. EdPuzzle is polarized, as shown in Fig. 3. Of all 40 participants,

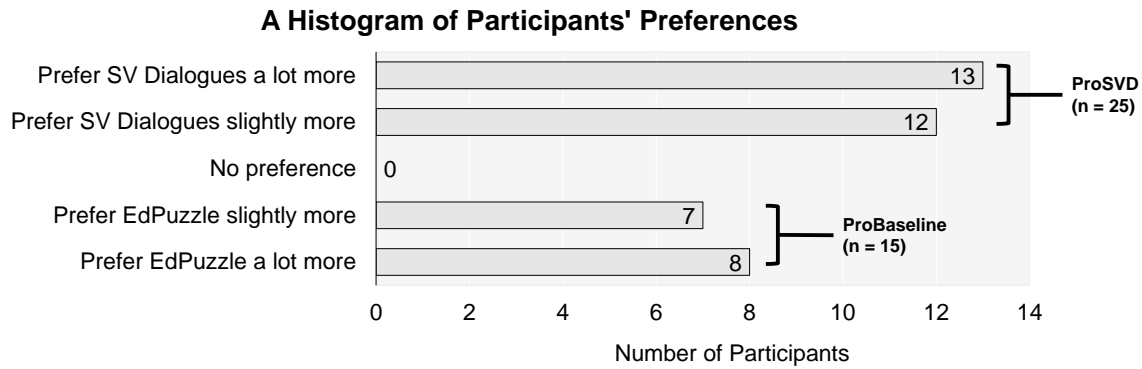


Figure 3: A histogram showing participants perfectly bimodal responses to the question "Between EdPuzzle and virtual characters' interactions [SVD], which one do you prefer?"

everyone had a preference; no one picked the “no difference” options. 25 participants preferred SVD, while 15 preferred EdPuzzle. Since the preference responses showed potential connection with all dependent variables, we decided to use this variable as an independent variable in our data analysis. Hereon, we will denote the first group “ProSVD” and the second group “ProBaseline”.

We observed from their qualitative feedback that there were three potential factors for their preference:

- (1) Conciseness (ProBaseline) vs. interpersonal interaction (ProSVD): participants who preferred SVD valued and enjoyed interpersonal interaction reminiscent of synchronous classrooms, whereas participants who preferred EdPuzzle valued straightforward text and viewed prompting questions and roundabout way of arriving at an answer meandering, taking their focus away from the content.
- (2) Uncomfortable (ProBaseline) vs. comfortable (ProSVD) with characters' virtualness: participants who preferred SVD were comfortable with the virtual characters and the knowledge that the dialogues were scripted, whereas participants who preferred EdPuzzle found the virtual interaction uncanny and distracting. This contrast is most stark in the feedback about social chit-chat.
- (3) Unwilling (ProBaseline) vs. willing (ProSVD) to constantly interact with the system: participants who preferred EdPuzzle found clicking through the messages one-by-one tedious, distracting them from the content of the interaction itself.

In the following chapters of the findings, we will go into details about what participants reported regarding these factors and how each factor affected the results of different engagement types.

5.1 Emotional engagement

The most significant result is the emotional engagement score. The difference in emotional engagement scores is significant when distinguished by preference, and they correspond with the direction of the preference.

- ProSVD: $\beta_1 = 5.936$, p-value = 0.002 **, power = 92.90%, $\eta^2 = 0.35$

- ProBaseline: $\beta_1 = -4.875$, p-value = 0.008 **, power = 86.00%, $\eta^2 = 0.42$
- ALL: $\beta_1 = 1.895$, p-value = 0.201 (n.s.)

Our analysis of the qualitative feedback illuminates two possible reasons why SVD increases the participants' emotional engagement: a sense of social connectedness and characters' relatability.

The first reason is that SVD creates a sense of social connectedness and a sense of being in a synchronous classroom. When asked to pick which system provided a sense of social connectedness, 37/40 picked SVD over EdPuzzle, while the remaining 3 picked the “no difference” option. The responses were mainly due to the interface layout and the social chit-chat. Many participants said that the Zoom-like layout reminded them of their classroom, because “this is the kind of set up I had during lectures whilst we were going through the worst parts of the pandemic.” (P29) The non-academic chit-chat at the beginning of the class also contributes to the sense of social connectedness and consequently emotional engagement, as P28 pointed out:

“Jumping straight into the educational content makes you feel like the professor is only there to do the bare minimum of their job and doesn't actually care about the students. Therefore, the conversation of people's weekends before the lecture actually began felt as though I was more a part of something and I also felt like there was a sense of belonging. This enhanced my enjoyment of the lecture because it felt like a friendly environment and not a judgy one...”

A sense of social connectedness is closely related to the first preference factor (conciseness vs. interpersonal interaction), and so it may have a positive or negative effect on emotional engagement depending on the participants. For P30, who preferred straightforward and concise lessons, SVD was detrimental to their emotional engagement: “I want to learn, and I want to just get on with it, without distractions. [...] Fake conversation and such just act as distractions that remove me from that learning mood.” On the other hand, those who value interaction found that social connectedness made them “interested in the topic that was being discussed” (P29), and even if they asked uninterested questions, they would “still

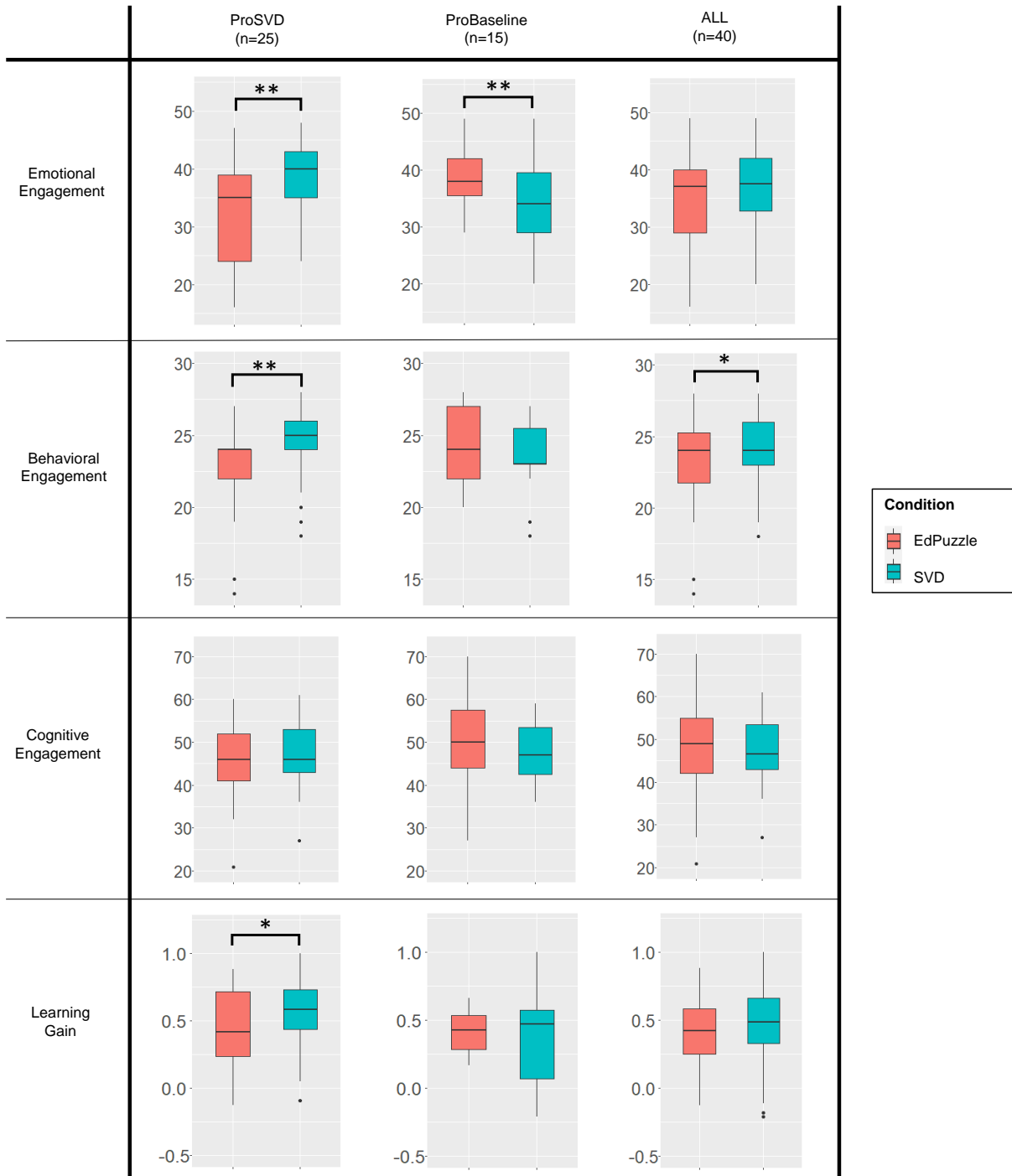


Figure 4: Box plots of learning gain, behavioral, emotional, and cognitive engagement for different subject groups: all participants (n=40), for only participants who preferred EdPuzzle over SVD (n=15), and for only participants who preferred SVD over EdPuzzle (n=25). The statistically significant comparisons are marked with two asterisks, the comparison with potential significance (p-value between 0.05 and 0.06) are marked with one asterisk.

interested nevertheless as I do enjoy the company of other students.” (P37)

The second reason is that participants who were comfortable with the virtualness of the system found the interaction natural and related to the virtual students. This reason is closely connected to the second preference factor (comfort with virtual characters) and a direct result of vicarious learning theory and our design guideline to distinguish between the real and virtual components (DG1). P34 found the interaction natural and so the learning experience was “enjoyable because you can learn along, as frequently students may have similar questions to each other.” And even if she compared herself to the characters, it was a “motivation to better your knowledge.” In particular, many participants found the students’ wrong answers realistic and the TA’s attitude supportive and encouraging, making them comfortable and interested in the dialogues:

“You can see there are others who would also get this wrong and it’s nothing to feel dumb about, whereas, in EdPuzzle, it was [a] plain long text which only made the reading dull and less interesting, constantly if someone were to get something wrong, it would be a real down breaker and demotivate them.” (P33)

It should be noted that there were a few outliers who compared themselves to the characters and felt pressured while watching SVD, such as P21, who explained that “I guess even though I wasn’t replying to the questions myself, I still mentally treated it as if I could be asked one.” We will further discuss the spectrum of learning styles that are suitable for SVD in Section 6.1.

5.2 Behavioral engagement

We found a trend result of SVD increasing behavioral engagement for ALL participants, but the difference was particularly significant for those who preferred SVD. As for those who preferred EdPuzzle, unlike emotional engagement, SVD produced no negative effect on their behavioral engagement.

- ProSVD: $\beta_1 = 1.577$, p-value = 0.002 **, power = 92.20%, $\eta^2 = 0.34$
- ProBaseline: $\beta_1 = -0.4375$, p-value = 0.494 (n.s.)
- ALL: $\beta_1 = 0.816$, p-value = 0.0501 *, power = 52.40%, $\eta^2 = 0.10$

In qualitative analysis, we attended participants’ reports specifically on two indications of behavioral engagement: concentration and attention. We found two possible reasons that participants report as effective for their behavioral engagement: the requirement to self-pace and the dialogue format of the augmented content.

The first reason is that SVD may have forced participants to self-pace their learning at a micro level, keeping their attention on the learning task and not allowing them to zone out. They had to click through each message one by one. This, therefore, “encourage[s] involvement and participation, rather than being [a] passive bystander” (P58). This control also allowed those who already knew the content or found the conversation meandering to speed up the interaction at will: “you don’t turn off and fall asleep since you can skip through it.” (P56) It should be noted that, although this reason is closely related to the third preference factor (willingness to constantly interact with the system), those who were unwilling to interact with the system might still find that the

system kept their attention and concentration, even though they did not enjoy the experience.

Secondly, participants who valued interpersonal interaction reported that the dialogues helped them concentrate more on the answers to the questions, compared to paragraph texts. The dialogues contrasted with the lecture and attracted their attention, since “you don’t just get used to one thing happening and let your mind wander, there’s always something a bit new happening.” (P26) Moreover, P22 reported that the scenarios between the virtual TA and students raised their focus, curiosity, and anticipation for the answers. Finally, the dialogue format could affect the participants’ desire to participate, thus increasing their focus on the augmented content and making them think on the topic more elaborately, as P55 explained:

“In the virtual conversations, although I couldn’t actually respond in the chat, I felt more as if I was part of a conversation and in my head, I thought out what my answer would be to a question even if it was directed at the other characters. For the EdPuzzle questions, I had less of a desire to elaborate and think deeply about my answer and the answers I gave. I would write down simple 1- or 2-word answers to quickly move on through the lecture and check if my knowledge was correct.”

5.3 Cognitive engagement

We found no significant difference, positive or negative, for any other comparison of cognitive engagement scores.

- ProSVD: $\beta_1 = 1.3782$, p-value = 0.119 (n.s.)
- ProBaseline: $\beta_1 = -2.2143$, p-value 0.197 (n.s.)
- ALL: $\beta_1 = 0.0363$, p-value = 0.965 (n.s.)

This was expected from the theoretical underpinning. Direct learning like EdPuzzle, where students are forced to think, is generally more cognitively engaging than vicarious learning like in SVD. However, the simulated dialogue likely increases the cognitive engagement compared to the monologue form of knowledge, as stated in Social Constructivism, so both systems ended up producing no significant difference.

We found three interesting observations based on participants’ qualitative feedback that could explain how SVD affected participants’ cognitive engagement: (1) vicarious learning alleviates performance anxiety; (2) self-pacing helps eliminate cognitive overload; and (3) students think more actively on questions in EdPuzzle, but think more actively on the answers in SVD.

Due to vicarious learning, SVD reduced the performance anxiety associated with synchronous classrooms without decreasing cognitive engagement. According to Sutton et al, the benefits of vicarious interaction depend on learning styles, and those who think actively via vicarious interaction are usually uncomfortable participating directly in synchronous classrooms [59]. Even though there was no synchronous classroom condition to compare in this study, many participants mentioned that SVD made them think actively on the question, as much as with EdPuzzle, while feeling less anxious or pressured. This sentiment was best reflected in the following quote by P35:

“It definitely encourages me to try and think and get involved rather than just waiting to see the answer. It makes me feel almost like I am the student and the prompts from the TA encourage me to think and expand on my answers, and therefore become more memorable because I am having to think about them deeply. It feels very realistic and like[s] what a TA or a lecturer would do in an education setting... I actually felt like I was involved in the conversation, I think it's a great way to get students to join in, and a massive advantage for individuals who are shy or do not have a lot of confidence about them[selves] - it's less pressure or 'on the spot' but also makes you feel involved at the same time.”

Finally, we found that participants thought more actively on questions in EdPuzzle, but they thought more actively on the answers in SVD. Although both systems framed the knowledge in Q&A format, EdPuzzle forced the participants to answer, so the participants thought on the questions hard and then memorized the answer but not thinking actively further on it (P37, P41, P47). P43 explained that this was because “there was already an opportunity for me to think about the point it was reinforcing before receiving the information.” On the other hand, with SVD, participants were not forced to answer the question, so many of them immediately proceeded to see the answer. However, since the answer was slowly built into a dialogue, they “see what classmates have to say and reflect on their answers and the information for myself” (P44). P38 said that this was possibly another benefit of vicarious learning, “Being an observer gives me room to think and analyze other people's responses.” This observation illustrates that the instructor's purpose whether to stress the question or the answers could inform if SVD would be useful in the application context.

5.4 Learning gain

For learning gain, which is the consequence of all three types of engagement, we found a trend of an increase for ProSVD participants and no significant difference for ProBaseline participants:

- ProSVD: $\beta_1 = 0.1312$, $p\text{-value} = 0.062^*$, $\text{power} = 52.00\%$, $\eta^2=0.14$
- ProBaseline: $\beta_1 = -0.0902$, $p\text{-value} = 0.305$ (n.s.)
- ALL: $\beta_1 = 0.0489$, $p\text{-value} = 0.375$ (n.s.)

For both subject groups, we did not observe significant learning gain differences by question types – memorization vs. conceptual application questions.

6 DISCUSSION

6.1 Result interpretation and implications

Through this study, we found that SVD has positive effects on the engagement of a specific group of learners, that is, learners who valued interpersonal interaction, are comfortable with virtual characters, and like actively controlling their learning pace. Quantitative results show that learners with such characteristics experience high emotional engagement with SVD due to the enacted sense of social connectedness and the relatability of the virtual characters. These learners also experience increased behavioral engagement

with SVD. This increase is empirical evidence that the vicarious interaction theory may apply when learning vicariously through virtual characters. It also supports our hypothesis that behavioral engagement is a benefit that can be transferred from direct to vicarious interactors and refines the gap in the vicarious interaction theory as we mentioned in Section 3.1. Although we do not find a significant difference in cognitive engagement, we saw evidence that SVD decreases performance anxiety and may lead to more active thinking in some scenarios. Finally, we see that learners who prefer SVD are likely to have a higher learning gain with it as a second-degree effect.

Although we present the factors influencing preference between EdPuzzle and SVD as binary variables as indicated in Section 5, the two systems are two ends of a spectrum. For each factor, there are SVD parameters that could be adjusted or new attributes that could be included to suit the learner's needs.

- **Conciseness vs. Interpersonal Interaction:** depending on a learner's learning style or situation, they should be able to choose which presentation to learn from. Instead of a separate system, SVD should be an alternative for presenting Q&A, integrated into a system that already has a more straightforward presentation (e.g., EdPuzzle). Moreover, the middle-ground options can include existing works that utilize social interactions to increase educational engagement such as a direct Q&A with virtual TA's presence [42], vicarious Q&A with only one virtual student [43], etc.
- **Characters' virtuality:** many parameters can affect one's perception of virtuality and the effect of virtuality on the learners. Previous research shows that interactive traits, such as gaze [25], speech pattern [50], gesture, and mannerism [35] of a virtual agent can increase the agent's social presence and the user's trust in the agent. Although SVD is not an agent-based system, our data shows many instances where participants who enjoyed SVD expressed that they “like” some characters over the other. To increase the relatability and likability of the characters, future work can consider adding the interactive attributes to the characters. However, we also need to be careful to follow DG1 and not make the characters too real. Another important consideration from our data is that some learners may view emotional or social attributes of interactions as hindrance to their learning process. A clear example of this is the social chat at the beginning of the class, which is shown to have polarized feedback. We recommend that every virtual attribute other than the characters' identities and interaction frequency should be customizable by the learners instead of a fixed component of the experience.
- **Pacing through the dialogues:** self-paced learning has been shown to improve student's engagement and metacognitive controls, especially in traditionally passive learning environment such as watching lecture videos [6, 45]. Our findings agree with the existing literature. Specifically, participants expressed that the self-pacing feature accommodates learners with various reading speeds well. However, we also found that too fine-grained control of the pacing could potentially lead to cognitive overload and disengagement in

some participants. Based on this, we recommend that the interaction scheme could be improved such that, at each point, the learner can choose whether to open one more message or to play the thread automatically, and they can pause the automatic play to revert to clicking one-by-one whenever they want.

6.2 Vicarious Dialogues vs. Interactive Virtual Agents

As mentioned in Section 1, SVD focuses on dialogue presentation and differs from interactive agents, which simulate dialogues based on students' behaviors and performances [8, 39]. Therefore, some effective characteristics of SVD are not exclusive to instructor-authored scripts and can be applied to virtual agents to "structure" and "present" the instructional interaction more engagingly. These characteristics are:

- **Vicarious learning:** according to our qualitative findings on cognitive engagement (Section 5.3), the passive nature of vicarious learning is useful for teaching the thought process (e.g., thinking on each step of building up an answer, seeing the possible wrong directions) and learning attitude (e.g., it's okay to answer incorrectly). A teachable agent can also be used as a virtual student to demonstrate such behaviors to the learners.
- **Instructor's control:** on top of the information privacy issue, agents could make biased comments or deliver the content in inefficient ways [3], whereas instructors already have the content and experience of how to deliver it effectively to the learners (e.g., which mistakes to highlight, which leading questions to ask). Virtual agents can also leverage this knowledge from instructors to deliver the lesson more effectively.
- **Multiple-people classroom setting:** Our findings in Section 5.1 shows that having multiple people is good for portraying multiple perspectives and creating a sense of social connectedness. However, designing multiple roles is a difficult social design problem, as the designer has to pay attention to the group dynamic arising from interactions of multiple identities, stereotypes, and personality traits, as we present in Section 3.3.2. These guidelines could be transferred to designing virtual agents, as well as social robots and multi-robot systems [57].

6.3 Scripting an effective dialogue

Although the design of SVD contributes to the results, as we explained in Section 5, it is undeniable that the quality of the script also affects the results. In this section, we will discuss the script's effects on participants' perception of SVD and the challenge of writing a good script for SVD. It should be noted that the contents and timings of each dialogue/Q&A are the same across the two conditions to ensure fairness of the comparison. Therefore, we will only discuss the quality of the script's prose, not its content.

We found that the script mainly affects the participants' perceived naturalness of the interaction, which is a factor for emotional

engagement. Too formal language ("the way students phrased questions did seem exaggeratedly formal compared to what I have actually observed in a Zoom environment, even in rigorous university courses." - P43) and fake friendliness ("trying to make something fake more human [...] they will just feel awkward and clunky" - P21) can make the virtual characters especially uncanny for some learners. These factors are a subset of the factors contributing to naturalness of a chat bot or voice user interfaces [37], but with additional consideration for the social dynamics between multiple virtual characters that is unique to SVD.

We spent a long time revising the scripts for each video, and we reflected that there are many dimensions to scripting the dialogue. The two most important dimensions are the conciseness of each bubble and the discussion length.

- **Conciseness:** We would like each chat bubble to be concise so that learners can grasp the knowledge easily without too much distraction from filler words (e.g., "well, actually..." or "I kind of think..."). However, we also want to make it lively and natural. Finding the right balance is a challenging problem, which depends on the content and purpose of the dialogues.
- **Discussion length:** We would like to have a back-and-forth discussion for knowledge construction with multiple chat bubbles. But how many leading questions should the TA ask before revealing the answer? How much should the virtual students figure out on their own, and how much should the virtual TA tell them? A long discussion will help the learner construct the knowledge thoroughly but will also take the learner's focus away from the lecture. This balance is also difficult to determine, similarly to the conciseness problem above.

6.4 Students' perception of SVD's authenticity

On top of the realism in the characters' representation dictated by DG1, another variable that could affect the students' perception of SVD's authenticity is the source of dialogue content. That is, even if the students are constantly aware that the dialogues are scripted by the instructor, the students may still suspect that the instructor bases the script off real dialogues from a class they previously taught. This nuanced variable was not resolved in our design, but it generates both ethical and practical concerns.

The ethical consideration is whether the perceived authenticity creates deception. From our participatory design, many instructors express that they would use SVD to present frequently asked or interesting questions from their previous classes. On one hand, a question may truly come from a real student, but on the other hand, the collection of questions are not reflective of a real class, as they are intentionally compiled by the instructors. Therefore, we cannot directly apply existing framework [10, 65] to judge if SVD is deceptive. It is our hope that, by informing the students that the script is written by the instructor (as noted in Section 3.4), the students will at least be aware of the instructor's potential influence on the dialogues and not believe that the dialogues are a complete copy of a past class.

On the practical side, it is crucial to consider how this variable may affect the students' engagement and self-efficacy. There is a

tension between these two values. On one hand, the more authentic the students perceive the dialogues to be, the higher engagement (e.g., interest, attention) they are likely to have [48]. On the other hand, since the instructors intentionally select interesting or intelligent questions for the script, the more authentic the students perceive the dialogues to be, the lower self-efficacy they are likely to have [19]. Although we tried to compromise these two contrasting effects when designing SVD, we did not measure relationship between individual student's perception of dialogue authenticity, their engagement, and self-efficacy. We recommend this as a potential area for future work.

6.5 Needs and feasibility of developing an authoring tool

With the challenges outlined in the previous section, we see that writing the scripts can be laborious and time-consuming for the instructors. During our participatory design study, instructors expressed concern about the authoring process of SVD. Turning existing Q&A into natural-sounding dialogues takes time, which they are reluctant to spare. For SVD to be practical, it needs an authoring tool to mitigate such concerns.

Below, we demonstrate the feasibility of using a natural-language processing model, specifically GPT-3, to perform the task of turning standard Q&A texts into dialogues between one TA and three students. Fig. 5 shows an output from GPT-3 after training with 5 examples (more examples in Appendix D). The examples show that it is feasible to create an authoring tool to turn any Q&A in EdPuzzle form into SVD dialogues without any extra effort required from the instructors.

6.6 Limitation

There are three main limitations in this study: the videos, the quizzes, and the scripts. The videos are all on one subject, World History, which may affect the generalizability of the results. Specifically, the different types of knowledge and knowledge activities associated with different subjects may induce different type of cognitive engagement in the learners. For example, learners would mainly perform memorization when learning World History but would perform more application when learning Mathematics. The lengths of the videos are also very short, so participants may not have a chance to forget or get disengaged from the content. For the quiz, the number of questions in each quiz is small, so it cannot capture big differences in learning gain, and the multiple-choice questions may also introduce noise in the measured learning gain. Finally, as discussed above in Section 6.3, the engagement with SVD partially depends on the quality of the scripts, but we were not able to quantify this effect and assumed that the scripts for both videos had the same quality and effectiveness.

7 CONCLUSION

In this paper, we proposed SVD as a technique for increasing isolated, video-based learners' engagement by scripting vicarious scripted instructional dialogues between a virtual TA and students around an educational video. We designed SVD based on participatory design with instructors and preliminary feedback from

learners. Our evaluative study shows that the effects of SVD depends on individual need for social connectedness, perception of virtuality, and learning strategies. Those who prefer SVD had significantly higher emotional and behavioral engagement with SVD compared to EdPuzzle. SVD also contributes to cognitive engagement in critical discourse. We positioned SVD in relation to other social video-watching and vicarious learning experiences and discussed how the insights from SVD can be generalized to improve other methods of augmenting video-based learning. Future research should look into other contexts where SVD could be beneficial and design an effective SVD's authoring tool for instructors.

ACKNOWLEDGMENTS

This work was supported by the NSERC Discovery Grant and CREATE programs, as well as the KEIT grant funded by the Korean government (MOTIE, No. 20009940). Also, this work was partially supported by the Huawei-UBC Joint Lab on SoC and AI Research Program. Finally, we would like to thank the members of ViDeX lab, D-lab, and MUX lab (UBC) for the valuable discussions and feedback.

REFERENCES

- [1] Oqab Alrashidi, Huy P Phan, and Bing H Ngu. 2016. Academic Engagement: An Overview of Its Definitions, Dimensions, and Major Conceptualisations. *International Education Studies* 9, 12 (2016), 41–52.
- [2] Patricia Alves-Oliveira, Pedro Sequeira, Francisco S Melo, Ginevra Castellano, and Ana Paiva. 2019. Empathic robot for group learning: A field study. *ACM Transactions on Human-Robot Interaction (THRI)* 8, 1 (2019), 1–34.
- [3] Ryan S Baker and Aaron Hawn. 2022. Algorithmic bias in education. *International Journal of Artificial Intelligence in Education* 32, 4 (2022), 1052–1092.
- [4] Albert Bandura. 2001. Social cognitive theory: An agentic perspective. *Annual review of psychology* 52, 1 (2001), 1–26.
- [5] Jack Barokas, Markus Ketterl, and Christopher Brooks. 2010. Lecture capture: student perceptions, expectations, and behaviors. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*. Association for the Advancement of Computing in Education (AACE), 424–431.
- [6] Romiro Gordo Bautista. 2015. Optimizing classroom instruction through self-paced learning prototype. *Journal of Technology and Science Educaton (JOTSE)* 5, 3 (2015), 184–193.
- [7] Mobina Beheshti, Ata Taspolat, Omer Sami Kaya, and Hamza Fatih Sapanca. 2018. Characteristics of Instructional Videos. *World Journal on Educational Technology: Current Issues* 10, 1 (2018), 61–69.
- [8] Gautam Biswas, Krittaya Leelawong, Daniel Schwartz, Nancy Vye, and The Teachable Agents Group at Vanderbilt. 2005. Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence* 19, 3-4 (2005), 363–392.
- [9] YANG Bo. 2020. Online Learning Interaction in Danmaku Videos: A Pilot Study. In *EdMedia+ Innovate Learning*. Association for the Advancement of Computing in Education (AACE), 1262–1270.
- [10] Aaron W Calhoun, May CM Pian-Smith, Robert D Truog, David M Gaba, and Elaine C Meyer. 2015. Deception and simulation education: issues, concepts, and commentary. *Simulation in Healthcare* 10, 3 (2015), 163–169.
- [11] Ang Chen, Paul W Darst, and Robert P Pangrazi. 1999. What constitutes situational interest? Validating a construct in physical education. *Measurement in physical education and exercise science* 3, 3 (1999), 157–XXX.
- [12] Chih-Ming Chen and Chung-Hsin Wu. 2015. Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education* 80 (2015), 108–121.
- [13] Gwo-Dong Chen, Jih-Hsien Lee, Chin-Yeh Wang, Po-Yao Chao, Liang-Yi Li, and Tzung-Yi Lee. 2012. An empathic avatar in a computer-aided learning program to encourage and persuade learners. *Journal of Educational Technology & Society* 15, 2 (2012), 62–72.
- [14] Yue Chen, Qin Gao, and Pei-Luen Patrick Rau. 2017. Watching a movie alone yet together: understanding reasons for watching Danmaku videos. *International Journal of Human-Computer Interaction* 33, 9 (2017), 731–743.
- [15] Michelene TH Chi and Ruth Wylie. 2014. The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist* 49, 4 (2014), 219–243.

INPUT Q&A

Q: Why do you think these empires needed to fight for more lands even though they already had so much?

A: There are many reasons why these major powers of Europe accumulated colonies and went to war with each other. Some of these empire building was about ego and about spreading one's influence and power. A lot of it was based on ethnic beliefs about civilization. These are rationalizations to take control of other people's resources. And a lot of it was that we were in a world where access to resources - in particular, access to raw materials, and especially oil - could, to some degree, define whether a power was a power at all.

OUTPUT DIALOGUE

AH: Let's take a moment to think about this question. Why do you think these empires needed to fight for more lands even though they already had so much?

JW: I think they wanted to show their power to other countries.

AH: That's right, JW. A lot of it was about ego and about spreading one's influence and power. But there are other reasons as well. Can you think of any, KM?

KM: I guess they also wanted to get more resources?

AH: That's correct. A lot of it was that we were in a world where access to resources - in particular, access to raw materials, and especially oil - could, to some degree, define whether a power was a power at all.

DH: I see. So, it was about power and resources.

AH: Yes, power and resources. But there's one more reason that's a bit more complicated. It has to do with beliefs about civilization.

KM: You mean, like, the white man's burden?

AH: Exactly. These are rationalizations to take control of other people's resources. So, to sum up, there are three main reasons: ego, power, and resources.

Figure 5: An example of virtual characters' dialogues generated by a GPT-3 model after training with 5 examples. The input is a standard Q&A. The output is a dialogue between one TA (AH in this script) and 3 students (JW, KM, and DH) that covers the same content.

- [16] Konstantinos Chorianopoulos. 2018. A taxonomy of asynchronous instructional video styles. *International Review of Research in Open and Distributed Learning* 19, 1 (2018).
- [17] M Dewan, Mahbub Murshed, and Fuhua Lin. 2019. Engagement detection in online learning: a review. *Smart Learning Environments* 6, 1 (2019), 1–20.
- [18] Brian Dorn, Larissa B Schroeder, and Adam Stankiewicz. 2015. Piloting TrACE: Exploring spatiotemporal anchored collaboration in asynchronous learning. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 393–403.
- [19] Patrick E Downes, Eean R Crawford, Scott E Seibert, Adam C Stoverink, and Elizabeth M Campbell. 2021. Referents or role models? The self-efficacy and job performance effects of perceiving higher performing peers. *Journal of Applied Psychology* 106, 3 (2021), 422.
- [20] Ilana Dubovi and Iris Tabak. 2020. An empirical analysis of knowledge co-construction in YouTube comments. *Computers & Education* 156 (2020), 103939.
- [21] Evren Eryilmaz, Brian Thoms, Justin Mary, Rosemary Kim, and Jakko van der Pol. 2015. Instructor versus peer attention guidance in online learning conversations. *AIS Transactions on Human-Computer Interaction* 7, 4 (2015), 234–268.
- [22] Jennifer A Fredricks, Phyllis C Blumenfeld, and Alison H Paris. 2004. School engagement: Potential of the concept, state of the evidence. *Review of educational research* 74, 1 (2004), 59–109.
- [23] Anne C Frenzel, Thomas Goetz, Oliver Lüdtke, Reinhard Pekrun, and Rosemary E Sutton. 2009. Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of educational psychology* 101, 3 (2009), 705.
- [24] Meredith D Gall. 1970. The use of questions in teaching. *Review of educational research* 40, 5 (1970), 707–721.
- [25] Maia Garau, Mel Slater, Vinoba Vinayagamoorthy, Andrea Brogni, Anthony Steed, and M Angela Sasse. 2003. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 529–536.
- [26] Barbara A Greene. 2015. Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. *Educational Psychologist* 50, 1 (2015), 14–30.
- [27] Philip J Guo, Juho Kim, and Rob Rubin. 2014. How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the first ACM conference on Learning@ scale conference*. 41–50.
- [28] Yi-Chen Huang, Chih-Ming Chen, and Jian-Cheng Guo. 2016. Effects of attention monitoring and alarm mechanism on promoting e-Learning performance. In *ICEL: Proceedings of International Conference on e-Learning*. 229–237.
- [29] Eugene Hwang and Jeongmi Lee. 2021. Attention Guidance Technique Using Visual Subliminal Cues And Its Application On Videos. In *ACM International Conference on Interactive Media Experiences*. 167–177.
- [30] Sanna Järvelä. 1998. Socioemotional aspects of students' learning in a cognitive-apprenticeship environment. *Instructional science* 26, 6 (1998), 439–472.
- [31] Shane R Jimerson, Emily Campos, and Jennifer L Greif. 2003. Toward an understanding of definitions and measures of school engagement and related terms. *The California School Psychologist* 8, 1 (2003), 7–27.
- [32] Adam N Joinson. 2007. Disinhibition and the Internet. In *Psychology and the Internet*. Elsevier, 75–92.
- [33] Ji-Youn Jung, Sihang Qiu, Alessandro Bozzon, and Ujwal Gadiraju. 2022. Great Chain of Agents: The Role of Metaphorical Representation of Agents in Conversational Crowdsourcing. In *CHI Conference on Human Factors in Computing Systems*. 1–22.
- [34] Judita Kasperuniene, Meet Jariwala, Egidijus Vaskevicius, and Saulius Satkauskas. 2016. Affective engagement to virtual and live lectures. In *International Conference on Information and Software Technologies*. Springer, 499–508.
- [35] Kangsoo Kim, Luke Boelling, Steffen Haesler, Jeremy Bailenson, Gerd Bruder, and Greg F Welch. 2018. Does a digital assistant need a body? The influence of visual embodiment and social behavior on the perception of intelligent virtual agents in AR. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 105–114.
- [36] Yanghee Kim and Amy L Baylor. 2006. A social-cognitive framework for pedagogical agents as learning companions. *Educational technology research and development* 54, 6 (2006), 569–596.
- [37] Yelim Kim, Mohi Reza, Joanna McGrenere, and Dongwook Yoon. 2021. Designers characterize naturalness in voice user interfaces: their goals, practices, and challenges. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

- [38] Xi Lin, Mingyu Huang, and Leslie Cordie. 2018. An exploratory study: using Danmaku in online video-based lectures. *Educational Media International* 55, 3 (2018), 273–286.
- [39] Ati Suci Dian Martha and Harry B Santoso. 2019. The design and impact of the pedagogical agent: A systematic literature review. *Journal of Educators Online* 16, 1 (2019), n1.
- [40] Jeffrey D Marx and Karen Cummings. 2007. Normalized change. *American Journal of Physics* 75, 1 (2007), 87–91.
- [41] J Terry Mayes. 2015. Still to learn from vicarious learning. *E-Learning and Digital Media* 12, 3-4 (2015), 361–371.
- [42] Caitlin Mills, Nigel Bosch, Kristina Krasich, and Sidney K D’Mello. 2019. Reducing mind-wandering during vicarious learning from an intelligent tutoring system. In *International Conference on Artificial Intelligence in Education*. Springer, 296–307.
- [43] Ari Nugraha, Izhar Almizan Wahono, Jianpeng Zhanghe, Tomoyuki Harada, and Tomoo Inoue. 2020. Creating Dialogue Between a Tutee Agent and a Tutor in a Lecture Video Improves Students’ Attention. In *International Conference on Collaboration Technologies and Social Computing*. Springer, 96–111.
- [44] John O’Donoghue, Gurmak Singh, and Charmaine Green. 2004. A comparison of the advantages and disadvantages of IT based education and the implication upon students. *Digital Education Review* 9 (2004), 63–76.
- [45] George Palaigeorgiou and Anthea Papadopoulou. 2019. Promoting self-paced learning in the elementary classroom with interactive video, an online course platform and tablets. *Education and Information Technologies* 24, 1 (2019), 805–823.
- [46] Cecilia Panigutti, Andrea Beretta, Fosca Giannotti, and Dino Pedreschi. 2022. Understanding the impact of explanations on advice-taking: a user study for AI-based clinical Decision Support Systems. In *CHI Conference on Human Factors in Computing Systems*. 1–9.
- [47] Christine Pleines. 2020. Understanding vicarious participation in online language learning. *Distance Education* 41, 4 (2020), 453–471.
- [48] Louise Pouliot and Paul S Cowen. 2007. Does perceived realism really matter in media effects? *Media Psychology* 9, 2 (2007), 241–259.
- [49] Susen Rabold, Stuart Anderson, John Lee, and Neil Mayo. 2008. YouTute: Online social networking for vicarious learning. In *Proceedings of the ICL2008*. Kassel University Press, 1–14.
- [50] Angélique Remacle, Stéphane Bouchard, Anne-Marie Etienne, Marie-Christine Rivard, and Dominique Morsomme. 2021. A virtual classroom can elicit teachers’ speech characteristics: evidence from acoustic measurements during in vivo and in virtuo lessons, compared to a free speech control situation. *Virtual Reality* 25, 4 (2021), 935–944.
- [51] Deborah Richards and Virginia Dignum. 2019. Supporting and challenging learners through pedagogical agents: Addressing ethical issues through designing for values. *British Journal of Educational Technology* 50, 6 (2019), 2885–2901.
- [52] Joni Salminen, Soon-gyo Jung, and Bernard J Jansen. 2021. Suggestions for Online User Studies. In *International Conference on Human-Computer Interaction*. Springer, 127–146.
- [53] Kerstin Schankweiler. 2020. Reaction Images and Metawitnessing. *Parallax* 26, 3 (2020), 254–270.
- [54] Laura Scheepmaker, Kay Kender, Christopher Frauenberger, and Geraldine Fitzpatrick. 2021. Leaving the field: Designing a socio-material toolkit for teachers to continue to design technology with children. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [55] Umair Uddin Shaikh, Shakir Karim, and Zaheeruddin Asif. 2017. Re-Thinking Vygotsky: Applying social constructivism to asynchronous online courses utilizing the power of crowdsourcing. (2017).
- [56] Ellen Skinner, Carrie Furrer, Gwen Marchand, and Thomas Kindermann. 2008. Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of educational psychology* 100, 4 (2008), 765.
- [57] Matthijs Smakman, Paul Vogt, and Elly A Konijn. 2021. Moral considerations on social robots in education: A multi-stakeholder perspective. *Computers & Education* 174 (2021), 104317.
- [58] Gail M Sullivan and Anthony R Artino Jr. 2013. Analyzing and interpreting data from Likert-type scales. *Journal of graduate medical education* 5, 4 (2013), 541–542.
- [59] Leah A Sutton. 2001. The principle of vicarious interaction in computer-mediated communications. *International Journal of Educational Telecommunications* 7, 3 (2001), 223–242.
- [60] Lauren Szcurek, Benoît Monin, and James J Gross. 2012. The stranger effect: The rejection of affective deviants. *Psychological science* 23, 10 (2012), 1105–1111.
- [61] Keith S Taber. 2015. Meeting educational objectives in the affective and cognitive domains: Personal and social constructivist perspectives on enjoyment, motivation and learning chemistry. In *Affective dimensions in chemistry education*. Springer, 3–27.
- [62] Mohammad Tahaei and Kami Vaniea. 2022. Recruiting Participants With Programming Skills: A Comparison of Four Crowdsourcing Platforms and a CS Student Mailing List. In *CHI Conference on Human Factors in Computing Systems*. 1–15.
- [63] Yassin Taskin, Tobias Hecking, H Ulrich Hoppe, Vania Dimitrova, and Antonija Mitrovic. 2019. Characterizing comment types and levels of engagement in video-based learning as a basis for adaptive nudging. In *European Conference on Technology Enhanced Learning*. Springer, 362–376.
- [64] Yi-Miau Tsai, Mareike Kunter, Oliver Lüdtke, Ulrich Trautwein, and Richard M Ryan. 2008. What makes lessons interesting? The role of situational and individual factors in three school subjects. *Journal of Educational Psychology* 100, 2 (2008), 460.
- [65] Cristian Vaccari and Andrew Chadwick. 2020. Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media+ Society* 6, 1 (2020), 2056305120903408.
- [66] Kati Vasalampi, Joonas Muotka, Sanni Pöysä, Marja-Kristiina Lerkkanen, Anna-Maija Poikkeus, and Jari-Erik Nurmi. 2016. Assessment of students’ situation-specific classroom engagement by an InSitu Instrument. *Learning and Individual Differences* 52 (2016), 46–52.
- [67] Lev Semenovich Vygotsky and Michael Cole. 1978. *Mind in society: Development of higher psychological processes*. Harvard university press.
- [68] Xu Wang, Miaomiao Wen, and Carolyn P Rosé. 2016. Towards triggering higher-order thinking behaviors in MOOCs. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. 398–407.
- [69] Justin D Weisz, Sara Kiesler, Hui Zhang, Yuqing Ren, Robert E Kraut, and Joseph A Konstan. 2007. Watching together: integrating text chat with video. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 877–886.
- [70] James V Wertsch. 1993. *Voices of the mind: Sociocultural approach to mediated action*. Harvard University Press.
- [71] Qunfang Wu, Yisi Sang, and Yun Huang. 2019. Danmaku: a new paradigm of social interaction via online videos. *ACM Transactions on Social Computing* 2, 2 (2019), 1–24.
- [72] Yabo Zhang, Aiping Qian, Zhongling Pi, and Jiumin Yang. 2019. Danmaku related to video content facilitates learning. *Journal of Educational Technology Systems* 47, 3 (2019), 359–372.

A PRE- AND POST-TESTS

There are ten questions per video. Questions 1-5 are memorization questions (answers can be found directly in the video or the augmented content) and questions 6-10 are application questions (students have to make inference based on what they learn from the video.) Note that the pre- and post-tests for each video is the same. The correct answers are highlighted in green. For questions that ask for rationale, we give full points if the student gives a response that does not contradict with the content in the video.

A.1 World War I video

- (1) What are the differences between the borders of European countries now and in 1914? (Select all that apply)
 - (a) Austria and Hungary used to be one empire.
 - (b) Germany used to be smaller than it is now.
 - (c) Russia used to be bigger than it is now.
 - (d) Serbia used to be a part of the Ottoman Empire.
 - (e) Bosnia and Herzegovina used to be a part of Italy.
- (2) Which countries (or parts of countries) were once a part of the Ottoman Empire? (Select all that apply)
 - (a) Israel
 - (b) Russia
 - (c) Turkey
 - (d) Saudi Arabia
 - (e) India
 - (f) Greece
- (3) What were the major powers in Europe doing in 1914 that contributed to the tension between them? (Select all that apply)
 - (a) Acquiring more natural resources
 - (b) Connecting people of the same religious belief
 - (c) Assassinating each other's military leaders
 - (d) Spreading their ethnic beliefs to their colonies
 - (e) Racing each other in arms accumulation
- (4) Which country exercised control over South Africa at the beginning of WW1?
 - (a) Germany
 - (b) France
 - (c) Russia
 - (d) England
 - (e) Italy
- (5) What was one similarity between Bosnia and Serbia in 1914?
 - (a) They were both colonies of France.
 - (b) They spoke similar dialects/languages
 - (c) They both left the Ottoman Empire.
 - (d) They colonized countries in the northern part of Africa.
 - (e) They were a part of the same political alliance.
- (6) During the time leading up to WW1, Russia's economy was becoming more and more dependent on the Ottoman Empire. Why do you think this is the case? (Hint: think about the

geographical location of both empires.)



- (a) Because Russia needed to import many goods from the Ottoman Empire.
 - (b) Because Russia needed to pay extra tax to trade in the Ottoman Empire.
 - (c) Because Russia had to export goods through ports of the Ottoman Empire.
 - (d) Because Russia lost their trade rights with England.
- (7) How do you think the empires benefitted from their colonies DURING the war? (Select all that apply)
 - (a) Raw materials to make weapons
 - (b) Routes to transport supplies
 - (c) Extra manpower
 - (d) Extra space for building weapon factories
 - (8) In 1908, the pan-Slavic ("pro-Slavic") movement in Russia mobilized to protest against Austria-Hungary. What did Austria-Hungary do that provoked this action?
 - (a) Invaded Serbia
 - (b) Attempted an assassination on Slavic leaders
 - (c) Deported Slavic people from Austria-Hungary
 - (d) Annexed Bosnia and Herzegovina

- (9) Prior to WW1, Germany initiated an alliance called the League of Three Emperors to avoid war on two fronts (east and west borders). This alliance controlled the Eastern Europe. Which of the two following empires did Germany allied with? Explain the rationale of your answer in 1-2 sentences. (A correct answer without rationale will not be marked.)



(a) The French Empire

(b) **The Austro-Hungarian Empire**

Reason: (Example answer) Germany bordered three empires, the French Empire, the Austro-Hungarian Empire, and the Russian Empire. But France was in the west, so it wouldn't contribute to the Eastern European power.

- (10) Based on the previous two questions, which of the following two empires do you think the British Empire allied with at the beginning of World War 1? Explain the rationale of your answer in 1-2 sentences. (A correct answer without rationale will not be marked.)

(a) **Russia**

(b) The Austro-Hungarian Empire

Reason: (Example answer) Russia is opposite to Austro-Hungarian (because of pan-slavism). Austro-Hungary allied with Germany. Germany and Britain were on opposite sides.

A.2 Napoleonic War video

- (1) What are possible causes of Napoleon's death? (Select all that apply)

(a) Lung cancer

(b) **Stomach cancer**

(c) Mercury poisoning

(d) **Arsenic poisoning**

(e) Suicide

- (2) What are the nations that sent armies to join the Great Britain army under the Duke of Wellington? (Select all that apply)

(a) **The Netherlands**

(b) Prussia

(c) **Hanover**

(d) **Brunswick**

(e) **Nassau**

(f) Spain

- (3) What happened in France after Napoleon had been exiled to Elba that made French people unhappy? (Select all that apply)

(a) Taxes were raised to help the country recover from war loss

(b) **King Louis XVIII returned with the nobility.**

(c) **Army veterans were treated badly.**

(d) France was under a treaty that put them at a disadvantage in trades.

(e) **The French Empire shrank back down to its original boundary.**

- (4) What event in 1812 decimated the French Grand Armee?

(a) The invasion of Egypt

(b) The naval battle in Spain

(c) **The invasion of Russia**

(d) The invasion of Prussia

(e) The withdrawal of Austria from the Coalition

- (5) What was Napoleon's strategy for the Battle of Waterloo that didn't work out due to miscalculation of timing?

(a) He planned to wait out and take measure of the nations that joined the new Coalition Army.

(b) He planned to get backup from Russia before meeting the English army.

(c) He planned to launch an attack on Poland and Germany for extra manpower.

(d) He planned to cross the English channel and attack England first.

(e) **He planned to launch an attack on each Coalition Army before they could come together.**

- (6) What do you think is Prussia's administrative system DURING the Napoleonic War?

(a) **Absolute Monarchy (the monarchs have all the powers)**

(b) Democratic Monarchy (the monarchs are under the constitution/parliament)

(c) Military Dictatorship (a dictatorship primarily enforced by the military.)

(d) Democratic Republic (elected individuals represent the citizen body)

- (7) What do you think is France's administrative system AFTER the Napoleonic War?

(a) Absolute Monarchy (the monarchs have all the powers)

(b) **Democratic Monarchy (the monarchs are under the constitution/parliament)**

(c) Military Dictatorship (a dictatorship primarily enforced by the military.)

(d) Democratic Republic (elected individuals represent the citizen body)

- (8) During Napoleon's exile in Elba, the French war prisoners were returned to France from Russia, Germany, Britain, and Spain. How do you think this news affected his decision to return to France? Explain the rationale behind your answer in 1-2 sentences. (A correct answer without supporting rationale will not be marked.)

(a) **It further encouraged him to return to France.**

(b) It discouraged him from returning to France.

Reason: (Example answer) these prisoners were Napoleon's soldiers who got captured in battles in the aforementioned countries, i.e., they were the people who would rally behind him.

(9) What do you think was Napoleon's choice of weapons that were not suitable for muddy terrain? (Select all that apply)

(a) Handguns/pistols



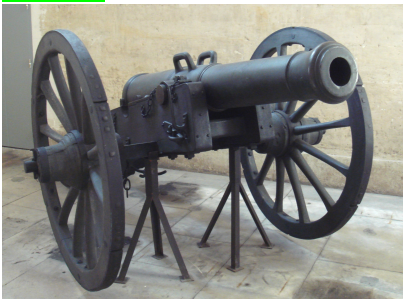
(b) Heavy cavalry



(c) Light cavalry



(d) Cannons



(e) Trebuchets



(10) 2 days before the decisive battle at Waterloo, Napoleon's army engaged with the Prussian army, one of the Coalition army, in the Battle of Ligny. Who do you think won? Why?

(a) France won.

(b) Prussia won.

Reason: (Example answer) the victory made the French complacent. They thought their plan worked, not knowing that the surviving Prussian force managed to retreat and join up with the rest of the Coalition force anyway.

B ENGAGEMENT QUESTIONNAIRE

The following statements are for self-evaluation in SVD condition terms. In places where the questionnaire differs for EdPuzzle condition, the different term is in parentheses. After each item, we include the reference to the original questionnaires (Situational Interest scale [11], Original Cognitive Engagement scale [26], Engagement vs. Disaffection (EvsD) scale [56], and Situational engagement scale [66]) that the item is taken from as well as the number of the item in the original questionnaire.

INSTRUCTION: Please evaluate your experience while watching the video by rating the following statement on a scale of "Completely disagree" to "Completely agree". (7-point Likert scale)

Behavioral engagement:

- I concentrated during the lesson. (Situational engagement scale, Item Beh5)
- I watched the video very carefully during the lesson. (EvsD scale, Behavioral Engagement section, Item 5)
- I was persistent during the lesson. (Situational engagement scale, Item Beh6)
- I want to find out more about the subject matter. (Situational Interest scale, Exploration Intention section, Item 2)

Emotional engagement and disengagement:

- I liked the lesson. (Situational engagement scale, Item Emo1)
- The lesson was exciting. (Situational Interest scale, Instant Enjoyment scale, Item 1)
- The lesson was enjoyable. (Situational Interest scale, Instant Enjoyment scale, Item 2)
- I was tired during the lesson. (Situation engagement scale, Item Daff 2)
- When the virtual characters interacted with each other (When the pop-up quizzes appeared), I felt bored. (EvsD scale, Emotional Disaffection, Item 1a)
- When the instructor in the video first explained the new material, I felt bored. (EvsD scale, Emotional Disaffection, Item 1c)
- When the virtual characters interacted with each other (When the pop-up quizzes appeared), I felt interested. (EvsD scale, Emotional Engagement, Item 2)

Cognitive engagement:

- I mentally took notes of details as I watched the video. (Original Cognitive Engagement, Item 30)
- I mentally took notes of main ideas as I watched the video. (Original Cognitive Engagement, Item 31)
- I tried to memorize the content of the discussion between the virtual characters (the content of the pop-up quizzes). (Original Cognitive Engagement, Item 32)
- I tried to memorize exactly what the instructor said during the lecture. (Item 36)
- I tried to memorize lists of new terms and definitions. (Original Cognitive Engagement, Item 39)
- I put together ideas or concepts and drew conclusions that were not directly stated in the video. (Original Cognitive Engagement, Item 19)

- While learning new concepts, I tried to think of implications and practical applications. (Original Cognitive Engagement, Item 21)
- I tried to learn new material by mentally associating new ideas with similar ideas that I already knew. (Original Cognitive Engagement, Item 23)
- I evaluated the usefulness of the ideas presented in the video. (Original Cognitive Engagement, Item 24)
- Before starting the POST-TEST, I made sure I understood the lesson content. (Original Cognitive Engagement, Item 26)

C POST-TASK SURVEY

- (1) When you watched the two videos in this study, what were your **LEARNING GOALS**? (i.e. what would you like to be able to do with the knowledge you learned?) (Select all that apply)
- Remember (recall facts and basic concepts)
 - Understand (explain ideas or concepts)
 - Apply (use information in new situations)
 - Analyze (draw connections among ideas)
 - Evaluate (justify a stand or decision)
 - Create (produce new or original work)
- (2) Between EdPuzzle and virtual characters' interactions, which one do you prefer?
- Prefer EdPuzzle a lot more
 - Prefer EdPuzzle slightly more
 - No preference
 - Prefer virtual characters' interactions (SVD) slightly more
 - Prefer virtual characters' interactions (SVD) a lot more
- (3) For each statement below, pick the design you prefer and explain your reason in the corresponding textbox. Comparing the virtual characters' interactions to EdPuzzle's quizzes, which one ...

	SVD	EdPuzzle	No difference
Helps you think more actively?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Helps you remember the ADDITIONAL content (e.g., the chat or the quiz feedback) better?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Creates more feeling of social connectedness?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Helps you to focus on the lesson content more?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- (4) In what ways does the **classroom scenario** of the scripted interaction affect your learning experience? (E.g., how does it affect your attention? Your boredom? Your inclination to think along?)
Answer: _____

- (5) How does **being an observer** of the virtual TA's interactions with virtual students, rather than an active participant, affect your learning?
Answer: _____

- (6) How much do you agree with the following statements about the **VIRTUAL CHARACTERS**? (on a 5-point Likert scale from "Completely Disagree" to "Completely Agree")
- The characters' identities are diverse and inclusive.
 - The characters' dialogue sound natural and/or realistic.
 - I feel comfortable watching the interactions between the characters.

- (7) Were there any moments when you compared yourself or related to the characters?
 Yes. Please describe the moment(s): _____
 No.
- (8) Were there any moments when the characters' dialogue... (Select all that apply)?
 Felt particularly supportive or offensive? Please describe the moment(s): _____
 Felt particularly unnatural or natural? Please describe the moment(s): _____
 Made your learning experience particularly tedious or enjoyable? Please describe the moment(s): _____
- (9) For each of the components of virtual characters' interactions below, evaluate how it affected your learning experience and explain your reason in the corresponding textbox. (On a 5-point Likert scale from "It has a very negative effect" to "It has a very positive effect")
- The chat messages
 - The emoji reactions
 - The conversation between the characters before the video starts
 - The questions asked by virtual students or prompted by the virtual TA
 - The correct answers from the virtual students
 - The wrong answers from the virtual students
- (10) In the current interface, you can click "Click to continue" to proceed through the virtual characters' interaction. How much do you agree with the following statements regarding the navigation method? Please explain your reason in the corresponding textbox. (On a 5-point Likert scale from "Completely Disagree" to "Completely Agree")
- Clicking for messages makes the learning experience more engaging.
 - Clicking for messages helps me self-pace my learning.
 - Clicking for messages is convenient.
- (11) When did you click the button "Click to continue"? (For example, did you pause to think over some dialogues before you continued? Or did you immediately proceed to see the answers? Did you open and read one message at a time or open many messages and then read the whole thread?)
Answer: _____

- (12) Imagine that each message shows up automatically after a few seconds. What do you think would be the pros and cons of that approach compared to clicking through the text messages at your own pace?
Answer: _____

- (13) Do you have any other feedback or opinions about virtual characters' interaction that you would like to share?
Answer: _____

D DIALOGUE GENERATION EXAMPLES

D.1 Training examples

Below are 6 pairs of corresponding Q&A (from EdPuzzle augmentation) and virtual characters' scripts that we manually wrote. (Note that AH is the TA in this script; DH, KM, and JW are students.)

Example 1:

[Question:] Can you spot the differences between this map from 1914 and what Europe looks like nowadays (the previous map that was just shown)? Hint: think about what countries existed and where the borders were!

[Answer:] There are three main differences between Europe in 1914 and nowadays: (1) Germany was bigger than it is now; (2) Russia was also bigger than it is now; (3) Austria and Hungary are one country, and it's called the Austro-Hungarian Empire.

[Dialogue:]

AH: Before the professor tells us the answer, let's take a moment, try and see if you can spot the differences between this map from 1914 and what Europe looks like nowadays. Specifically, think about what countries existed and where the borders were.

DH: Germany seems pretty big here.

KM: Austria and Hungary are one country?

AH: Yes, and yes! We can see two big countries in the middle that look different from the modern map, the German Empire and Austria Hungary, also known as the Austro-Hungarian Empire. Is there any other country that seemed bigger back in the day?

JW: Russia?

AH: That's right, Russia.

Example 2:

[Question:] As the video mentioned, most countries in the Middle East were in the Ottoman Empire in 1914. Which countries do you think are there today? And which countries do you think were in the Ottoman Empire at its peak (though no longer a part of it in 1914)?

[Answer:] The Ottoman Empire consisted of modern day Turkey and much of the modern Middle East, especially the Arab world - Syria, Iraq, Lebanon, much of where modern day Israel is, some of Saudi Arabia. (One common mistake that my students often make in the exam is thinking that India is a part of the Ottoman Empire. India is further to east of Persia and has no connection whatsoever with the Ottoman Empire.) What you see on this map was actually the dying state of the Ottoman Empire. At its peak, it controlled much of the Muslim world. In addition to the bright red area on the map, it also had North Africa, a little bit of the west side of Persia, and even a little bit of the Balkan/Southeast Europe, including Greece. However, this was hundreds and hundreds of years before this map.

[Dialogue:]

AH: Let me pause here for a bit. As they mentioned,

most countries in the Middle East were in the Ottoman Empire. Do you know which countries are there today?

KM: Iraq?

AH: Right, KM. Iraq is one. What else?

JW: Saudi Arabia?

AH: Yes, although not all of Saudi Arabia, but a part of it was in the Ottoman Empire. What other countries do you think there are, DH?

DH: India?

AH: India is actually a bit further down and to the right, so it was never part of the Ottoman Empire. Some other countries that were in the bright red region on the map are Syria, Lebanon, and most of Israel.

DH: I see. But I heard from somewhere that the Ottoman Empire was bigger than this...?

AH: DH said in the chat that he heard that the Ottoman Empire was supposedly bigger than this, and that's not wrong! What you see here is the Ottoman Empire in its dying state. At the peak of its power, long before this, it actually controlled most of the Muslim world. There is North Africa - Egypt right there. Also a little bit to the right, going into Persia. A little bit to the top, going into Southeast Europe. Greece, in particular. This is a map from 1914, so the Empire already shrunk down to only a few countries.

KM: Did it just start declining around the time WW1 started?

AH: KM asked whether the Ottoman Empire started declining around 1914. The answer is no. Actually, its peak was hundreds and hundreds of years ago, and it had been declining for some time already.

KM: Got it. Thank you!

Example 3:

[Question:] Why do you think these empires needed to fight for more lands even though they already had so much?

[Answer:] There are many reasons why these major powers of Europe accumulated colonies and went to war with each other. Some of these empire building was about ego and about spreading one's influence and power. A lot of it was based on ethnic beliefs about civilization. These are rationalizations to take control of other people's resources. And a lot of it was that we were in a world where access to resources - in particular, access to raw materials, and especially oil - could, to some degree, define whether a power was a power at all.

[Dialogue:]

JW: Question! Why do these empires need to fight for more lands if they already have so much?

AH: JW just asked a great question here in the chat. What's the point of all of these expansions? Well, one could say it's just ambition and ego. But there

were also reasons like, we want to civilize these other lands, to spread our ethnic beliefs. It sounds absurd now. But these are considered a rational justification for colonizations at that time.

KM: Fight for natural resources?

AH: KM suggested in the chat that the fight for natural resources could be a factor. Yes, it is definitely one of the reasons for war. The race to possess natural resources, raw materials, and oil, is one of the ways for these empires to gain power. Does this make sense, JW?

JW: Yes. Thanks!

Example 4:

[Question:] 10 months after his relatively comfortable exile in Elba, Napoleon decided to return to France and regain his title of an emperor. The video just told us a few reasons for this, but there was another reason: the people of France were unhappy about something, and so Napoleon thought he might be able to rally them and retake the force. What do you think the French people were unhappy about?

[Answer:] Firstly, the monarchy returned with King Louis XVIII, and the nobility in power did not treat the army veterans very well, so there was a good chance that Napoleon could retake control of the force. Also, the French Empire got shrunken back to its original boundary after Napoleon was gone, and French people didn't feel proud about that. So, Napoleon thought that maybe he could do something about it.

[Dialogue:]

AH: Let's pause here. So, 10 months after his relatively comfortable exile in Elba, Napoleon decided to return to France and regain his title of an emperor. The professor just told us a few reasons for this, but there was another reason: the people of France were unhappy about something, and so Napoleon thought he might be able to rally them and retake the force. What do you think the French people were unhappy about?

DH: Were they unhappy because they became weaker?

AH: That's right, but how were they weaker?

JW: Maybe they had to give back the lands they had conquered.

AH: That's right, JW. The French Empire was reduced down to its pre-war borders, and the French people didn't feel proud about this. This is one reason. But there's another reason that has to do with the rulers of the Empire. Can you think of what that might be?

KM: The return of monarchy?

AH: Exactly! With Napoleon gone, King Louis XVIII returned to rule France with the aristocracy. And they didn't treat the army veterans from Napoleon's force very well. So, Napoleon saw an opportunity to retake military control and decided to return to France.

Example 5:

[Question:] Although the video said that France met the Great Britain force, that force was actually a combination of small armies from other nations, led by a British general. Who do you think threw in their forces with Britain?

[Answer:] The Great Britain army is actually a coalition of units from multiple nations: the United Kingdom, the Netherlands, Hanover, Brunswick, and Nassau.

[Dialogue:]

AH: Let's pause here for a moment to talk about the enemies of the French Empire. Although the professor said that France met the Great Britain force, that force was actually a combination of small armies from other nations, led by a British general. Who do you think threw in their forces with Britain?

KM: The Netherlands?

AH: Yes, the Netherlands is one. What else? There are three other nations that no longer exist today.

JW: Brunswick?

AH: That's correct. Brunswick is one. What other countries do you think were there, DH?

DH: Prussia?

AH: Prussia was actually not in the Great Britain army but sent their own separate army to fight in the Battle of Waterloo. The other two countries were actually Hanover and Nassau.

Example 6:

[Question:] What do you think are the main reasons that Napoleon lost the Battle of Waterloo?

[Answer:] There are two main reasons. One is that it rained the night before the attack, so the ground was muddy, and Napoleon's weapons weren't suitable for such terrain. The other reason is that Napoleon's strategy was to launch a surprise attack on the coalition armies before they combined. However, there was a miscalculation in timing and the 7th Coalition managed to assemble anyway.

[Dialogue:]

AH: Let's pause here a bit to discuss the battle. Would anyone like to guess the reasons for Napoleon's loss?

DH: Maybe he thought that his enemies had not caught the news of his return and planned attacks yet, but they did, so he failed.

AH: You're partially right. Napoleon planned to attack the separate armies before they could assemble and form another Coalition. But the reason that he failed was not because they heard the news, but because he made a miscalculation in the timing of his attack, so the other side managed to assemble their force anyway.

JW: Is this assembled force the 7th Coalition?

AH: Yes. It's the 7th Coalition. Its formation was very disadvantageous for the French army, but there was one other factor that played a role in Napoleon's loss and it has to do with weather. Can anyone guess what happened?

KM: It rained?

AH: That's right. It rained the night before the attack. Unfortunately, Napoleon didn't prepare for this, and the weapons in his army weren't suitable for the muddy terrain. So, that's our two reasons: the muddy ground and the miscalculation of attack timing. Let's return to the video and see what happened after the war was lost.

D.2 Generated dialogues

We tried generating each of the 6 dialogues above by training the GPT-3 model (engine: davinci; temperature: 0.55) with the other 5 dialogues. For example, we used Example 2-6 as training examples then input the Question and Answer of Example 1 for the model to generate the corresponding dialogue. Note that these are the raw outputs from GPT-3. In order to use them as scripts, we need to embellish the dialogues via simple rule-based algorithms, e.g., add student's name at the end of a response or replace some texts with emojis.

Generated Dialogue 1:

AH: Let me pause here for a bit. Can you spot the differences between this map from 1914 and what Europe looks like nowadays?
 JW: Germany is smaller now.
 AH: That's right, JW. Germany is smaller now. What else?
 KM: Russia is also smaller.
 AH: Yes, Russia is also smaller. They lost a lot of land after World War I and World War II. What else?
 DH: Austria and Hungary were one country.
 AH: Yes, that's correct. They were called the Austro-Hungarian Empire. Now, they're two separate countries.

Generated Dialogue 2:

AH: As the professor mentioned, most countries in the Middle East were in the Ottoman Empire in 1914. So, which countries do you think are there today? And which countries do you think were in the Ottoman Empire at its peak (though no longer a part of it in 1914)?
 KM: Turkey, Syria, Iraq, Lebanon, Israel, Saudi Arabia?
 AH: Yes, those are all countries that were once a part of the Ottoman Empire. What you see on this map was actually the dying state of the Ottoman Empire. At its peak, it controlled much of the Muslim world. In addition to the bright red area on the map, it also had North Africa, a little bit of the west side of Persia, and even a little bit of the Balkan/Southeast Europe, including Greece. However, this was hundreds and hundreds of years before this map.
 JW: What about India?
 AH: India is actually not a part of the Ottoman Empire. India is further to east of Persia and has no connection whatsoever with the Ottoman

Empire.

Generated Dialogue 3:

AH: Let's take a moment to think about this question. Why do you think these empires needed to fight for more lands even though they already had so much?
 JW: I think they wanted to show their power to other countries.
 AH: That's right, JW. A lot of it was about ego and about spreading one's influence and power. But there are other reasons as well. Can you think of any, KM?
 KM: I guess they also wanted to get more resources?
 AH: That's correct. A lot of it was that we were in a world where access to resources - in particular, access to raw materials, and especially oil - could, to some degree, define whether a power was a power at all.
 DH: I see. So, it was about power and resources.
 AH: Yes, power and resources. But there's one more reason that's a bit more complicated. It has to do with beliefs about civilization.
 KM: You mean, like, the white man's burden?
 AH: Exactly. These are rationalizations to take control of other people's resources. So, to sum up, there are three main reasons: ego, power, and resources.

Generated Dialogue 4:

AH: The video just told us a few reasons for this, but there was another reason: the people of France were unhappy about something, and so Napoleon thought he might be able to rally them and retake the force. What do you think the French people were unhappy about?
 JW: The monarchy returning?
 AH: That's one reason, yes. The monarchy returned with King Louis XVIII. But there's another reason that's connected to the army.
 DH: The nobility in power not treating the army veterans well?
 AH: That's right! The nobility in power did not treat the army veterans very well, so there was a good chance that Napoleon could retake control of the force. Also, the French Empire got shrunken back to its original boundary after Napoleon was gone, and French people didn't feel proud about that. So, Napoleon thought that maybe he could do something about it.
 KM: That makes sense.

Generated Dialogue 5:

AH: Let's pause here for a bit. The video said that France met the Great Britain force, but actually, that force was a combination of small armies from other nations, led by a British general. Who do you think threw in their forces with Britain?

JW: The Netherlands and Hanover?

AH: That's right! In addition to the United Kingdom, the Netherlands, Hanover, Brunswick, and Nassau also contributed their units to the Great Britain army.

KM: Why did they do that?

AH: KM asked in the chat why these countries decided to side with Great Britain. Do you have any guesses?

DH: For protection?

AH: Yes, that's one reason. But there's also another reason that has to do with Napoleon. Can you think of what that is?

JW: They didn't want Napoleon to rule them?

AH: That's right! They saw Napoleon as a threat to their own sovereignty.

Generated Dialogue 6:

AH: Let's pause here. What do you think are the main reasons that Napoleon lost the Battle of Waterloo?

JW: I think the weather played a role?

AH: That's right, JW. It rained the night before the attack, so the ground was muddy, and Napoleon's weapons weren't suitable for such terrain. What else?

DH: I think his troops were tired from previous battles?

AH: That could be a factor, but I don't think it's one of the main reasons. What else?

KM: I think he miscalculated the time?

AH: That's exactly right, KM. Napoleon's strategy was to launch a surprise attack on the coalition armies before they combined. However, there was a miscalculation in timing and the 7th Coalition managed to assemble anyway.