Does Dynamically Drawn Text Improve Learning? Investigating the Effect of Text Presentation Styles in Video Learning

Ashwin Ram
Shengdong Zhao
ashwinram@u.nus.edu
zhaosd@comp.nus.edu.sg
NUS-HCI Lab, Department of Computer Science
National University of Singapore
Singapore, Singapore

ABSTRACT

Dynamically drawn content (e.g., handwritten text) in learning videos is believed to improve users' engagement and learning over static powerpoint-based ones. However, evidence from existing literature is inconclusive. With the emergence of Optical Head-Mounted Displays (OHMDs), recent work has shown that video learning can be adapted for on-the-go scenarios. To better understand the role of dynamic drawing, we decoupled dynamically drawn text into two factors (font style and motion of appearance) and studied their impact on learning performance under two usage scenarios (while seated with desktop and walking with OHMD). We found that although letter-traced text was more engaging for some users, most preferred learning with typeface text that displayed the entire word at once and achieved better recall (46.7% higher), regardless of the usage scenarios. Insights learned from the studies can better inform designers on how to present text in videos for ubiquitous access.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in ubiquitous and mobile computing**.

KEYWORDS

Dynamic drawing, Text presentation styles, Video learning, Heads-up computing, Mobile HCI, Smart-glasses, OHMD

ACM Reference Format:
Ashwin Ram and Shengdong Zhao. 2022. Does Dynamically Drawn Text Improve Learning? Investigating the Effect of Text Presentation Styles in Video Learning. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3491102.3517499

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04…$15.00
https://doi.org/10.1145/3491102.3517499

1 INTRODUCTION

Characterized by dynamically drawn content on a digital blackboard, Khan Academy's online learning videos have been immensely popular, attracting millions of users every month [24, 26, 40]. Its drawing-based production style is considered a key element to its success; learners tend to be more engaged and receptive to Khan Academy's video content than they are to static PowerPoint-based videos [19].

Several studies have endeavored to investigate this dynamic drawing effect [6, 12–14, 19], albeit the results and insights have been largely inconclusive. For instance, Cross et al. found that some users preferred dynamically drawn text (handwriting) in Khan Academy videos because of its engaging properties whereas other users preferred standard typeface text for its clarity and professionalism [6]. Fiorella et al., on the other hand, showed that dynamically drawn video content (text and diagrams) improved users' learning but was no more engaging than static pre-drawn videos [13]. Perusing literature, it seems that the following questions remain to be answered: 1) Does dynamic drawing improve learning outcomes? 2) Is it more engaging?

Without clear conclusions, it is difficult to make informed design decisions on the presentation style when creating learning videos.

In an attempt to draw a clearer conclusion regarding how dynamic drawing affects learning we decided to conduct a systematic investigation of dynamic drawing as applied to textual content. Traditionally, video learning is mostly performed under stationary desktop settings. But, the ubiquitous access to online learning content on mobile devices has enabled learning on the go. While smartphones are currently the predominant platform for on-the-go learning, prior investigations have already shown that smartphones, due to the heads-down nature of interaction, can lead to decreased situational awareness and a decline in learning performance [5, 50].

In contrast, the emergent Optical Head-Mounted Displays (OHMDs) platform has recently been gaining traction as a suitable platform for on-the-go information acquisition [17, 39]. As a see-through wearable display, it enables digital content to be viewed alongside the environment, facilitating both task performance and path navigation, as has been demonstrated in the contexts of reading [39] and editing text on the go [17]. Furthermore, OHMDs naturally induce a desirable heads-up posture which can reduce the risk of posture related injuries such as turtleneck in the long run [20]. Recent work has also explored its potential for adapting video learning for on-the-go scenarios such as walking, noting that OHMDs can improve recall while enabling faster walking speed as compared to smartphones [38]. Thus, we focus our investigation on the desktop platform for stationary learning and OHMD platform for mobile learning scenarios.

For our analysis, we look into the two main aspects in which handwritten and typeface text differ: 1) font style: handwritten text appears more natural and free-style, whereas typeface text is visibly more rigid. 2) motion of the letter as it appears on-screen: handwritten letters are traced out gradually, while typeface text appears immediately either in a letter-by-letter manner or as an entire word. These two aspects seem to play different roles in the benefits of dynamic handwriting suggested in literature. To delineate the precise effects of each factor, we conducted two controlled studies.

In the first study, we compared font style (Handwritten, Typeface) and motion (Appear-letter, Appear-word, Trace-letter) by using controlled pseudoword-based videos, so as to eliminate experimentally confounding effects that may accompany the use of real educational videos. Against our expectations, we found that the majority of users preferred to have the entire word appear immediately, i.e., appear-word. Learning outcomes were also more effective with appear-word than trace-letter, regardless of the context of usage (desktop while sitting vs. OHMD while walking). Furthermore, this positive effect from appear-word was strengthened when combined with typeface fonts, resulting in a 53.1% improvement of average recall scores over handwriting.

Subsequently, we conducted a validation study with Khan Academy videos to understand how these findings transfer to realistic settings. Our results revealed that although a few users found handwriting to be more engaging and naturalistic, most users still preferred typeface fonts with appear-word motion, which improved recall scores by 46.7% on average over handwritten text, thereby reinforcing our results from the initial study.

Together, our studies indicate that dynamically drawn text does not improve users' learning outcomes in terms of recall. Instead, we recommend using typeface fonts that display the entire word immediately as a more general-purpose text presentation style for videos. However, handwriting can still be useful for improving video engagement when dealing with familiar content in stationary contexts.

The contribution of this paper is threefold: 1) Empirically investigated the role of dynamic drawing for text in improving fundamental learning outcomes such as recognition and recall, 2) Empirically compared the effect of text presentation styles on two usage conditions (desktop while sitting vs. OHMD while walking), and 3) Design implications for existing and future videos to support on-the-go learning tasks on desktop and OHMD-based video learning.

2 RELATED WORK

2.1 Dynamic Drawing in Traditional Desktop-based Learning

The effects of dynamic drawing on learning have been noted in literature to various degrees over time. However, the conclusions are debatable.

One set of results favored dynamic handwritten visuals. For instance, Guo et al., who compared different video production styles in terms of engagement, noticed that "Khan-style tutorial videos were more engaging than PowerPoint slides and/or code screencasts" [19]. Similarly, Fiorella et al. showed that watching an instructor's hand dynamically draw content for physics videos improved users' performance in transfer tests when compared to viewing the content in a static pre-drawn manner [12]. This could be because dynamic drawing naturally combines several multimedia learning principles that together improve learning [32]. By directing learners' attention to relevant on-screen information using cues, it follows the signaling principle. By sequentially presenting content, it reduces the overall cognitive demand on the learner, thereby obeying the segmenting principle. By temporally aligning the drawing creation with the instructor's explanation, it helps learners integrate the visuals with the audio, thereby adhering to the temporal contiguity principle. Further experiments indicated that these effects of dynamic drawing can be observed even without the presence of instructor hands/cursor, are not specific to subject domains [13] and can support generative learning strategies [14].

On the other hand, other investigations have shown that a significant portion of the participants prefer typeface fonts instead of handwritten text. An investigation that compared two variants of Khan academy videos, one that used the default dynamic handwritten visuals and another where the text was converted into typeface font, found that among 50 participants, 19 (37%) users preferred handwritten text since it was "more naturalistic" and reminiscent of classroom learning but 31 (63%) users preferred the typeface fonts owing to its clarity and professionalism [6]. This finding could also be governed by cognitive load theory which posits that the cognitive load on learners' working memory should be minimised for effective learning [37]. Typeface fonts, which are easier to read, results in a lower extrinsic cognitive load on users' working memory, as opposed to handwriting, which are potentially more difficult to read. Moreover, no significant engagement benefits of dynamic drawing was observed [13].

Taken together, these results suggest a knowledge gap in our understanding of the exact factors contributing to the relative strengths and weaknesses of dynamic drawing. In order to fill this knowledge gap, we conduct a systematic investigation of the underlying factors that could be responsible for the suggested benefits of dynamic drawing, with a specific focus on textual content. Moreover, current literature is limited to traditional desktop-based learning scenarios and the transferability of these effects to on-the-go video learning on OHMDs is still unclear.

2.2 Text Presentation in OHMDs for On-the-go Information Acquisition

Given the recent traction for on-the-go information acquisition on OHMDs, there has been renewed research interest to understand how information should be presented on the OHMD platform for improved acquisition on the go [17, 38, 39]. These works can be broadly separated into two categories.

The first category focuses on the static properties of text which include color, font, and layout that are suitable for viewing on OHMD. A comparison of different text styles against varying realistic backgrounds showed that designs with contrasting text and background colors can improve legibility in outdoor environments [15, 16]. Investigations comparing different font styles found that thin horizontal and vertical lines can be difficult to perceive due to the impact of head-shaking [31]. In terms of text layout, Chua et al. noted that for improved noticeability in dual-task scenarios, the text should be displayed in the middle-center or bottom-center portions of the display [4].

The second category has explored dynamic aspects of text appearance on OHMDs. For simple OHMD-based reading tasks, Rzayev et al. noted that text presented using an RSVP mode of presentation improves reading comprehension while users sit, whereas scrolling is a better alternative for reading while users walk [39]. Other work investigating peripheral vision for reading showed that word-based typewriter animation improved users' reading accuracy [28]. In contrast, when viewing more cognitively demanding dynamic content such as educational videos on the go, presenting content sequentially and allowing it to persist on the screen, can improve users' recall and recognition with minimal impact on their walking capabilities [38].

Complementary to previous work, our focus is on the font and dynamic aspects of text appearance, focusing on how different text presentation styles can impact learning on OHMDs in on-the-go situations.

3 STUDY 1: COMPARISON OF TEXT PRESENTATION STYLES

Font style	Motion	t	t +
Handwritten	Appear-letter		lor
	Appear-word		lorem
	Trace-letter		lorem
Typeface	Appear-letter		lor
	Appear-word		lorem
	Trace-letter		lorem

Figure 1: The six text presentation styles compared in Study 1 to understand the impact of each font and motion factor. The dotted lines indicate the path of tracing motion. Refer to the video figure for a better understanding of the styles.

We began our investigation by breaking down dynamically drawn text into its two fundamental factors, namely the Font style and Motion of appearance of the letter or word.

Font. This describes the style of the text. It could either be a naturalistic style created using a digital pen (Handwritten) or standard printed font typically used in PowerPoint presentations (Typeface).

Motion. A characteristic property of dynamic handwriting is that each letter is traced out in a continuous manner. Although it is well known that motion can draw users' attention and improve learning [22], it is unclear whether continuous tracing of each letter (Trace-letter) has any benefits over the discrete letter-by-letter display (Appear-letter) or the more common way of presenting text which involves displaying the entire word at once (Appear-word).

Given that the relative importance of either factor for dynamic drawing is unclear, we conducted a controlled study with the following research questions and hypotheses.

3.1 Research questions and Hypotheses

RQ1: What impact does font style have on users' learning performance and cognitive demand in different usage contexts?

In on-the-go situations, handwritten fonts can be significantly more difficult to read than typeface fonts due to the shakiness of the OHMD display caused by users' expected physical movements. This can impose an extraneous cognitive load on the learner, which according to cognitive load theory [37] leads to a decline in learning performance. However, for stationary scenarios where the impact of legibility is lesser, we could observe the learning-related benefits of handwriting [13]. Thus, we hypothesize that

H1.1: Handwritten font will decrease learning scores (recall/recognition) compared to typeface font while walking using an OHMD

H1.2: Handwritten font will impose higher cognitive demand (PPWS/NASA-TLX) on users compared to typeface font while walking using an OHMD

H1.3: Handwritten font will improve learning scores (recall/recognition) over typeface font in stationary contexts using a desktop.

H1.4: Handwritten font will be no different than typeface in terms of cognitive demand (PPWS/NASA-TLX) in stationary contexts using a desktop.

RQ2: What is the impact of text motion on learning performance in different usage contexts?

Both appear-letter and trace-letter motions introduce a word letter-by-letter and thereby follows the multimedia learning principles of segmenting, temporal contiguity and signaling discussed earlier [32]. However, the tracing motion of trace-letter may be a more effective way of realising the signaling principle as it continuously evokes users' attention. Hence, we hypothesize that

H2: Trace-letter will improve learning scores (recall/recognition) over Appear-letter and Appear-word irrespective of usage context.

3.2 Experiment Design

A within-subject design with 2 Fonts styles (Handwritten, Typeface) x 3 Motions (Appear-letter, Appear-word, Trace-letter) x 2 Usage contexts (Stationary using desktop, On-the-go using OHMD) was used, which resulted in 12 conditions per participant. A balanced Latin square design counterbalanced the order of Font style and Motion variables blocked by usage context. Since our focus is not to compare the performance results under different levels of difficulties, we follow previous approaches [49] by administering the usage context in the increasing order of difficulty, i.e. stationary using desktop followed by on-the-go using OHMD.

3.2.1 Usage context. We considered two contexts: (1) a baseline stationary scenario where the user watched videos on a desktop monitor while seated and (2) an on-the-go context which required the user to walk along a path taped on the floor (36 meters long and 30 cm wide) similar to that used by Vadas et al. [44]. We chose this path as it requires both simple (during the straight portions of the path) and complex motor skills (during the 8-figure segment of the path) [21]. Following Barnard et al., the direction the participants walked on the path (clockwise or counter-clockwise) was randomized to minimize learning effect [1]. Given that video visibility on the OHMD is affected by outdoor lighting [15], all scenarios were designed in indoor lighting conditions to minimize any platform-specific bias due to video visibility.

3.2.2 Video design. To remove the confounding effects present in real learning videos, we transformed the video into a representative skeleton format where the diagram consisted of a square box with the numbers 1-8 presented as shown in Figure 2. Each number in Figure 2(b) was associated with a word using a line and the number-word pair constituted a concept as in the real learning video shown in Figure 2(a).

It was also necessary for the words to be unfamiliar for all users. Hence, we used pseudowords generated from an artificial corpus similar to Macedonia and Knösche [30]. The pseudowords were generated using the Wuggy pseudoword generator as it facilitated the generation of polysyllabic pseudowords which followed English phonotactic constraints [23]. These constraints ensured that the pseudowords were pronounceable by participants fluent in English. All pseudowords had a word length of 5 as this is the average word length of the English language [45]. For uniformity, the audio consisted of the word pronunciations spoken by a text-to-speech service. Each word was pronounced twice, once at the beginning of its appearance and once at its end.

We created the six text presentation styles (shown in Figure 1) using a video editing tool[1]. A sans serif font was used for the typeface text as recommended by previous work on text readability on OHMDs [31]. The handwritten font was created using an Apple Pencil on an iPad. A fully saturated green hue was used for the text as previous research suggests the global effectiveness of green color for viewing on OHMD [15]. Since Khan academy videos took an average of 5 seconds to write a 10-lettered word, for all text presentation styles, we set 2.5 seconds to display the 5-lettered word.

3.3 Measures

The following dependent variables were measured to assess users' learning performance, cognitive load and their perceptions about the various text presentation styles.

3.3.1 Learning Performance. We collected two measures to assess learning performance immediately after the stimuli.

Recognition. This tests users' ability to recognize whether a given piece of information is familiar. This is an easier "remember" level task than recall due to the availability of more cues for memory retrieval [41]. The recognition test consisted of multiple-choice questions. The normalized participant's score (0-1) was used as a measure of their recognition ability.

Recall. This tests users' ability to retrieve information from memory with no or minimal cues provided to aid the memory retrieval process. Users were first asked to write down as many words as they could (free recall) and then answer listing questions such as "(1) is called" (cued recall). The normalized participant's score (0-1) was used as a measure of their recall ability.

Scoring of Quiz. For all tests, correct answers received 1 point. Following Ram et al., if the word spelling was incorrect in the recall task, we calculated the error using the Levenshtein distance which counts the minimum number of deletions, insertions, and substitutions needed to rectify the spelling [38]. The word was marked incorrect if the error was greater than half the word length. For instance, if the word 'grepo' (word length = 5) was misspelled as 'gerop', no points were awarded as the words differed by a Levenshtein distance of 3 units, which is greater than half the word length.

3.3.2 Cognitive Load. We measured cognitive load using two aspects.

Percentage of Preferred Walking Speed (PPWS). This measures how much slower a user walks as compared to their normal walking speed. A lower PPWS is indicative of a higher cognitive load imposed by the stimuli [17]. To obtain PPWS, users' normal walking speed (in m/s) is measured by recording the time taken by them to walk a fixed distance without administering any stimuli. The PPWS (0-100%) is then obtained by dividing the walking speed during the task by the normal walking speed.

NASA-TLX. Participants filled out an unweighted NASA-TLX questionnaire to report their subjective task load after each video condition.

3.3.3 Subjective Measures. A post-experiment questionnaire item was provided after each usage context block, asking participants to rank the text presentation styles they viewed. A subsequent semi-structured interview captured the reasons for their ranking and the process each individual followed to complete the task under different conditions.

3.4 Participant & apparatus

12 participants (7 female; M = 23.83 years, SD = 1.89 years) who had at least self-reported professional working fluency in English were recruited from within the university community. None had previous experience with OHMDs. Each participant was compensated ≈ USD 7.33/h for their time.

For the study, we used an Epson BT-300 binocular optical see-through OHMD, which places the display in the center of the user's line of sight. It has a 1280x720 px resolution display, 23° FoV with a projected distance of 80 inches at 5m. The videos were pre-loaded and viewed in an offline manner on the default video player application in Epson. For the stationary context the video was displayed on a MacBook Pro 13.3 inch computer (resolution = 2560 x 1600 px)

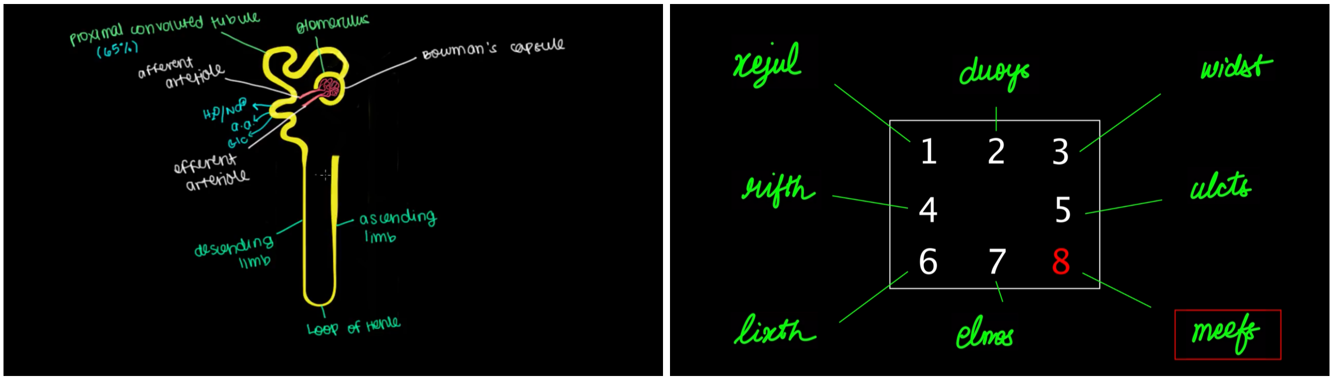
3.5 Procedure

Each participant performed the experiment in one session lasting approximately one hour. The session was blocked by usage context, with a participant watching the six different text presentation styles in each context.

The experiment started by providing participants with instructions regarding the tasks. Next, they were provided with an initial warm-up session where they watched a demo video and attempted a quiz to familiarize themselves with the task.

After watching each video, participants took a quiz that tested their recognition and recall based on the content learned in that video. After each usage block, participants had to fill up the post-questionnaire to rank the videos they watched in that block. The

[1] Link to study material: https://bit.ly/3v4sgrq



(a) (b)

Figure 2: (a) A real Khan Academy learning video [25] contrasted with (b) an equivalent artificial pseudoword video. Each number and its corresponding word is highlighted in red as the word is pronounced.

experiment ended with an interview to understand users' preference of the text presentation styles.

4 STUDY 1: RESULTS

To analyse the results we applied factorial repeated measures analysis of variance (ANOVA) (if parametric and normal) or transformed the data using Aligned Rank Transform (ART) (if non-parametric or non-normal, [46]) for evaluation. Normality and sphericity were tested using the Shapiro-Wilk and Mauchly test. For violations of sphericity, Greenhouse-Geisser ($\epsilon < 0.75$) corrections were used to adjust the degrees of freedom. Post-hoc contrast tests on the data transformed with ART were performed using the ART-C approach [10]. Interviews were transcribed and coded thematically by the author. Themes were inductively crafted from participants' statements based on both their frequency of occurrence and perceived significance.

4.1 Learning Performance

An ANOVA was performed individually on the ART scores of recognition and recall scores ~ Font x Motion x Usage context. The results are shown in Figure 3.

4.1.1 Recognition. A significant main effect was seen in the Font ($F_{1,11} = 10.60, p = 0.001, \eta_p^2 = 0.08$) and Usage ($F_{1,11} = 24.87, p < 0.001, \eta_p^2 = 0.17$). For font, typeface ($0.72 \pm 0.25$) was better than handwritten (0.60 ± 0.28) ($p_{bonf} = 0.001, d = 0.54$). No significant interaction effects were observed.

4.1.2 Recall. The ANOVA revealed a significant main effect for Font ($F_{1,11} = 21.16, p < 0.001, \eta_p^2 = 0.66$), Motion ($F_{2,22} = 3.99, p = 0.033, \eta_p^2 = 0.26$) and Usage ($F_{1,11} = 6.51, p = 0.027, \eta_p^2 = 0.37$). Overall, scores were higher when stationary (0.46 ± 0.25) than while walking (0.35 ± 0.18) ($p_{bonf} < 0.001, d = 1.328$). Participants scored higher using typeface font (0.45 ± 0.22) compared to handwritten font (0.35 ± 0.22) ($p_{bonf} = 0.027, d = 0.73$). In addition, scores were lower when using the trace-letter motion (0.37 ± 0.19) than appear-word (0.44 ± 0.23) ($p_{bonf} = 0.019, d = 0.97$).

A significant interaction effect between Font x Usage ($F_{1,11} = 12.21, p = 0.005, \eta_p^2 = 0.52$) was found. Post hoc analysis showed that typeface (0.43 ± 0.15) significantly outperformed handwritten (0.36 ± 0.17) in the on-the-go context ($p_{bonf} < 0.001, d = 1.66$).

4.2 Cognitive Load

4.2.1 Percentage of Preferred Walking speed. The ANOVA showed a significant main effect of Font ($F_{1,11} = 5.43, p = 0.04, \eta_p^2 = 0.33$). Participants had significantly higher walking speed using typeface (0.70 ± 0.18) font than handwritten (0.73 ± 0.16) ($p_{bonf} = 0.04, d = 0.67$), showing that handwritten font imposes higher cognitive load during on-the-go situations. No significant interaction effects were observed.

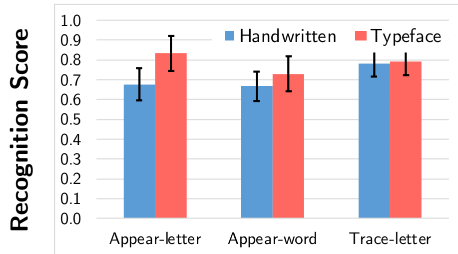
4.2.2 NASA-TLX. ANOVA performed on the overall unweighted NASA-TLX scores showed a significant main effect for Font ($F_{1,11} = 17.57, p = 0.002, \eta_p^2 = 0.61$), Motion ($F_{1.32,14.54} = 7.03, p = 0.013, \eta_p^2 = 0.39$) and Usage ($F_{1,11} = 19.91, p < 0.001, \eta_p^2 = 0.64$). Overall, typeface ($40.76 \pm 18.16$) had a lower task load than handwritten fonts (52.25 ± 21.23) ($p_{bonf} = 0.002, d = 1.21$). In addition, appear-word motion (45.32 ± 20.72) imposed less cognitive demand than trace-letter (48.65 ± 19.78) ($p_{bonf} = 0.043, d = 0.83$). Significant interaction effects were also seen in Font x Usage ($F_{1,11} = 6.60, p = 0.026, \eta_p^2 = 0.37$), with post-hoc analysis showing higher task load using handwritten text (64.88 ± 17.73) than typeface (50.64 ± 17.31) while walking ($p_{bonf} = 0.001, d = 1.39$).

4.3 Subjective Preference

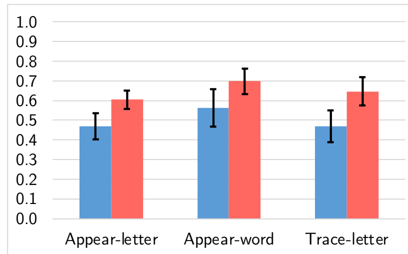
In both usage conditions, most users (10 of 12) preferred typeface with appear-word motion as shown in Figure 3e and this style was also within the top-3 choices for all users. Furthermore, a majority of the users opted for typeface-based styles within their top-3 choices.

4.4 Discussion

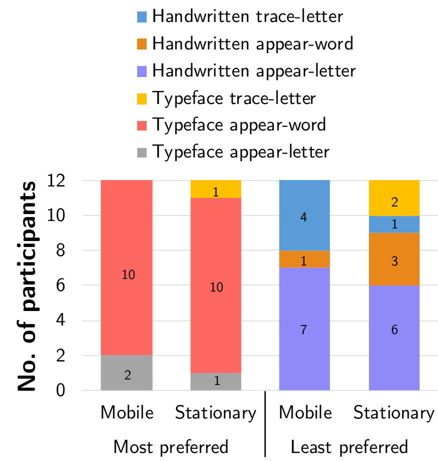
We organized the discussion below into parts that answer each of the research questions that we put forth earlier.



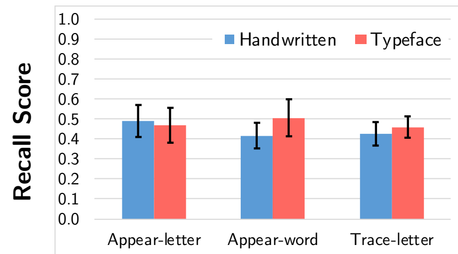
(a) Recognition: Stationary using Desktop



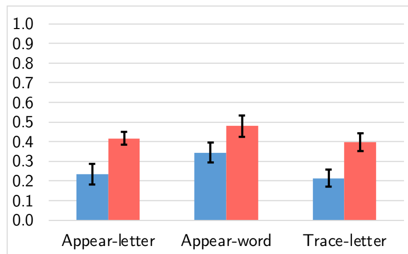
(b) Recognition: On-the-go using OHMD



(e) Preference of text writing styles



(c) Recall: Stationary using Desktop



(d) Recall: On-the-go using OHMD

Figure 3: Measured learning performance for recognition (a)-(b), recall (c)-(d) and users' preference of text presentation styles (e) in different usage contexts for Study 1.

RQ1: What impact does font style have on users' learning performance and cognitive demand in different usage contexts?

In our hypothesis H1.1, we had predicted that handwritten fonts would decrease learning performance when used in walking context. Our results support this, showing a decline of 39.5% in recall when using handwritten fonts as compared to typeface. As evidenced by participants' comments, this was because the handwritten font style was difficult to focus on and read especially when the OHMD display was shaking: "I could barely read the words… also the screen was constantly shaking making it worse." (P1). This also led to increased cognitive load and slower walking speeds while using handwritten fonts, supporting H1.2.

Contrary to our expectations, the handwritten font did not show benefits over typeface in terms of learning even in stationary scenarios. This suggests that the free-style appearance of handwritten text may not be a critical contributor to the benefits of dynamic drawing. While a few participants described the handwritten font style to be more "engaging" as reported by Cross et al. [6], they still preferred the typeface font for learning, and thus does not provide support for H1.4. This indicates that legibility is of primary importance to learners irrespective of usage context.

These findings support prior research by Turner [42] who showed the superiority of typeface fonts to cursive handwriting in terms of legibility. While Turner's findings were considered in non-digital (paper-based) contexts, our results extend them to digital contexts, while also shedding light on its impact on users' learning in different usage contexts.

RQ2: What is the impact of text motion on learning performance in different usage contexts?

Overall, we were surprised to find out that recall scores were higher when the whole word appeared immediately than when the letters were traced out. This finding was also reflected in participants' preferences, with most of them finding trace motion to be "slow" for the learning task as opposed to appear-word motion which was of the "right speed": "[for trace motion videos] the words took so much time to appear... I just wanted it to go faster" (P5).

Further research into visual chunking has pointed out the possible reason behind the surprising results. The theory of visual chunking states that our visual working memory encodes the components of stimuli and the relations between them as a single representation or chunk rather than as separate memory representations [35]. In our case, since appear-word motion presents the entire word to the learner, it naturally facilitates chunking and encoding by the visual working memory. On the other hand, with trace-letter or appear-letter motion, learners tend to focus on the text in a letter-by-letter manner which potentially hampers the process of chunking. In fact, participants who disliked these styles expressed a similar feeling: "I felt that I was focusing on each letter as it appeared but later on (during the test) I couldn't remember a lot of words"(P2).

Moreover, there was no significant difference in learning scores between the appear-letter and trace-letter motion, even in the stationary scenarios. This does not provide support for H2 and suggests that the trace motion is no more advantageous than immediate appearance motions in terms of learning performance. Although a few participants (3/12) found the tracing motion to be "naturalistic" when combined with handwritten font, for most participants legibility was more important for learning: "[Tracing] guides me during learning, but I still prefer typeface font as it's easier to read".

In addition, trace motion was found to be "unnatural" when paired with a typeface font.

In summary, our quantitative results suggest that handwriting has no particular advantage and that typeface font along with appear-word motion is preferable in any usage situation. This result is different from previous studies, which indicated that dynamic drawing improves learning outcomes and engagement. We suspect that two factors may have caused the observed result.

- First, the preference for appear-word motion could be due to the word length (length=5) used in our learning content as five-lettered words consist of at most two syllables which form a single chunk. However, this result may not hold for longer polysyllabic words which consist of multiple chunks, in which case tracing motion might still offer a chunk-wise focus.
- Second, the audio content in our videos was controlled and consisted only of the pronunciation of the words. On the other hand, in generic learning videos, a significant amount of the content is expressed via audio. In such situations, tracing motion may help guide users' attention as they experience mind wandering [11].

Given the need for further investigation to delineate any benefits of handwriting, we conducted a follow-up study using real learning videos.

5 STUDY 2: COMPARING TEXT PRESENTATION STYLES IN REAL LEARNING VIDEOS

To investigate the external validity of the result of Study 1, we conducted another controlled study comparing text presentation styles with trace motion against the predominantly preferred style in Study 1, i.e. typeface text with appear-word motion. This time, we use real learning videos as testing material.

We tested two variants of this study. In the more controlled variant, we modified the handwriting in the original Khan Academy videos to ensure a fair comparison with the typeface text. In particular, the handwriting was recreated at a larger size and thickness to match the typeface text properties fixed from the earlier study. We refer to this variant as the controlled font (CF) variant.

A potential confound in the CF variant is that the degree of cursiveness of the created handwritings were difficult to control. In the original Khan academy videos, the handwriting style is heavily dependent on the instructor and varies across the spectrum of manuscript to cursive writing (with inclination towards cursive in most videos). Although we adopted a handwriting style that is moderately cursive in the CF variant, it is possible that the less cursive handwritings will be more preferred by users. To clarify this, we conduct another variant of the study that is identical to CF Variant with the exception that the handwritten trace-letter videos now use the original handwriting in the Khan Academy videos which varies from manuscript style text to quite cursive style of handwriting (see Fig. 2(a)). We refer to this variant as the original font (OF) variant.

5.1 Research questions and Hypotheses

RQ1: How does text presentation style impact users' learning performance in real learning videos under different usage contexts?

Polysyllabic words in real learning videos consist of multiple chunks, for which the trace-letter motion could help provide a chunk-wise focus to users. In addition, the continuous tracing motion could facilitate the signaling principle more effectively by guiding users' attention as they experience mind wandering. Hence, we hypothesize that

H1: Text presentation styles with trace-letter motion will allow users to learn better (recall/recognition) than using typeface appear-word style, irrespective of usage contexts.

RQ2: How is user preference affected by the choice of text presentation style for real learning videos?

As noted by prior literature, tracing could act as a social cue that can make lecture videos more engaging by building a feeling of partnership between students and instructor [33]. But as we observed in the earlier study, this preference may only hold for handwritten fonts. Thus, we hypothesize that

H2: Handwritten trace-letter style will be preferred by users for video learning, irrespective of usage contexts.

5.2 Experiment Design

Both variants of the study followed a within-subject design with 3 text presentation Styles (Handwritten trace-letter, Typeface trace-letter, Typeface appear-word) and 2 Usage contexts (Stationary using desktop, On-the-go using OHMD) was used, which resulted in 6 conditions per participant. The order of styles was fully counterbalanced and blocked by usage context. With the exception of the video design which is described below, the rest of the experiment design, apparatus, and procedure is identical to that in Study 1.

5.2.1 Video design. We chose biology as the subject domain for the study as it contains more declarative knowledge [34], making them well-suited for assessing users' recall and understanding skills. The videos were adapted for on-the-go video learning on OHMDs by following the guidelines of LSVP, a video style that has been proven to work well for users to watch on OHMDs (smart glasses) [38]. Educational videos are recommended to be kept to a duration of 6 minutes or lesser in order to maximize engagement [2]. Hence, we set the video duration to be 3 minutes so that the videos have sufficient learning content for assessing users' learning while still allowing users to view the videos with minimal fatigue and complete engagement.

To choose videos of similar difficulty, we chose 8 biology videos of similar length, style, and difficulty of content from the Khan Academy YouTube channel. Next, we verified whether users can achieve a similar level of learning performance in each video. For this, we conducted a pilot study with 6 participants who were asked to watch and learn from the videos. After each video, they were asked to answer a recall/recognition quiz that tested their learning. The quiz was scored independently by two researchers (one author) using a predefined marking schema. Through this process, two biology videos were removed since participants scored significantly higher in these videos than the rest. The remaining videos were used as the learning material in the study. To reduce any further confounding effects, we also counterbalanced the videos across the different text presentation styles.

In both variants, the videos were adapted for on-the-go video learning on OHMDs by following the guidelines of LSVP, a video style that has been proven to work well for users to watch on OHMDs (smart glasses) [38]. The design of typeface appear-word and typeface trace-letter videos is similar to that in Study 1.

In case of handwritten trace-letter videos, in the CF variant the handwritings were recreated using an apple pencil at a size and thickness that matched the typeface text to ensure a fair comparison. In the OF variant, we maintained the handwriting from the original videos while only enlarging its size to match the typeface text. It should be noted that naive upscaling of the text in the video will result in blurring of the text. To preserve the video quality of the handwriting while enlarging it, we used detail-preserving upscale technique [7] along with an unsharp mask [43].

5.3 Dependent measures

We reused the immediate recognition and recall tests from Study 1 to measure learning performance. These tests allowed us to assess the participants at the Remember level of Bloom's Taxonomy [27] that is fundamental for learning. The recognition quiz consisted of multiple-choice questions with a single correct answer. The recall test consisted of a single open-ended question following Dori and Belcher: "Describe (in detail) what you learnt in this video." [8]. This form of free recall assessment provided a comprehensive indication of both the factual and conceptual knowledge users gained from the video.

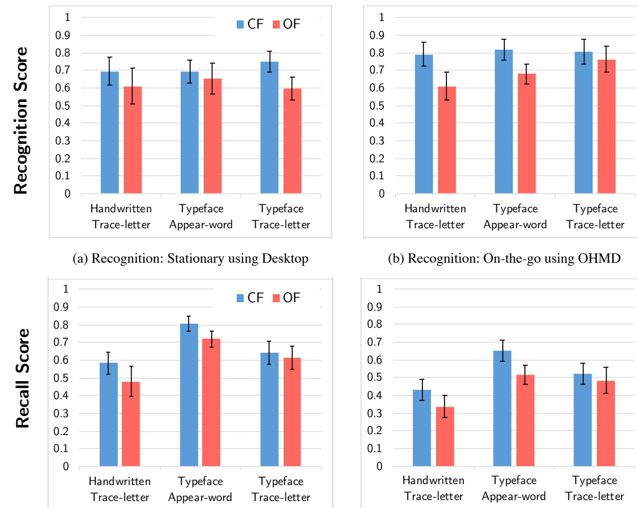
Scoring of Recall Quiz. Similar to Fiorella et al., we created a scoring rubric to assess recall [13]. One point was awarded for correctly recalling each conceptual keyword and its relationships. For example, one point was awarded for mentioning the term "sinoatrial node" and another point for describing its function as "an area of self-depolarizing cells" or any other semantically similar sentence. For misspelled words, the marking procedure follows that in Study 1.

5.4 Participants & apparatus

12 participants (6 female; M = 22.4 years, SD = 5.1 years) for the CF variant and 12 additional participants (6 female; M = 20.7 years, SD = 1.6 years, none of them participated in the CF variant) for the OF variant were recruited for the experiment from within the university community. None of them participated in our earlier studies or pilots. All participants had at least self-reported professional working fluency in English. 10 users in CF variant and 11 in OF variant had prior experience learning a course fully or mostly online. To minimize any potential bias due to prior knowledge, participants were ensured to be from a non-biology-based background. Each participant was compensated ≈ USD 7.33/h for their time. The apparatus was the same as in Study 1.

6 STUDY 2: RESULTS

The results analysis procedure was similar to study 1. Given the closely related nature of the study variants, we analysed the results together and present them in the format of CF variant stats followed by OF variant stats (i.e., $(p = 0.826)_{CF}; (p = 0.499)_{OF}$ indicates a p value of 0.826 for CF variant, and p value of 0.499 for OF variant). As a reminder, here **CF** refers to the variant where the font was controlled, and **OF** refers to the variant which uses the original Khan Academy handwriting style.



(a) Recognition: Stationary using Desktop (b) Recognition: On-the-go using OHMD

(c) Recall: Stationary using Desktop (d) Recall: On-the-go using OHMD

Figure 4: Measured learning performance for recognition (a)-(b) and recall (c)-(d) for CF and OF variant in Study 2. The difference in recognition scores is not significant.

6.1 Learning Performance

A repeated measures analysis of variance (ANOVA) was performed on the recognition and recall scores. The results are shown in Figure 4a-d.

6.1.1 Recognition. No significant main effect for Style was found in the recognition scores between different text presentation styles $(p = 0.826)_{CF}; (p = 0.554)_{OF}$. Overall, users scored slightly higher using typeface trace-letter $(0.78 \pm 0.22)_{CF}; (0.68 \pm 0.24)_{OF}$ than typeface appear-word $(0.76 \pm 0.22)_{CF}; (0.67 \pm 0.25)_{OF}$ and handwritten trace-letter $(0.74 \pm 0.25)_{CF}; (0.61 \pm 0.31)_{OF}$, but the high p value suggests that the difference is more likely to due to random chance, and they are comparable.

6.1.2 Recall. In both variants of the study the ANOVA revealed a significant main effect for Style $(F_{2,22} = 8.44, p = 0.002, \eta_p^2 = 0.43)_{CF}; (F_{2,22} = 16.28, p < 0.001, \eta_p^2 = 0.59)_{OF}$ and Usage $(F_{1,11} = 35.24, p < 0.001, \eta_p^2 = 0.76)_{CF}; (F_{2,22} = 21.91, p < 0.001, \eta_p^2 = 0.66)_{OF}$.

Post hoc tests with bonferroni correction showed that using typeface appear-word $(0.73 \pm 0.19)_{CF}; (0.62 \pm 0.19)_{OF}$ significantly improved recall over handwritten trace-letter $(0.51 \pm 0.22)_{CF}; (0.41 \pm 0.26)_{OF}$ $(p_{bonf} = 0.004, d = 1.24)_{CF}; (p_{bonf} = 0.002, d = 1.38)_{OF}$. Recall scores were also significantly higher for typeface trace-letter $(0.55 \pm 0.24)_{OF}$ as compared to handwritten trace-letter $(p_{bonf} = 0.014, d = 1.02)_{OF}$. Users also performed better with typeface appear-word than typeface trace-letter $(p_{bonf} = 0.077, d =$

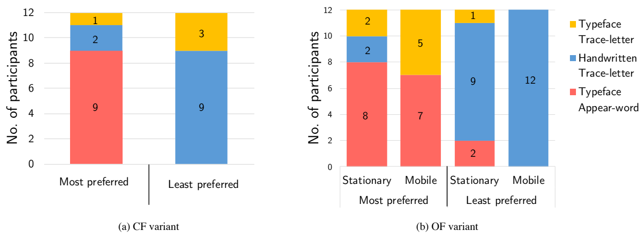


Figure 5: Users' preference of text presentation styles for (a) CF variant and (b) OF variant in Study 2. User preference is identical in both usage contexts in the CF variant.

$0.74)_{CF}; (p_{bonf} = 0.076, d = 0.74)_{OF}$, indicative of a statistical trend. No interaction effects were found from the data.

6.2 Cognitive load

6.2.1 Percentage of Preferred Walking speed. No significant differences were observed in the walking speed between different text presentation styles $(p = 0.644)_{CF}; (p = 0.462)_{OF}$. No consistent trends were noticeable between handwritten trace-letter $(0.71 \pm 0.13)_{CF}; (0.66 \pm 0.11)_{OF}$, typeface trace-letter $(0.69 \pm 0.13)_{CF}; (0.67 \pm 0.09)_{OF}$ and typeface appear-word $(0.69 \pm 0.13)_{CF}; (0.64 \pm 0.09)_{OF}$, which is likely due to random variations.

6.2.2 NASA-TLX. The ANOVA on the overall unweighted NASA-TLX scores showed a significant main effect of Style in the CF variant and a statistical trend in OF variant $(F_{2,22} = 5.47, p = 0.012, \eta_p^2 = 0.33)_{CF}; (F_{1.24,13.59} = 3.46, p = 0.078, \eta_p^2 = 0.24)_{OF}$. Usage was significant in both variants $(F_{1,11} = 19.88, p < 0.001, \eta_p^2 = 0.64)_{CF}; (F_{1,11} = 13.48, p = 0.004, \eta_p^2 = 0.55)_{OF}$.

Overall, typeface appear-word $(44.86 \pm 13.80)_{CF}; (39.05 \pm 17.81)_{OF}$ had a significantly lower task load than handwritten trace-letter $(51.99 \pm 13.07)_{CF}; (45.25 \pm 18.76)_{OF}$ in the CF variant $(p_{bonf} = 0.049, d = 0.81)_{CF}; (p_{bonf} = 0.23, d = 0.56)_{OF}$. Typeface trace-letter $(50.0 \pm 13.21)_{CF}; (41.14 \pm 16.46)_{OF}$ also had a lower task load than handwritten trace-letter in OF variant $(p_{bonf} = 1.00, d = 0.26)_{CF}; (p_{bonf} = 0.06, d = 0.77)_{OF}$, indicative of a statistical trend. A similar trend was observed with typeface appear-word being lower than typeface trace-letter $(p_{bonf} = 0.064, d = 0.77)_{CF}; (p_{bonf} = 1.00, d = 0.27)_{OF}$ in the CF variant.

6.3 Subjective preference

Most users $(9 \text{ of } 12)_{CF}; (8 \text{ of } 12)_{OF}$ preferred learning with the typeface appear-word style as shown in Figure 5. On the other hand, the most disliked style was handwritten trace-letter primarily due to issues of legibility. However, a few users $(2 \text{ of } 12)_{CF}; (2 \text{ of } 12)_{OF}$ mentioned that the handwritten trace-letter style felt more engaging and preferred them more than the typeface-based styles.

6.4 Discussion

RQ1: How does text presentation style impact users' learning performance in real learning videos in different usage contexts?

Using typeface appear-word style resulted in a 46.7% improvement in recall scores on average than handwritten trace-based styles, against H1. While similar improvements were noted in Study 1 for mono/disyllabic words, insights from interviews in the current study suggest that a deeper reason may be at play when dealing with polysyllabic words. In particular, this improvement may have more to do with how users process and learn a new word than with visual chunking.

Users processed a new word in two different ways. Initially, they automatically formed a phonemic representation on hearing the word via the audio, as suggested in speech comprehension models [36]. This speech-based representation, however, is often erroneous or incomplete especially when encountering unfamiliar words, for e.g the word "sinoatrial" may be interpreted as "cynoareal". In such cases, users naturally establish another more accurate representation by reading the text from the display [9]. Users then tried to resolve any mismatch between the two representations using the reading-based representation as a reference while learning the word: "I think it spells a certain way after hearing, then I check the text [by reading] to see if it matches… this [process] was better when the whole word appeared at once" (P6, CF variant). The appear-word style facilitated this correction process by allowing users to form the reading-based representation for reference immediately as opposed to trace-based styles which forced users to create the representation in a part-by-part manner.

RQ2: How is user preference affected by the choice of text presentation style for real learning videos?

The overall results from study 2 are consistent with our findings in study 1 showing that even in real lecture videos, most users (17 of 24 for stationary and 16 of 24 on the go) prefer to learn with a typeface font which displays the entire word immediately. This was because users found the attention evoked by tracing motion felt undue, especially in stationary scenarios: "My attention is on the word being written out, but then I miss out on the [instructor's] explanation" (P11, OF variant). This anchoring of attention to the word while the oral explanation continued, created a "lag" in how users processed information. Consequently, most users felt that trace motion did not improve their learning.

However, we did find substantial evidence in support of H2 indicating the benefits of trace motion, i.e. several users (7 of 24 for stationary and 8 of 24 on the go) found tracing motion to be more engaging and evoking attention to incoming words: "It was more engaging [and] naturalistic like how a professor teaches" (P3, CF variant) "I feel it [tracing] helps me pay more attention to the word being written out " (P4, OF variant). Yet the fact that users prefer to immediately see the entire word whereas trace-based styles delay the appearance of the word, ultimately outweighed this engagement effect for most users. "It was kind of annoying to wait for the word to finish. I liked the video where the text appears immediately" (P1, CF variant).

In summary, this study reinforced our findings from study 1, showing that handwriting does not improve learning outcomes. Despite being engaging for some, handwriting was not preferred by most users as the tracing motion was found to interfere with users' processing of new words. Consequently, the typeface word-appear style was the most preferred text writing style for users and it improves recall.

Usage Condition	Topic Familiarity	Recommended Text Presentation Style		
		Handwritten Trace-letter	Typeface Trace-Letter	Typeface Appear-word
Stationary using Desktop	Familiar	✓	×	✓
	Unfamiliar	×	×	✓
On-the-go using OHMD	Familiar	×	✓	✓
	Unfamiliar	×	✓	✓

Figure 6: Suitability[2] of text presentation styles depending on the usage context and familiarity of the learning content.

7 OVERALL DISCUSSION & LIMITATIONS

Taken together, our results from both studies provide a more comprehensive answer to the question of whether dynamic handwritten text has benefits over typeface text for video learning. By systematically analyzing the fundamental characteristics of text presentation i.e., font and writing motion, study 1 provided preliminary evidence that typeface font coupled with whole word appearance improved users' learning and is the most preferred text presentation style for videos. Testing such fine-grained characteristics of text presentation, however, came at the cost of using artificial pseudoword-based videos (to avoid confounding factors) which may have suppressed the benefits of handwritten text observed in real learning videos, in particular, that of tracing motion. To overcome these limitations, in study 2 we tested the impact of text presentation styles on real videos using two study variants and found that while the initial results hold true, tracing motion (with handwritten font) does make the video more engaging. However, most users dismissed this engagement factor in favour of getting the entire text quickly as it facilitated their learning, especially when dealing with unfamiliar words and concepts.

Our findings support the predictions of cognitive multimedia learning theory [32]. In particular, while both appear-word and trace-motion styles incorporate multimedia learning principles such as, segmenting visuals for sequential presentation, temporally matching the visuals with oral explanation and signalling the information being referenced in the visuals, it was typeface appear-word that improved recall. This suggests that immediate appearance might be more effective in terms of the temporal contiguity principle than tracing motion, allowing learners to quickly integrate the oral explanation with the visuals. Furthermore, our findings indicate that tracing motion can act as a social cue, improving learners' engagement. This supports the social agency theory [33], which posits that social cues can foster a feeling of partnership with the instructor, thereby increasing their motivation to engage with the learning material.

Figure 6 summarizes how our results can be used to help create future educational videos. In terms of suitability of different text presentation styles based on the usage context of the learning videos, we found that for on-the-go conditions involving walking, legibility is of primary concern. In particular, if users use an OHMD-based heads-up platform for on-the-go video learning, the text should be presented such that information acquisition is possible despite the shaking of OHMDs and divided attention between the environment and OHMD display. In this regard, the illegibility of handwritten fonts can negatively impact users' learning. Instead, we recommend using typeface fonts owing to its improved readability in on-the-go situations.

In terms of motion of appearance of text, a few users preferred trace-letter (with typeface font) as it helped them focus their attention and engage better with the content on the go, but the consensus favoured appear-word motion and it also improved recall. Hence, the results suggest that typeface trace-letter is a viable alternative especially for content that requires more focus and engagement, albeit with lesser emphasis on long term remembrance (e.g. providing navigation instructions). However, when learning is of primary concern, we recommend using typeface fonts with whole word appearance as a standard text format, with information presented following the guidelines of LSVP [38] for future educational videos that aim to support on-the-go video learning on OHMDs.

For stationary desktop-based learning, on the other hand, designers can also consider using handwritten font style with trace-letter motion in order to enhance lecture engagement. But given that the delay in word appearance caused by tracing motion hinders users' processing of unfamiliar text, we suggest using this style only when dealing with topics that are more familiar to the user. When creating videos on topics that users will encounter for the first time or when unsure of the level of audience familiarity, the typeface font with whole word appearance would serve as a better alternative, and thus can be treated as a general-purpose text presentation style for video learning.

7.1 Limitations

Despite studying the effect of text-based dynamic drawing in-depth, our approach has limitations. Firstly, incentivising participants in our studies may have acted as an extrinsic motivation for them to focus and learn the lecture content. In real online learning settings, however, students may also need to be intrinsically motivated, in which case engagement may play a more critical role. Additional investigations in such unsupervised settings are necessary to better understand this impact of engagement.

Secondly, we limited our exploration to textual content in videos as reading is an integral part of video learning, on the go or otherwise. In addition, we focused on the biology subject domain to understand how learners' fundamental recognition and recall skills are affected by text writing styles. Further research is necessary to establish the impact of dynamic drawing on images/diagrams in videos and to understand whether these findings extend to other subject domains such as mathematics where higher learning objectives need to be acquired.

Another limitation is the smaller sample size (n=12) considered in our studies, which imposes the risk that the observed large effect is due to chance. Although such sample size (n=12) is not uncommon in the CHI literature [18, 29, 47, 48], readers should take this possibility into consideration when using our results. We also considered head-locked content in which case the shakiness of the display (due to head movement) can decrease text readability. World-locked content, on the other hand, may be more robust in this regard. Moreover, our study was conducted under indoor conditions to limit the interference of external light on text visibility. We hope

[2]While this table provides a designer-friendly overview of the different text writing styles, the binarized standpoint may not provide the full picture of the strengths and weaknesses of each style; as reality is usually more complex than theoretical abstractions. Readers are encouraged to refer to the discussion below for a more holistic picture of the suitability of text writing styles.

that future work can study the effects of these limitations, while also strengthening current findings through replication [3].

8 CONCLUSION

In this work we conducted a systematic investigation of the benefits of dynamic drawing as applied to text, focusing on the underlying factors of font and letter/word motion. In addition, we also examined the impact of these factors with respect to the usage condition, i.e., when used on a desktop while sitting and on an OHMD while walking. Our experimental results showed that although handwritten text has engaging properties, users preferred typeface font which displays the entire word at once, especially for OHMD-based on-the-go learning. We also found evidence indicating that the latter style improves users' recall. Our findings contribute to the existing literature and also provide implications for designing future online learning videos, for both traditional desktop-based and on-the-go OHMD-based video learning situations.

ACKNOWLEDGMENTS

This project is funded by the National Research Foundation, Singapore under its AI Singapore Programme (award number:AISG2-RP-2020-016). We also thank the reviewers for their insightful comments that helped to improve the paper.

REFERENCES

[1] Leon Barnard, Ji Soo Yi, Julie A. Jacko, and Andrew Sears. 2005. An Empirical Comparison of Use-in-Motion Evaluation Scenarios for Mobile Computing Devices. 62, 4 (April 2005), 487–520. https://doi.org/10.1016/j.ijhcs.2004.12.002

[2] Cynthia J. Brame. 2016. Effective Educational Videos: Principles and Guidelines for Maximizing Student Learning from Video Content. *CBE—Life Sciences Education* 15, 4 (2016), es6. https://doi.org/10.1187/cbe.16-03-0125

[3] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 981–992. https://doi.org/10.1145/2858036.2858498

[4] Soon Hau Chua, Simon T. Perrault, Denys J. C. Matthies, and Shengdong Zhao. 2016. Positioning Glass: Investigating Display Positions of Monocular Optical See-Through Head-Mounted Display. In *Proceedings of the Fourth International Symposium on Chinese CHI* (San Jose, USA) *(ChineseCHI2016)*. Association for Computing Machinery, New York, NY, USA, Article 1, 6 pages. https://doi.org/10.1145/2948708.2948713

[5] Joke Coens, Ellen Degryse, Marie-Paul Senecaut, Jorge Cottyn, and Geraldine Clarebout. 2011. Listening to an educational podcast while walking or jogging: Can students really multitask? *International Journal of Mobile and Blended Learning (IJMBL)* 3, 3 (2011), 23–33. https://doi.org/10.4018/jmbl.2011070102

[6] Andrew Cross, Mydhili Bayyapunedi, Edward Cutrell, Anant Agarwal, and William Thies. 2013. TypeRighting: Combining the Benefits of Handwriting and Typeface in Online Educational Videos. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 793–796. https://doi.org/10.1145/2470654.2470766

[7] Adobe After Effects 2021. *Using the Detail-preserving Upscale effect*. Adobe After Effects. Retrieved Decemeber 29, 2021 from https://helpx.adobe.com/after-effects/using/detail-preserving-upscale-effect.html#:~:text=The%20Detail%2Dpreserving%20Upscale%20effect%20is%20capable%20of%20scaling%20up,to%20digital%20cinema%20frame%20sizes.

[8] Yehudit Judy Dori and John Belcher. 2005. How Does Technology-Enabled Active Learning Affect Undergraduate Students' Understanding of Electromagnetism Concepts? *Journal of the Learning Sciences* 14, 2 (2005), 243–279. https://doi.org/10.1207/s15327809jls1402_3

[9] Linnea C. Ehri. 2005. Learning to Read Words: Theory, Findings, and Issues. *Scientific Studies of Reading* 9, 2 (2005), 167–188. https://doi.org/10.1207/s1532799xssr0902_4

[10] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. arXiv:2102.11824 [stat.ME]

[11] James Farley, Evan Risko, and Alan Kingstone. 2013. Everyday attention and lecture retention: the effects of time, fidgeting, and mind wandering. *Frontiers in Psychology* 4 (2013), 619. https://doi.org/10.3389/fpsyg.2013.00619

[12] Logan Fiorella and Richard E Mayer. 2016. Effects of observing the instructor draw diagrams on learning from multimedia messages. *Journal of Educational Psychology* 108, 4 (2016), 528–546. https://doi.org/doi:10.1037/edu00000

[13] Logan Fiorella, Andrew T Stull, Shelbi Kuhlmann, and Richard E Mayer. 2019. Instructor presence in video lectures: The role of dynamic drawings, eye contact, and instructor visibility. *Journal of Educational Psychology* 111, 7 (2019), 1162–1171. https://doi.org/10.1037/edu0000325

[14] Logan Fiorella, Andrew T. Stull, Shelbi Kuhlmann, and Richard E. Mayer. 2020. Fostering generative learning from video lessons: Benefits of instructor-generated drawings and learner-generated explanations. *Journal of Educational Psychology* 112, 5 (2020), 895–906. https://doi.org/10.1037/edu0000408

[15] J. L. Gabbard, J. E. Swan, and D. Hix. 2006. The Effects of Text Drawing Styles, Background Textures, and Natural Lighting on Text Legibility in Outdoor Augmented Reality. *Presence* 15, 1 (2006), 16–32. https://doi.org/10.1162/pres.2006.15.1.16

[16] Joseph L. Gabbard, J. Edward Swan, Deborah Hix, Si-Jung Kim, and Greg Fitch. 2007. Active Text Drawing Styles for Outdoor Augmented Reality: A User-Based Study and Design Implications. In *2007 IEEE Virtual Reality Conference*. 35–42. https://doi.org/10.1109/VR.2007.352461

[17] Debjyoti Ghosh, Pin Sym Foong, Shengdong Zhao, Can Liu, Nuwan Janaka, and Vinitha Erusu. 2020. EYEditor: Towards On-the-Go Heads-Up Text Editing Using Voice and Manual Input. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376173

[18] Tovi Grossman and Ravin Balakrishnan. 2005. The Bubble Cursor: Enhancing Target Acquisition by Dynamic Resizing of the Cursor's Activation Area. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA) *(CHI '05)*. Association for Computing Machinery, New York, NY, USA, 281–290. https://doi.org/10.1145/1054972.1055012

[19] 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* (Atlanta, Georgia, USA) *(L@S '14)*. Association for Computing Machinery, New York, NY, USA, 41–50. https://doi.org/10.1145/2556325.2566239

[20] Ewa Gustafsson, Sara Thomee, Anna Grimby-Ekman, and Mats Hagberg. 2017. Texting on mobile phones and musculoskeletal disorders in young adults: a five-year cohort study. *Applied ergonomics* 58 (2017), 208–214. https://doi.org/10.1016/j.apergo.2016.06.012

[21] Rebecca J. Hess, Jennifer S. Brach, Sara R. Piva, and Jessie M. VanSwearingen. 2010. Walking Skill Can Be Assessed in Older Adults: Validity of the Figure-of-8 Walk Test. *Physical Therapy* 90, 1 (01 2010), 89–99. https://doi.org/10.2522/ptj.20080121

[22] Anne P Hillstrom and Yu-Chin Chai. 2006. Factors that guide or disrupt attentive visual processing. *Computers in human behavior* 22, 4 (2006), 648–656. https://doi.org/10.1016/j.chb.2005.12.003

[23] E. Keuleers and M. Brysbaert. 2010. Wuggy: A multilingual pseudoword generator. *Behavior Research Methods* 42, 3 (2010), 627–633. https://doi.org/10.3758/BRM.42.3.627

[24] Khan Academy 2020. *2018 in review: Annual Report*. Khan Academy. Retrieved September 7, 2021 from https://blog.khanacademy.org/2018-in-review-annual-report/

[25] Khan Academy 2020. *Countercurrent multiplication in the kidney | Renal system physiology | NCLEX-RN | Khan Academy*. Khan Academy. Retrieved August 26, 2021 from https://www.youtube.com/watch?v=Vqce2dtg45U&ab_channel=khanacademymedicine

[26] Khan Academy 2020. *Khan Academy*. Khan Academy. Retrieved August 25, 2021 from http://www.khanacademy.org/

[27] David R. Krathwohl. 2002. A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice* 41, 4 (2002), 212–218. https://doi.org/10.1207/s15430421tip4104_2

[28] Pin-Sung Ku, Yu-Chih Lin, Yi-Hao Peng, and Mike Y. Chen. 2019. PeriText: Utilizing Peripheral Vision for Reading Text on Augmented Reality Smart Glasses. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 630–635. https://doi.org/10.1109/VR.2019.8798065

[29] Zhen Li, Joannes Chan, Joshua Walton, Hrvoje Benko, Daniel Wigdor, and Michael Glueck. 2021. Armstrong: An Empirical Examination of Pointing at Non-Dominant Arm-Anchored UIs in Virtual Reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 123, 14 pages. https://doi.org/10.1145/3411764.3445064

[30] Manuela Macedonia and Thomas R. Knösche. 2011. Body in Mind: How Gestures Empower Foreign Language Learning. *Mind, Brain, and Education* 5, 4 (2011), 196–211. https://doi.org/10.1111/j.1751-228X.2011.01129.x

[31] Yuki Matsuura, Tsutomu Terada, Tomohiro Aoki, Susumu Sonoda, Naoya Isoyama, and Masahiko Tsukamoto. 2019. Readability and Legibility of Fonts Considering Shakiness of Head Mounted Displays *(ISWC '19)*. Association for Computing Machinery, New York, NY, USA, 10 pages. https://doi.org/10.1145/3341163.3347748

[32] Richard E. Mayer. 2002. Multimedia learning. Psychology of Learning and Motivation, Vol. 41. Academic Press, 85 – 139. https://doi.org/10.1016/S0079-7421(02)80005-6

[33] Richard E. Mayer. 2005. *Principles of Multimedia Learning Based on Social Cues : Personalization, Voice, and Image Principles.* Cambridge University Press, 201–212. https://doi.org/10.1017/CBO9780511816819.014

[34] Jennifer Momsen, Tammy Long, Sara Wyse, and Diane Ebert-May. 2010. Just the Facts? Introductory Undergraduate Biology Courses Focus on Low-Level Cognitive Skills. *CBE life sciences education* 9 (12 2010), 435–40. https://doi.org/10.1187/cbe.10-01-0001

[35] M. R. Nassar, J. C. Helmers, and M. J. Frank. 2018. Chunking as a rational strategy for lossy data compression in visual working memory. *Psychological review* 125, 4 (2018), 486–511. https://doi.org/10.1037/rev0000101

[36] Dennis Norris, Anne Cutler, James M. McQueen, and Sally Butterfield. 2006. Phonological and conceptual activation in speech comprehension. *Cognitive Psychology* 53, 2 (2006), 146–193. https://doi.org/10.1016/j.cogpsych.2006.03.001

[37] Fred Paas, Alexander Renkl, and John Sweller. 2003. Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist* 38, 1 (2003), 1–4. https://doi.org/10.1207/S15326985EP3801_1

[38] Ashwin Ram and Shengdong Zhao. 2021. LSVP: Towards Effective On-the-Go Video Learning Using Optical Head-Mounted Displays. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 1, Article 30 (March 2021), 27 pages. https://doi.org/10.1145/3448118

[39] Rufat Rzayev, Paweł W. Woundefinedniak, Tilman Dingler, and Niels Henze. 2018. Reading on Smart Glasses: The Effect of Text Position, Presentation Type and Walking. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3173574.3173619

[40] Clive Thompson. 2011. *How Khan Academy is changing the rules of education.* Retrieved August 25, 2021 from https://www.wired.com/2011/07/ff-khan/

[41] Endel Tulving and Zena Pearlstone. 1966. Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior* 5, 4 (1966), 381 – 391. https://doi.org/10.1016/S0022-5371(66)80048-8

[42] Olive G. Turner. 1930. The Comparative Legibility and Speed of Manuscript and Cursive Handwriting. *The Elementary School Journal* 30, 10 (1930), 780–786. http://www.jstor.org/stable/996624

[43] Adobe After Effects 2021. *Apply Blur and Sharpen in After Effects.* Adobe After Effects. Retrieved Decemeber 29, 2021 from https://helpx.adobe.com/after-effects/using/blur-sharpen-effects.html

[44] Kristin Vadas, Nirmal Patel, Kent Lyons, Thad Starner, and Julie Jacko. 2006. Reading On-the-Go: A Comparison of Audio and Hand-Held Displays. In *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services* (Helsinki, Finland) *(MobileHCI '06)*. Association for Computing Machinery, New York, NY, USA, 219–226. https://doi.org/10.1145/1152215.1152262

[45] Bochkarev Vladimir, Shevlyakova Anna, and Solovyev Valery. 2012. Average word length dynamics as indicator of cultural changes in society. *Social Evolution and History* 14 (2012), 153–175.

[46] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[47] Xing-Dong Yang, Tovi Grossman, Pourang Irani, and George Fitzmaurice. 2011. TouchCuts and TouchZoom: Enhanced Target Selection for Touch Displays Using Finger Proximity Sensing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2585–2594. https://doi.org/10.1145/1978942.1979319

[48] Hui-Shyong Yeo, Xiao-Shen Phang, Steven J. Castellucci, Per Ola Kristensson, and Aaron Quigley. 2017. Investigating Tilt-Based Gesture Keyboard Entry for Single-Handed Text Entry on Large Devices. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4194–4202. https://doi.org/10.1145/3025453.3025520

[49] Shengdong Zhao and Ravin Balakrishnan. 2004. Simple vs. Compound Mark Hierarchical Marking Menus. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology* (Santa Fe, NM, USA) *(UIST '04)*. Association for Computing Machinery, New York, NY, USA, 33–42. https://doi.org/10.1145/1029632.1029639

[50] Yue Zhao, Tarmo Robal, Christoph Lofi, and Claudia Hauff. 2018. Stationary vs. Non-Stationary Mobile Learning in MOOCs. In *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization (UMAP '18)*. Association for Computing Machinery, New York, NY, USA, 299–303. https://doi.org/10.1145/3213586.3225241