Your Eyes on Speed: Using Pupil Dilation to Adaptively Select Speed-Reading Parameters in Virtual Reality

JESSE W. GROOTJEN, LMU Munich and Munich Center for Machine Learning (MCML), Germany
PHILIPP THALHAMMER, LMU Munich, Germany
THOMAS KOSCH, HU Berlin, Germany

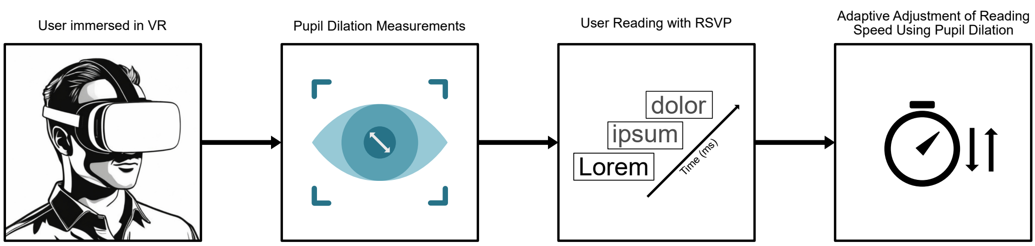


Fig. 1. We utilize pupil dilation measurements in virtual reality to assess the perceived mental workload levels while reading with rapid serial visual presentations. We assess how different presentation speeds affect pupil diameter and present an adaptive reading system that adjusts the presentation speed automatically. Part of this teaser figure was generated with Fooocus.

Rapid Serial Visual Presentation (RSVP) improves the reading speed for optimizing the user's information processing capabilities on Virtual Reality (VR) devices. Yet, the user's RSVP reading performance changes over time while the reading speed remains static. In this paper, we evaluate pupil dilation as a physiological metric to assess the mental workload of readers in real-time. We assess mental workload under different background lighting and RSVP presentation speeds to estimate the optimal color that discriminates the pupil diameter varying RSVP presentation speeds. We discovered that a gray background provides the best contrast for reading at various presentation speeds. Then, we conducted a second study to evaluate the classification accuracy of mental workload for different presentation speeds. We find that pupil dilation relates to mental workload when reading with RSVP. We discuss how pupil dilation can be used to adapt the RSVP speed in future VR applications to optimize information intake.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**.

Additional Key Words and Phrases: Mental Workload, RSVP, Reading, Eye Tracking, Workload-Aware Interfaces, Working Memory

ACM Reference Format:
Jesse W. Grootjen, Philipp Thalhammer, and Thomas Kosch. 2024. Your Eyes on Speed: Using Pupil Dilation to Adaptively Select Speed-Reading Parameters in Virtual Reality. *Proc. ACM Hum.-Comput. Interact.* 8, MHCI, Article 284 (September 2024), 17 pages. https://doi.org/10.1145/3676531

Authors' Contact Information: Jesse W. Grootjen, LMU Munich and Munich Center for Machine Learning (MCML), Munich, Germany, jesse.grootjen@ifi.lmu.de; Philipp Thalhammer, LMU Munich, Munich, Germany, thalhammerphilipp@gmail.com; Thomas Kosch, HU Berlin, Berlin, Germany, thomas.kosch@hu-berlin.de.

ACM 2573-0142/2024/9-ART284
https://doi.org/10.1145/3676531

1 Introduction

The widespread adoption of wearable mobile computing has created a pressing challenge to effectively deliver text on displays with limited screen real estate. Rapid Serial Visual Presentation (RSVP) has emerged as a popular solution to read text on mobile devices, allowing users to fixate on text on the display center. Words are presented sequentially at a user-defined frequency [23], making full-text presentations possible on various wearable devices. RSVP eliminates the need for eye movements and enables faster reading speeds, offering significant time-saving advantages [45]. Consequently, users can read full texts on smartphones [29, 55], smartwatches [18, 67], and augmented [67] or Virtual Reality (VR) [66]. Furthermore, the presentation speed determines how fast a reader can read the displayed text. However, changing the presentation speed impacts the reading demand and text comprehension [45]. Consequently, low presentation speeds can lead to improved text comprehension at the cost of time. In contrast, high presentation speeds decrease text comprehension while increasing time savings. Hence, the optimal presentation speed between processing capabilities and presentation speed can maintain text comprehension while maximizing time savings.

Using RSVP to improve the reading performance on mobile devices was subject to previous research. For example, Dingler et al. [19] enhanced reading efficiency on electronic devices by introducing stimuli that guide users' eye movements to increase reading speed. The study found that such stimuli can effectively increase the reading speed to 150% of the normal rate while maintaining nearly stable comprehension rates, with initial mental load strain decreasing significantly over time. Consequently, RSVP has been successfully used on mobile general-purpose devices [18, 66]. Concerning reading in VR, Gabel et al. [27] compared continuous and discrete user interface panels for displaying and interacting with long texts in VR. The study found no significant differences in reading performance across different text panel variants, but it did identify significant effects on certain aspects of user experience. Although long texts provide an overview of the text, we hypothesize that using RSVP in VR allows text to be displayed closer to the user for increasing readability. However, the challenge of adapting the RSVP presentation speed to the user's mental capacities remains.

Presentation speeds require continuous calibrations to match the user's presentation speed preferences and change the speed with prolonged reading times that may deteriorate the reading performance [45]. Yet, manual presentation speed adjustments disturb the reading flow, asking for adaptive reading flow adjustments and thus increasing the user reading efficiency. Hence, past research recommended presentation speed adaptations, for example, by analyzing mental processes in real-time and adjusting the presentation speed in-situ [1].

To overcome this challenge, we explore pupil dilation to infer mental workload for automatically adjusting RSVP reading speeds. In this context, pupil dilation has been used to measure mental workload in psychology [34] and Human-Computer Interaction (HCI) [44, 57]. Since pupil dilation is susceptible to changing lighting conditions, previous work investigated how pupil dilation can be modeled in light-changing environments [20, 57]. Such predictive models are suitable for VR environments where lighting conditions can be controlled.

Inspired by previous work, we conducted two user studies to understand pupil changes for different background lighting and evaluate the classification accuracy for different reading speeds based on pupil diameter (see Figure 1). The first study reveals that a gray background maximizes the discrimination of pupil diameter when reading with different reading speeds. The second study shows an increasing linear trend for the pupil diameter with increasing reading speeds. We discuss our results and the use of pupil dilation to adapt the reading speed in VR applications.

2 Related Work

The following section iterates previous research on RSVP, how different RSVP parameters influence mental workload, and how mental workload can be sensed using pupil dilation.

2.1 Reading with RSVP

Words are sequentially displayed one at a time in a fixated spatial arrangement when reading with RSVP. This process entails updating the currently displayed word at a predefined frequency, following a single-word-at-a-time approach [30]. RSVP enhances the reading speed by presenting words continuously within the same visual field, thereby minimizing time-consuming eye movements by jumping from word to word while improving the reading concentration [60]. Lengthy text can be displayed legibly on devices with small screen space or in VR, where the text can be brought closer to the user on a word-by-word basis to circumvent the limitations given by the screen resolution and text distance placement.

RSVP enhances the overall reading velocity by presenting words within a consistent area of vision, eliminating the need for time-consuming eye movements when transitioning from one word to another [36, 61]. For example, studies by Rubin and Turano [64] compared reading speeds and comprehension levels between conventional single-page reading and calibrated RSVP techniques. Their research indicates that RSVP can accelerate reading without compromising understanding compared to regular reading. However, this strategy depends on a calibration process where users read aloud, and researchers or users must manually adjust the RSVP speed based on the accuracy of the reading.

Given its single-word presentations, RSVP overcomes space limitations associated with display devices with restricted screen space, allowing for the display of longer texts. Furthermore, single words can be moved closer to the user to enhance readability. Examples of this are the successful implementation of RSVP on devices such as mobile phones [46, 70], smartwatches [17], and mobile augmented reality devices [67]. Yet, the requirements for RSVP design may vary depending on the device and the user's mental capacity at the time of use [37]. Users may favor slower word presentation speeds in the later hours when they are already mentally exhausted during prolonged reading sessions.

2.2 RSVP Parameter Design

Building upon the RSVP concept, the specific designs of RSVP parameters enhance the overall reading experience. As a method of presenting single words or short phrases sequentially, RSVP eliminates the need for eye movements and allows for precise control over the rate at which readers access textual information. Previous research has focused on identifying key factors that optimize reading performance in RSVP, namely text alignment and presentation speed, were explored in previous work [45]. In their study, the authors investigate different RSVP parameters and their impact on the user's perceived workload by using subjective (e.g., the NASA-TLX questionnaire [32, 33]) and objective measures (e.g., Electroencephalography (EEG) [24, 25]).

Text alignment, often called the centering of text, influences the reader's fixation location and how words or phrases are presented. An optimal text alignment ensures that the displayed content is fixated without requiring unnecessary eye movements. Previous studies have investigated different ways of implementing this. Some commercial RSVP implementations, such as Spritz Inc.[1], employ an Optimal Recognition Position (ORP) to reduce the saccades and minimize eye movements. However, prolonged use of ORPs can lead to reduced parafoveal processing, increased perceived workload (i.e., working memory [3]), and a higher frequency of blinks, indicating visual fatigue

[1] https://spritz.com/ – last accessed 2024-07-29

[5]. Dingler et al. [19] explored alternative ORP representations, such as underlining instead of coloring them, and assessed their impact on perceived workload and text comprehension. Although no significant differences were found between underlined and colored ORPs regarding workload and text comprehension, participants reported quicker adaptation to colored ORPs. Following this, Kosch et al. [45] explored the combination of ORP and color and found no significant difference in perceived workload and alpha and theta power in their EEG signal. Their results show that the ORP is not a decisive factor influencing the perceived mental workload.

Another parameter in RSVP design is presentation speed, which determines the rate at which words are displayed. A faster presentation speed can reduce the time required to read a document. However, it comes at the cost of reduced cognitive processing time for each word. Previous work has shown a negative correlation between text comprehension and presentation speed [45, 65]. While text alignment showed no significant difference in various studies looking at perceived workload and measured mental workload, presentation speed does have a substantial effect on perceived workload and measured mental workload.

2.3 Sensing Mental Workload By Pupil Dilation

Estimating mental workload through eye tracking is a valuable and insightful approach in various fields, such as HCI, aviation, and cognitive psychology. Using eye tracking to monitor a person's eye movements and gaze patterns, researchers can gain insights into their cognitive processes and the demands on their visual attention. This method allows for evaluating mental workload in real-time, making it particularly advantageous for applications where task performance, safety, or user experience is of concern. Eye tracking can provide metrics, such as fixation duration, saccade frequency, and pupillary response, to gauge mental workload [4, 15, 57]. Pupil dilation serves as another indicator that studies have shown to be consistently associated with task difficulty, a concept established by the work of Hess and Polt [34]. Consequently, pupil dilation has been proposed as a near-real-time metric for measuring mental workload, effectively using mental workload in adaptive mixed reality interfaces [50]. In high mental workload tasks, the mean and standard deviation of the pupil diameter increase in response to the heightened visual processing [6, 54]. While more advanced pupil-based metrics have been proposed to measure mental workload, such as the index of pupillary activity [7, 20], it is crucial to acknowledge that environmental factors, such as scene colors, brightness, and movement, can influence pupil dilation, which could complicate its interpretation [35, 56].

2.4 Mental Workload and Task Performance in VR

Mental workload in VR mirrors the role of working memory in cognitive tasks. In VR, visual distractions and mental demands can significantly affect task performance, leading to increased response times, discomfort, physiological arousal, and cognitive overload [49, 52, 58]. Conversely, an excessively automated and low-fidelity VR environment causes user disengagement. Thus, HCI researchers must balance cognitive engagement and task demand [63]. As tasks become more challenging and success less certain, user effort decreases, and perceived workload rises, potentially resulting in impaired performance, increased perceived workload, and reduced engagement.

Understanding the physiological aspect of mental workload is important to understand the impact on user experience in VR. Electrodermal Activity (EDA) recordings can offer insights into task engagement, with higher skin conductance levels associated with reduced engagement and heightened stress demanding tasks [21]. Another example is the work from Chiossi et al. [11], where they created a physiologically adaptive system that improves task performance and perceived workload. Additionally, the use of EEG [10, 12, 24] has been used for adapting user interfaces based on mental workload. Yet, these methods require sensors that make skin contact with the user and

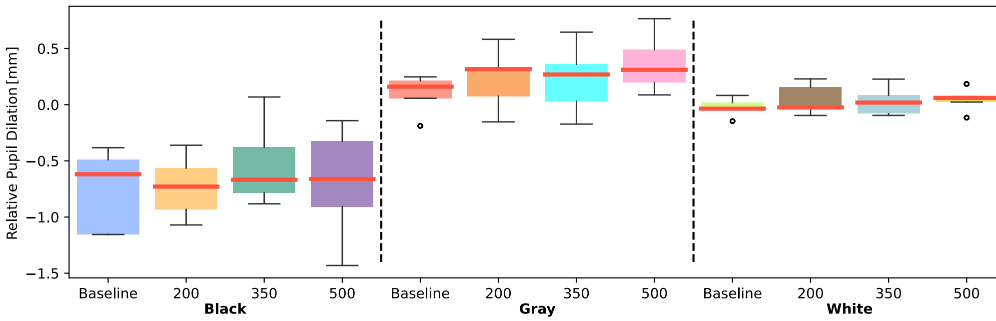


Fig. 2. Relative pupil dilation values of the first experiment for each condition and background color. We established a resting baseline prior before the reading conditions of the first experiment started to obtain the relative pupil diameter for the subsequent conditions (i.e., static baseline reading, 200 WPM, 350 WPM, and 500 WPM).

impair usability. Therefore, previous research investigated contactless methods to infer mental workload.

Consequently, pupil dilation has been exploited as a metric for mental workload. Previous work showed that the pupil can be robustly sensed using eye tracking [26] to infer the degree of perceived mental workload [41]. VR glasses especially benefit from using pupil dilation as a metric for mental workload in VR [51, 62]. We hypothesize that pupil dilation is a metric to automatically adjust the RSVP presentation speed based on the measured mental workload in real-time.

2.5 Summary

Previous work showed that RSVP is suitable for reading text in limited spaces. However, different presentation speeds elicit different levels of mental workload [45, 65], especially when reading for extended periods, forcing users to set their preferred presentation speed constantly to maximize the information processing throughput. At the same time, mental workload can be classified using pupil dilation and eye movement measures [7, 20, 34, 41, 42] by integrating eye tracking into VR headsets [51, 62]. Motivated by improving reading in VR environments using adaptive RSVP, we investigate how pupil dilation can indicate mental workload to adjust the presentation speed. We conducted two user studies in VR: (1) we investigated how different background lightings affect pupil diameter under different presentation speeds in a prestudy. (2) In a second study, we evaluate how fine-grained presentation speed levels influence pupil diameter using suitable background lighting from the prestudy. Finally, we discuss how our findings can be used to build adaptive RSVP applications that set the presentation speed automatically based on the perceived workload.

3 General Experiment Setup

We cover the description for the following two user studies, in which the experimental apparatus and materials were consistent throughout both experiments.

3.1 Apparatus and Materials

All studies were conducted using the HTC Vive Pro Eye Headset[2], which features 1440×1600 pixels per eye and tracks eye movement at 120 Hz. The interpupillary distance was calibrated for each participant. The VR reading tasks were implemented in Unity Version 2021.3.25f1 and executed on

[2]https://www.vive.com/us/product/vive-pro-eye/specs – last accessed 2024-07-29

a PC with Windows 10, an Intel Core i7 10700k at 3.80 GHz, 32 GB RAM, and an Nvidia Geforce 3070Ti. The text was presented in font size 36 and placed 60 centimeters away from the viewer in the *Space Mono Regular* font. All texts, the code of the prototype, and analysis scripts are made publicly available[3].

3.2 Procedure

Upon arrival, the participants were greeted with a verbal outline of the study regarding the procedure, rationale, and study goals. Afterward, the participants completed an informed consent form and provided their demographic data. Participants with visual impairments were excluded from the study. We then made the participants familiar with RSVP using a generic text with the Spritz-App[4] on a mobile phone. After this introduction, the participants were advised to wear the VR headset. Then, we calibrated the eye tracker before testing the different conditions for the two experiments. In total, we recruited 26 participants (11 female, 15 male), where the participant's age ranged between 17 and 59 (M = 28.2, SD = 11.4). All participants received a compensation of €10. All participants reported having minimal to no experience with RSVP, and participants self-reported an average of experience with VR with 2.5 (SD = 1.53)[5].

4 Experiment I: Evaluating Background Brightness

In the first experiment, we evaluated how different background levels affect pupil dilation under different presentation speeds. We explore how the background brightness, in combination with the presentation speed, influences the pupil diameter.

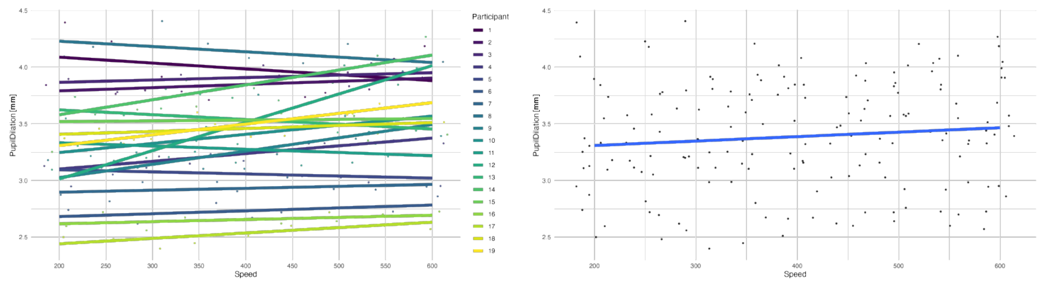
4.1 Study Design

The study utilized a repeated measures design, incorporating two independent variables: presentation speed and background brightness. The presentation speeds were 200 WPM, 350 WPM, and 500 WPM based on previous research [45]. For background brightness, three levels were used: black (#000000) with white text, gray (#7F7F7F) with black text, and white (#FFFFFF) with black text, resulting in a total of nine RSVP reading trials per participant. The three background brightness levels cover the brightness between the brightest level (i.e., white), the darkest level (i.e., dark), and an in-between level (i.e., gray). Additionally, three baseline tasks were included, where participants read text presented in one block at each background brightness level to establish baseline measurements for pupil dilation and reading speed. Each participant thus participated in twelve reading trials. We showed a blank screen for 15 seconds before each trial to avoid bleed-over effects from previous conditions and give the pupil time to adjust to the light levels. The texts for the reading conditions were sourced from the work of Hahn et al. [31], each about 130 words long and written at the level of a sixth grade in German. We evaluate the pupil dilation as the only measure in experiment I.

4.2 Participants

Seven participants (three male and four female) participated in the experiment. They were recruited via our university mailing list and snowball system. The participants' ages ranged between 24 and 59 (M = 36.6, SD = 15.8). All participants received a compensation of €10. All participants reported having minimal to no experience with RSVP, and participants self-reported an average of experience with VR with 2.57 (SD = 1.4).

[3] https://github.com/mimuc/RSVP
[4] https://spritz.com – last accessed 2024-07-29
[5] Reported on a scale from 1 to 5, with 1 being no experience, five being used at least once a week.



(a) Linear Relation per Participant.

(b) Mean Linear Relation.

Fig. 3. Pupil Dilation (in millimeter) plotted against the speed in RSVP (words per minute). **(a)** The lines represent a linear model plotted through the points per participant. **(b)** The line represents a linear model plotted through the data points of all participants. In both (**a**) and (**b**), observe a linear relation between the presentation speed and pupil dilation. However, the linear relation does not always follow a positive trajectory.

4.3 Data Processing and Results

We statistically analyze the participants' relative pupil dilation. We measure the mean pupil dilation during a prior resting condition to determine the individual pupil dilation. Then, the relative pupil dilation is calculated by subtracting the measured mean pupil dilation values from the mean resting baseline value. We conducted a Shapiro-Wilk test to test our measures for normality. We adopted an alpha level of .05 for statistical significance. Post hoc tests were performed between the conditions. Post hoc pair-wise comparisons to baseline measures derived from static reading were performed when applicable. We investigated if the relative pupil diameter for the three background colors originated from a normal distribution. A Shapiro-Wilk test showed no deviation of normality for all three background colors, $p > .05$. We applied repeated measures ANOVA separately on the relative pupil diameter measures. We did this separately for each background color to avoid interaction effects resulting from large changes in pupil diameter through the background color instead of the presentation speed. Figure 2 shows an overview of pupil dilation in the first experiment.

The black background did not show a significant main effect between the presentation speed conditions, $F(1, 20) = 1.40, p = .25, \eta_p^2 = 0.07$. However, the gray background showed a significant main effect for the presentation speeds, $F(1, 20) = 15.41, p < .001, \eta_p^2 = 0.44$. Bonferroni-corrected pairwise t-tests between the conditions did not show a significant effect between all conditions, $p > .05$. Finally, the white background showed a significant main effect for the presentation speeds, $F(1, 20) = 4.92, p = .04, \eta_p^2 = 0.20$. Again, Bonferroni-corrected pairwise t-tests between the conditions did not show a significant effect between all conditions, $p > .05$.

4.4 Discussion

We found that background brightness did not affect pupil dilation in static VR conditions. As shown in Figure 2, the average relative pupil dilation displayed slight variations across different presentation speeds. Nonetheless, the gray background exhibited a linear increase in pupil dilation. Therefore, we used the same gray background for experiment II to examine the impact of various presentation speeds on pupil dilation. The results indicate that the background brightness does not significantly influence pupil dilation under static conditions in a VR environment. This finding suggests that other factors might play a more prominent role in affecting pupil size when the visual scene is not changing. The marginal differences observed in average relative pupil dilation across

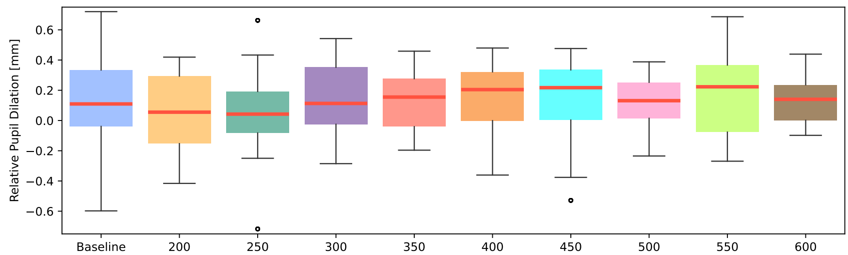


Fig. 4. Relative Pupil Dilation values for each presentation speed. The participants conducted a resting baseline before starting with the reading conditions. Afterward, the participants read static text (i.e., baseline) or the RSVP conditions in a counterbalanced order.

different presentation speeds suggest that the speed at which visual stimuli are presented might subtly impact pupil response and detection. For example, When a display is bright, the scattering of light forms a diffuse layer of retinal illumination, which lowers the perceived contrast of the image and consequently hinders detection performance while overstimulating the eyes [53]. Interestingly, the gray background showed a linear trend of increasing pupil dilation with increasing presentation speeds. This linear relationship implies that pupil dilation increases consistently when a gray background is used as the visual presentation progresses. This could be due to a gray background's neutral and non-distracting nature, allowing the eyes to respond more directly to the visual stimuli without interference from background brightness changes. Thus, we use the same gray background for experiment II.

5 Experiment II: Main Study

The goal of experiment II is to evaluate changes in pupil dilation under different presentation speeds in VR.

5.1 Study Design

Again, we use the presentation speed as the only independent variable, employing nine speeds ranging from 200 WPM to 600 WPM in 50 WPM increments to allow for more gradual data collection. Ten conditions were tested per participant: one static reading condition for baseline measurements of each participant's reading speed and nine RSVP conditions at the specified presentation speeds. Based on the results of experiment I, all conditions combined a gray background with black text. The order of conditions was counterbalanced, and texts were randomly assigned for each condition.

We include the pupil dilation as a dependent variable. Furthermore, we use the NASA-TLX [32, 33] to operationalize the task load and text questionnaires to quantify the reading comprehension after each condition. Each text stated three text comprehension questions, where we counted the number of correct answers for each text.

5.2 Apparatus and Materials

The texts for this experiment were sourced from a collection of standardized speed reading passages by Quinn et al. [59], which were translated into German using ChatGPT. This was followed by the experimenters' evaluation of the texts' plausibility. Each text was about 500 to 550 words long and was written in a low (German) language level. Instead of one right/wrong question, each text featured three single-choice questions with four possible answers each to avoid the participants only guessing the answer and still having a 50% success rate. For the static reading condition (i.e., baseline), a text by Hahn et al. [31] was used again, supplemented with three newly created single-choice questions.

5.3 Procedure

The procedure for experiment II is similar as described in Section 3.2. Participants conducted a total of ten conditions see Section 5.1, starting with the static reading task for baseline measurement, followed by the nine RSVP conditions.

5.4 Participants

Nineteen participants (seven female and twelve male), aged between 17 and 50 (M = 25.1, SD = 7.8), were recruited for the study. All participants were recruited via the mailing list for the study and snowball system. All participants received a compensation of €10 for participating. All participants reported having minimal to no experience with RSVP and self-reported a mean of experience with VR of 2.47 (SD = 1.6). All participants had to either have normal or corrected to normal vision with contacts, as no glasses could be worn under the VR headset to prevent problems with eye-tracking. Those with corrected to normal vision wore contacts. Additionally, all participants had to be native German speakers, so there would be no effect on the measurements because of varying language skills.

5.5 Data Processing and Results

We conduct a statistical analysis of the participants' relative pupil dilation. First, we measure the average pupil dilation during an initial resting state to establish a baseline for each individual. Then, we calculate the relative pupil dilation by subtracting the observed dilation values from the mean resting baseline value. The RSVP test conditions gave each participant three measurements (pupil dilation, task load, and text comprehension). We conducted a Shapiro-Wilk test to test the measure for normality. We adopted an alpha level of .05 for statistical significance. Post hoc tests were performed between the conditions. Post hoc pair-wise comparisons to baseline measures derived from static reading were performed when applicable.

5.5.1 Pupil Dilation. A Shapiro-Wilk test revealed a deviation from normality in our measures, $p < .05$. A Friedman test did not reveal a significant main effect, $\chi^2(9, N = 19) = 10.2, p = .332$. The effect size was calculated using Kendall's W, which was .06, indicating a small effect size. Figure 4 shows the relative pupil diameter differences.

5.5.2 Task Load. A Shapiro-Wilk test revealed a deviation from normality in our measures, $p < .05$. A Friedman test revealed a significant main effect, $\chi^2(9, N = 19) = 100.00, p < .001$. Kendall's W denotes a value of .59, indicating a large effect size. Pairwise Bonferroni-corrected Wilcoxon signed rank post hoc tests revealed significant effects between the subjective task load scores (see Table 1 and Figure 5a).

5.5.3 Text Comprehension. A Shapiro-Wilk test revealed a deviation from normality in our measures, $p < .05$. A Friedman test revealed a significant main effect, $\chi^2(9, N = 19) = 17.8, p = .038$. Kendall's W denotes a value of .10, indicating a small effect size. However, pairwise Bonferroni-corrected Wilcoxon signed-rank post hoc tests did not reveal a significant effect between the pairs. Figure 5b shows the averaged comprehension scores.

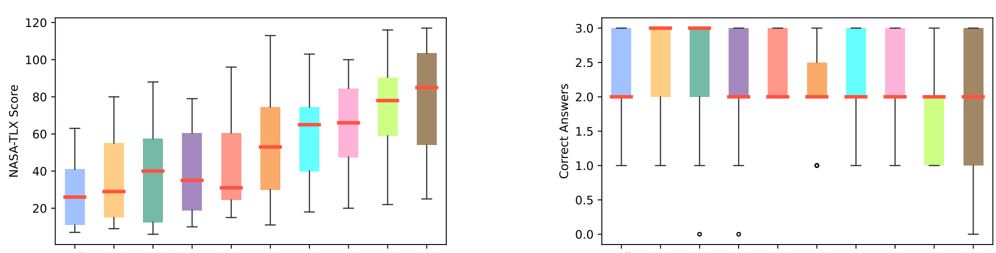
5.5.4 Bayesian Analysis of Pupil Dilation. We adopt a Bayesian approach for data analysis of pupil dilation, specifically employing Bayesian linear mixed models (BLMM). This approach recently gained traction [38, 47, 68] and offers several advantages over inference statistics. One of these

Table 1. p-values from the Bonferroni-corrected pairwise post hoc Wilcoxon signed-rank tests for the raw NASA-TLX scores. The effect size r denotes the rank-biserial correlation.

Pair		**p-value**	r
600 WPM	350 WPM	$p < .001$	.571
550 WPM	350 WPM	$p < .001$	.538
	300 WPM	$p < .001$	.602
	250 WPM	$p < .001$	.583
	200 WPM	$p < .001$	.638
500 WPM	300 WPM	$p < .001$	.462
	250 WPM	$p < .001$	.450
	200 WPM	$p < .001$	.524
450 WPM	200 WPM	$p = .009$	.469
Baseline	600 WPM	$p < .001$	.711
	550 WPM	$p < .001$	.744
	500 WPM	$p < .001$	.657
	450 WPM	$p = .002$	.588
	400 WPM	$p = .019$	.502

Table 2. Bayesian statistics results of the priors for pupil dilation. All results are contrasted against the whole dataset. We report the High-Density Interval (HDI) with a 95% credible interval.

	Pupil Dilation		
Speed	p_b	*Med.*	$HDI_{95\%}$
200	<.001	0.09	[0.00, 0.27]
250	<.001	0.09	[0.00, 0.27]
300	<.001	0.10	[0.00, 0.29]
350	<.001	0.14	[0.00, 0.33]
400	<.001	0.14	[0.00, 0.34]
450	<.001	0.12	[0.00, 0.31]
500	<.001	0.17	[0.00, 0.37]
550	<.001	0.16	[0.00, 0.36]
600	<.001	0.21	[0.00, 0.42]



(a) (b)

Fig. 5. **(a)** Perceived task load for all participants. **(b)** Number of correct answers to the comprehension questionnaires provided by the participants.

advantages, as presented by Kay et al. [39], is the incorporation of prior knowledge from eye-tracking data. Additionally, Bayesian statistics allows for effect estimation in small sample sizes. It will enable readers to evaluate effect size, including those close to zero, rather than determining the presence or absence of effects. Consequently, we utilize Bayesian parameter estimation to estimate effect sizes and quantify the uncertainty surrounding those estimations by leveraging the information in our data and the applied prior knowledge. For all our models, we use the package, *brms* to compute 4 Hamilton-Monte-Carlo chains with 40.000 iterations each and 10% warm-up samples. All Rubin-Gelman [28] statistics are well below 1.1 for effective sample size.

None of the explored different weakly informative priors affected the statistical inference. As a result, we chose priors to resemble only weakly informative priors when standardizing with a prior on a normal distribution of the data ($M = 4, SD = 2$) without allowing for negative numbers ($lb = 0$). Additionally, we accounted for the potential variability across participants by incorporating a random factor on the participant parameter. This approach acknowledges that different participants may exhibit varying characteristics and baselines and allows for more nuanced modeling. By explicitly modeling the participant-specific effect, we capture the heterogeneity and better account for the underlying structure of the data.

Effects were considered meaningful when a particularly low probability (p_b <= 2.5) of the effect being zero or the opposite. We calculated p_b through the relative proportion of posterior samples being zero or opposite to the median. This metric has similar properties to the classical p-value and is an accepted substitution cf. [40, 48, 68]. Still, it quantifies the proportion of probability that the effect is zero or the opposite, given the data observed. Note that this is the reverse of the classical approach to inferential statistics, where one measures data probability given the test statistic's null hypothesis. In addition to the median of the parameter, we calculated the High-Density Interval (HDI) at 95% of the posterior distribution for all parameters, which indicates the possible range of effects given the data alongside the median of the respective parameter. Simple mean comparisons were made on standardized outcome variables. Therefore, all $\tilde{b}$ represent an effect size in standard deviations from the mean (corresponding to Cohen's d for simple effects of categorical predictors with two levels). We found that all variations of presentation speed had a distinguishable effect on pupil dilation (see Table 2).

6 Discussion

We conducted a user study using eye tracking to investigate the impact of presentation speeds on text comprehension, subjective task load, and pupil dilation. The pilot study investigated the influence of background color on pupil dilation (see Section 4), indicating a linear trend with a gray background color. Subsequently, our main study (see Section 5) revealed that increased presentation speeds significantly increased subjective workload and reduced text comprehension.

6.1 Mental Workload and RSVP

Our results show the relationship between presentation speeds and RSVP task performance. While speeds exceeding 350 WPM demonstrated significant gains in reading speed compared to static reading, they incurred a substantial cost in subjective workload. These performance outcomes align with existing literature [45, 65], and our eye-tracking data further strengthens the impact on perceived task load as evidenced by our task load measures. Our findings highlight the complexity of analyzing physiological and subjective measures in research. The task load showed a significant effect with increasing presentation speeds. Thus, the presentation speed has successfully manipulated mental workload. However, text comprehension and pupil dilation did not show a significant difference. Adopting Bayesian statistics provided a robust framework for handling pupil dilation data. Bayesian statistics considers individual differences in pupil diameter data and their interpersonal change towards different workload types, lighting conditions, or age [13, 69]. Thus, we recommend collecting an individual baseline to obtain a calibration when using pupil dilation for assessing VR RSVP reading interfaces.

6.2 Real-Time Assessment for RSVP Parameter Selection

Pupil dilation emerges as a valuable real-time indicator for evaluating RSVP reader designs. Its practicality, particularly in VR systems where eye-tracking already is present, makes it an efficient choice compared to introducing additional modalities such as electrodermal activity [9, 11] or electroencephalography [25]. Our data supports the concept of perceptual workload-aware computer interfaces, demonstrating their ability to sense changes in reading speed based on pupil dilation.

The mean pupil dilation is a metric for measuring perceptual workload in RSVP. Our findings indicate an increase in pupil dilation with increasing presentation speeds, consistent with the motivational intensity model [63] and prior research on using pupil dilation [7] for adaptive interfaces [50]. We noted that pupil dilation declines when surpassing a threshold of approximately 500 WPM, contradicting our findings of subjective task workload and text comprehension questionnaires. Our findings are consistent with previous research [45], where the authors hypothesize that participants reduce reading activity during higher presentation speeds.

6.3 Using Pupil Dilation in Interactive Systems

Using pupil dilation in interactive systems was envisioned in previous research. Yet, further research is necessary before pupil dilation becomes viable in interactive systems. For example, the pupil diameter can change through other stimuli besides mental workload. In this context, one major drawback is its sensitivity to various external factors, such as changes in lighting conditions [57] or the user's emotional states [2], which can independently influence pupil size. Additionally, individual differences in baseline pupil size and reactivity can vary widely, making it challenging to standardize measurements across different users [13, 69]. Pupil dilation also tends to reflect general arousal rather than specific mental workload, complicating the interpretation of data. Moreover, continuous monitoring of pupil dilation requires specialized equipment, which may not be practical or comfortable for all users in real-world settings. These limitations require further research in this domain to establish reliable interaction in the future.

6.4 Study Limitations

Several design limitations in our study warrant consideration. The manipulation of presentation speeds was confined to nine levels, with a maximum of 600 WPM, leaving the effects on pupil dilation, subjective task load, and text comprehension unexplored for speeds beyond this range. Additionally, our sample comprised individuals unfamiliar with RSVP, limiting the generalizability of our results to non-trained users of RSVP.

6.5 Future Work and Outlook

Our findings provide a proof of concept for a potential adaptive system based on pupil dilation measurements to regulate presentation speed. However, challenges such as variations in reading speed, mobile use, interruptions, outdoor scenarios, and different viewing postures must be addressed before commercialization. We will begin with addressing individual differences in pupil dilation. We will achieve this by investigating how interactive VR systems can be calibrated to account for individual pupil diameter differences. Combining pupil dilation measurements with other eye gaze metrics, such as smooth pursuit [42] and saccadic [16] eye movements, or other psychophysiological metrics including heart rate variability [14] or electrodermal activity [43]. Furthermore, we will investigate how VR interface elements with heterogeneous brightness levels can be detected and compensated for the measured pupil values. We will combine these findings into a functional adaptive RSVP reading system. Finally, future research must consider individuals with visual impairments, ensuring that adaptive models for RSVP presentation speeds are inclusive [8, 22].

7 Conclusion

This study uses eye-tracking technology to investigate the impact of presentation speeds in Rapid Serial Visual Presentation (RSVP). We investigate in the first experiment the influence of pupil diameter variation under different background lighting when reading with different presentation speeds in Virtual Reality (VR). Our findings show an increasing linear trend for increasing presentation speeds when using gray background lighting. Thus, our second experiment evaluates pupil dilation and the subjectively perceived task load and text comprehension for different RSVP presentation speeds. While we found that the task load was successfully manipulated with increasing RSVP presentation speeds, we did not find a significant effect on the pupil diameter. Yet, Bayesian modeling shows that pupil diameter is distinguishable under specific circumstances, for example, for presentation speeds ranging from 200 WPM to 450 WPM. Designing adaptive reading interfaces that utilize pupil dilation to change the RSVP presentation speed in real time remains challenging. The pupil diameter is susceptible to external stimuli, such as emotional arousal or lighting conditions. Furthermore, factors related to reading, such as varying text difficulty, also impose a challenge. We are confident that both experiments extend the understanding of which lighting conditions maximize the discrimination of mental workload in pupillary data and how this knowledge can be used to investigate the impact of interfaces on mental workload. The Unity code, dataset, and analysis with all associated measures are publicly available for reproduction and use in future studies[6].

References

[1] Lena M Andreessen, Peter Gerjets, Detmar Meurers, and Thorsten O Zander. 2021. Toward neuroadaptive support technologies for improving digital reading: a passive BCI-based assessment of mental workload imposed by text difficulty and presentation speed during reading. *User Modeling and User-Adapted Interaction* 31 (2021), 75–104. https://doi.org/10.1007/s11257-020-09273-5

[2] Areej Babiker, Ibrahima Faye, Kristin Prehn, and Aamir Malik. 2015. Machine Learning to Differentiate Between Positive and Negative Emotions Using Pupil Diameter. *Frontiers in Psychology* 6 (2015). https://doi.org/10.3389/fpsyg.2015.01921

[3] Alan Baddeley. 1992. Working Memory. *Science* 255, 5044 (1992), 556–559. https://doi.org/10.1126/science.1736359 arXiv:https://www.science.org/doi/pdf/10.1126/science.1736359

[4] Roman Bednarik, Piotr Bartczak, Hana Vrzakova, Jani Koskinen, Antti-Pekka Elomaa, Antti Huotarinen, David Gil de Gómez Pérez, and Mikael von und zu Fraunberg. 2018. Pupil Size as an Indicator of Visual-Motor Workload and Expertise in Microsurgical Training Tasks. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications* (Warsaw, Poland) *(ETRA '18)*. Association for Computing Machinery, New York, NY, USA, Article 60, 5 pages. https://doi.org/10.1145/3204493.3204577

[5] Simone Benedetto, Andrea Carbone, Marco Pedrotti, Kevin Le Fevre, Linda Amel Yahia Bey, and Thierry Baccino. 2015. Rapid serial visual presentation in reading: The case of Spritz. *Computers in Human Behavior* 45 (2015), 352–358. https://doi.org/10.1016/j.chb.2014.12.043

[6] Anne-Marie Brouwer, Maarten A. Hogervorst, Bob Oudejans, Anthony J. Ries, and Jonathan Touryan. 2017. EEG and Eye Tracking Signatures of Target Encoding during Structured Visual Search. *Frontiers in Human Neuroscience* 11 (2017). https://doi.org/10.3389/fnhum.2017.00264

[7] Siyuan Chen, Julien Epps, Natalie Ruiz, and Fang Chen. 2011. Eye Activity as a Measure of Human Mental Effort in HCI. In *Proceedings of the 16th International Conference on Intelligent User Interfaces* (Palo Alto, CA, USA) *(IUI '11)*. Association for Computing Machinery, New York, NY, USA, 315–318. https://doi.org/10.1145/1943403.1943454

[8] Allen M.Y. Cheong, Gordon E. Legge, Mary G. Lawrence, Sing-Hang Cheung, and Mary A. Ruff. 2007. Relationship between slow visual processing and reading speed in people with macular degeneration. *Vision Research* 47, 23 (2007), 2943–2955. https://doi.org/10.1016/j.visres.2007.07.010

[9] Francesco Chiossi, Thomas Kosch, Luca Menghini, Steeven Villa, and Sven Mayer. 2023. SensCon: Embedding Physiological Sensing into Virtual Reality Controllers. *Proc. ACM Hum.-Comput. Interact.* 7, MHCI, Article 223 (sep 2023), 32 pages. https://doi.org/10.1145/3604270

[6] https://github.com/mimuc/RSVP.

[10] Francesco Chiossi, Changkun Ou, Carolina Gerhardt, Felix Putze, and Sven Mayer. 2023. Designing and evaluating an adaptive virtual reality system using eeg frequencies to balance internal and external attention states. *arXiv preprint arXiv:2311.10447* (2023).
[11] Francesco Chiossi, Yagiz Turgut, Robin Welsch, and Sven Mayer. 2023. Adapting Visual Complexity Based on Electrodermal Activity Improves Working Memory Performance in Virtual Reality. *Proc. ACM Hum.-Comput. Interact.* 7, MHCI, Article 196 (sep 2023), 26 pages. https://doi.org/10.1145/3604243
[12] Emily B.J. Coffey, Anne-Marie Brouwer, and Jan B.F. van Erp. 2012. Measuring workload using a combination of electroencephalography and near infrared spectroscopy. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 56, 1 (2012), 1822–1826. https://doi.org/10.1177/1071181312561367
[13] Joseph T. Coyne, Cyrus Foroughi, and Ciara Sibley. 2017. Pupil Diameter and Performance in a Supervisory Control Task: A Measure of Effort or Individual Differences? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 61, 1 (2017), 865–869. https://doi.org/10.1177/1541931213601689
[14] Stéphane Delliaux, Alexis Delaforge, Jean-Claude Deharo, and Guillaume Chaumet. 2019. Mental Workload Alters Heart Rate Variability, Lowering Non-linear Dynamics. *Frontiers in Physiology* 10 (2019). https://doi.org/10.3389/fphys.2019.00565
[15] Shannon Patricia Devlin, Jake Ryan Flynn, and Sara Lu Riggs. 2020. How Shared Visual Attention Patterns of Pairs Unfold Over Time When Workload Changes. In *ACM Symposium on Eye Tracking Research and Applications* (Stuttgart, Germany) *(ETRA '20 Short Papers)*. Association for Computing Machinery, New York, NY, USA, Article 41, 5 pages. https://doi.org/10.1145/3379156.3391339
[16] Leandro L Di Stasi, Rebekka Renner, Peggy Staehr, Jens R Helmert, Boris M Velichkovsky, José J Cañas, Andrés Catena, and Sebastian Pannasch. 2010. Saccadic peak velocity sensitivity to variations in mental workload. *Aviation, space, and environmental medicine* 81, 4 (2010), 413–417. https://doi.org/10.3357/ASEM.2579.2010
[17] Tilman Dingler, Rufat Rzayev, Valentin Schwind, and Niels Henze. 2016. RSVP on the Go: Implicit Reading Support on Smart Watches through Eye Tracking. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers* (Heidelberg, Germany) *(ISWC '16)*. Association for Computing Machinery, New York, NY, USA, 116–119. https://doi.org/10.1145/2971763.2971794
[18] Tilman Dingler, Rufat Rzayev, Alireza Sahami Shirazi, and Niels Henze. 2018. Designing Consistent Gestures Across Device Types: Eliciting RSVP Controls for Phone, Watch, and Glasses. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173993
[19] Tilman Dingler, Alireza Sahami Shirazi, Kai Kunze, and Albrecht Schmidt. 2015. Assessment of Stimuli for Supporting Speed Reading on Electronic Devices. In *Proceedings of the 6th Augmented Human International Conference* (Singapore, Singapore) *(AH '15)*. Association for Computing Machinery, New York, NY, USA, 117–124. https://doi.org/10.1145/2735711.2735796
[20] Andrew T. Duchowski, Krzysztof Krejtz, Izabela Krejtz, Cezary Biele, Anna Niedzielska, Peter Kiefer, Martin Raubal, and Ioannis Giannopoulos. 2018. The Index of Pupillary Activity: Measuring Cognitive Load Vis-à-Vis Task Difficulty with Pupil Oscillation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173856
[21] Stephen H. Fairclough and Louise Venables. 2006. Prediction of subjective states from psychophysiology: A multivariate approach. *Biological Psychology* 71, 1 (2006), 100–110. https://doi.org/10.1016/j.biopsycho.2005.03.007
[22] Elisabeth M Fine, Eli Peli, and Adam Reeves. 1997. Simulated cataract does not reduce the benefit of RSVP. *Vision Research* 37, 18 (1997), 2639–2647. https://doi.org/10.1016/S0042-6989(97)00051-5
[23] Kenneth I Forster. 1970. Visual perception of rapidly presented word sequences of varying complexity. *Perception & psychophysics* 8 (1970), 215–221. https://doi.org/10.3758/BF03210208
[24] Jérémy Frey, Maxime Daniel, Julien Castet, Martin Hachet, and Fabien Lotte. 2016. Framework for Electroencephalography-based Evaluation of User Experience. 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, 2283–2294. https://doi.org/10.1145/2858036.2858525
[25] Jérémy Frey, Christian Mühl, Fabien Lotte, and Martin Hachet. 2013. Review of the Use of Electroencephalography as an Evaluation Method for Human-Computer Interaction. arXiv:1311.2222 [cs.HC]
[26] Wolfgang Fuhl, Thiago Santini, Gjergji Kasneci, and Enkelejda Kasneci. 2016. PupilNet: Convolutional Neural Networks for Robust Pupil Detection. *CoRR* abs/1601.04902 (2016). arXiv:1601.04902 http://arxiv.org/abs/1601.04902
[27] Jenny Gabel, Melanie Ludwig, and Frank Steinicke. 2023. Immersive Reading: Comparison of Performance and User Experience for Reading Long Texts in Virtual Reality. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 170, 8 pages. https://doi.org/10.1145/3544549.3585895

[28] Andrew Gelman and Donald B. Rubin. 1992. Inference from iterative simulation using multiple sequences. *Statistical science* 7, 4 (1992), 457–472. https://doi.org/10.1214/ss/1177011136

[29] Tsvetozar Georgiev. 2012. Investigation of the User's Text Reading Speed on Mobile Devices. In *Proceedings of the 13th International Conference on Computer Systems and Technologies* (Ruse, Bulgaria) *(CompSysTech '12)*. Association for Computing Machinery, New York, NY, USA, 329–336. https://doi.org/10.1145/2383276.2383324

[30] Luther C Gilbert. 1959. Speed of processing visual stimuli and its relation to reading. *Journal of Educational Psychology* 50, 1 (1959), 8. https://doi.org/10.1037/h0045592

[31] GA Hahn, D Penka, C Gehrlich, Andr'e Messias, M Weismann, L Hyv"arinen, Markku Leinonen, M Feely, Gary Rubin, C Dauxerre, et al. 2006. New standardised texts for assessing reading performance in four European languages. *British Journal of Ophthalmology* 90, 4 (2006), 480–484. https://doi.org/10.1136/bjo.2005.087379

[32] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909

[33] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[34] Eckhard H. Hess and James M. Polt. 1964. Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science* 143, 3611 (1964), 1190–1192. https://doi.org/10.1126/science.143.3611.1190 arXiv:https://www.science.org/doi/pdf/10.1126/science.143.3611.1190

[35] Brendan John, Pallavi Raiturkar, Arunava Banerjee, and Eakta Jain. 2018. An Evaluation of Pupillary Light Response Models for 2D Screens and VR HMDs. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 19, 11 pages. https://doi.org/10.1145/3281505.3281538

[36] James F. Juola, David Haugh, Scott Trast, F. Richard Ferraro, and Michael Liebhaber. 1987. Reading With and Without Eye Movements. In *Eye Movements from Physiology to Cognition*, J.K. O'REGAN and A. LEVY-SCHOEN (Eds.). Elsevier, Amsterdam, 499–508. https://doi.org/10.1016/B978-0-444-70113-8.50071-0

[37] Lari Kärkkäinen and Jari Laarni. 2002. Designing for Small Display Screens. In *Proceedings of the Second Nordic Conference on Human-Computer Interaction* (Aarhus, Denmark) *(NordiCHI '02)*. Association for Computing Machinery, New York, NY, USA, 227–230. https://doi.org/10.1145/572020.572052

[38] Matthew Kay, Steve Haroz, Shion Guha, and Pierre Dragicevic. 2016. Special Interest Group on Transparent Statistics in HCI. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 1081–1084. https://doi.org/10.1145/2851581.2886442

[39] Matthew Kay, Gregory L. Nelson, and Eric B. Hekler. 2016. Researcher-Centered Design of Statistics: Why Bayesian Statistics Better Fit the Culture and Incentives of HCI. 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, 4521–4532. https://doi.org/10.1145/2858036.2858465

[40] Kristopher Kohm, John Porter, and Andrew Robb. 2022. Sensitivity to Hand Offsets and Related Behavior in Virtual Environments over Time. *ACM Trans. Appl. Percept.* 19, 4, Article 17 (nov 2022), 15 pages. https://doi.org/10.1145/3561055

[41] Thomas Kosch, Mariam Hassib, Daniel Buschek, and Albrecht Schmidt. 2018. Look into My Eyes: Using Pupil Dilation to Estimate Mental Workload for Task Complexity Adaptation. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188643

[42] Thomas Kosch, Mariam Hassib, Paweł W. Woźniak, Daniel Buschek, and Florian Alt. 2018. Your Eyes Tell: Leveraging Smooth Pursuit for Assessing Cognitive Workload. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3174010

[43] Thomas Kosch, Jakob Karolus, Havy Ha, and Albrecht Schmidt. 2019. Your skin resists: exploring electrodermal activity as workload indicator during manual assembly. In *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (Valencia, Spain) *(EICS '19)*. Association for Computing Machinery, New York, NY, USA, Article 8, 5 pages. https://doi.org/10.1145/3319499.3328230

[44] Thomas Kosch, Jakob Karolus, Johannes Zagermann, Harald Reiterer, Albrecht Schmidt, and Paweł W. Woźniak. 2023. A Survey on Measuring Cognitive Workload in Human-Computer Interaction. *ACM Computing Surveys* (jan 2023). https://doi.org/10.1145/3582272

[45] Thomas Kosch, Albrecht Schmidt, Simon Thanheiser, and Lewis L. Chuang. 2020. One Does Not Simply RSVP: Mental Workload to Select Speed Reading Parameters Using Electroencephalography. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New

York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376766

[46] Jari Laarni. 2002. Searching for Optimal Methods of Presenting Dynamic Text on Different Types of Screens. In *Proceedings of the Second Nordic Conference on Human-Computer Interaction* (Aarhus, Denmark) *(NordiCHI '02)*. Association for Computing Machinery, New York, NY, USA, 219–222. https://doi.org/10.1145/572020.572050

[47] Jisoo Lee, Erin Walker, Winslow Burleson, Matthew Kay, Matthew Buman, and Eric B. Hekler. 2017. Self-Experimentation for Behavior Change: Design and Formative Evaluation of Two Approaches. 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, 6837–6849. https://doi.org/10.1145/3025453.3026038

[48] Jan Leusmann, Carl Oechsner, Johanna Prinz, Robin Welsch, and Sven Mayer. 2023. A Database for Kitchen Objects: Investigating Danger Perception in the Context of Human-Robot Interaction. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 6, 9 pages. https://doi.org/10.1145/3544549.3585884

[49] J.J.-W. Lin, H.B.L. Duh, D.E. Parker, H. Abi-Rached, and T.A. Furness. 2002. Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Proceedings IEEE Virtual Reality 2002*. 164–171. https://doi.org/10.1109/VR.2002.996519

[50] David Lindlbauer, Anna Maria Feit, and Otmar Hilliges. 2019. Context-Aware Online Adaptation of Mixed Reality Interfaces. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 147–160. https://doi.org/10.1145/3332165.3347945

[51] Tiffany Luong, Nicolas Martin, Anaïs Raison, Ferran Argelaguet, Jean-Marc Diverrez, and Anatole Lécuyer. 2020. Towards Real-Time Recognition of Users Mental Workload Using Integrated Physiological Sensors Into a VR HMD. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 425–437. https://doi.org/10.1109/ISMAR50242.2020.00068

[52] Guido Makransky, Thomas S. Terkildsen, and Richard E. Mayer. 2019. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction* 60 (2019), 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

[53] Sebastiaan Mathôt and Yavor Ivanov. 2019. The effect of pupil size and peripheral brightness on detection and discrimination performance. *PeerJ* 7 (2019), e8220. https://doi.org/10.7717/peerj.8220

[54] Manuel Oliva. 2019. Pupil size and search performance in low and high perceptual load. *Cognitive, Affective, & Behavioral Neuroscience* 19, 2 (April 2019), 366–376. https://doi.org/10.3758/s13415-018-00677-w

[55] Gustav Öquist and Kristin Lundin. 2007. Eye Movement Study of Reading Text on a Mobile Phone Using Paging, Scrolling, Leading, and RSVP. In *Proceedings of the 6th International Conference on Mobile and Ubiquitous Multimedia* (Oulu, Finland) *(MUM '07)*. Association for Computing Machinery, New York, NY, USA, 176–183. https://doi.org/10.1145/1329469.1329493

[56] Oskar Palinko, Andrew L. Kun, Alexander Shyrokov, and Peter Heeman. 2010. Estimating Cognitive Load Using Remote Eye Tracking in a Driving Simulator. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (Austin, Texas) *(ETRA '10)*. Association for Computing Machinery, New York, NY, USA, 141–144. https://doi.org/10.1145/1743666.1743701

[57] Bastian Pfleging, Drea K. Fekety, Albrecht Schmidt, and Andrew L. Kun. 2016. A Model Relating Pupil Diameter to Mental Workload and Lighting Conditions. 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, 5776–5788. https://doi.org/10.1145/2858036.2858117

[58] Andrea C. Pierno, Andrea Caria, Scott Glover, and Umberto Castiello. 2005. Effects of increasing visual load on aurally and visually guided target acquisition in a virtual environment. *Applied Ergonomics* 36, 3 (2005), 335–343. https://doi.org/10.1016/j.apergo.2004.11.002

[59] Elizabeth Quinn, Ian Stephen Paul Nation, and Sonia Millett. 2007. Asian and Pacific speed readings for ESL learners. *ELI Occasional Publication* 24 (2007), 74.

[60] Mehrak Rahimi and Seyyed Abolfazl Babaei. 2020. The impact of Rapid Serial Visual Presentation (RSVP) on awareness of reading strategies. *International Journal on Emerging Technologies* 11, 3 (2020), 1041–1048.

[61] Keith Rayner. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin* 124, 3 (1998), 372. https://doi.org/10.1037/0033-2909.124.3.372

[62] Miriam Reiner and Tatiana M. Gelfeld. 2014. Estimating mental workload through event-related fluctuations of pupil area during a task in a virtual world. *International Journal of Psychophysiology* 93, 1 (2014), 38–44. https://doi.org/10.1016/j.ijpsycho.2013.11.002 Applied Neuroscience: Functional enhancement, prevention, characterisation and methodology. (Hosting the Society of Applied Neuroscience).

[63] M. Richter, G.H.E. Gendolla, and R.A. Wright. 2016. Chapter Five - Three Decades of Research on Motivational Intensity Theory: What We Have Learned About Effort and What We Still Don't Know. Advances in Motivation Science, Vol. 3.

Elsevier, 149–186. https://doi.org/10.1016/bs.adms.2016.02.001

[64] Gary S. Rubin and Kathleen Turano. 1992. Reading without saccadic eye movements. *Vision Research* 32, 5 (1992), 895–902. https://doi.org/10.1016/0042-6989(92)90032-E

[65] Mark C. Russell and Barbara S. Chaparro. 2001. Exploring Effects of Speed and Font Size with RSVP. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 45, 6 (2001), 640–644. https://doi.org/10.1177/154193120104500614

[66] Rufat Rzayev, Polina Ugnivenko, Sarah Graf, Valentin Schwind, and Niels Henze. 2021. Reading in VR: The Effect of Text Presentation Type and Location. 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 531, 10 pages. https://doi.org/10.1145/3411764.3445606

[67] Rufat Rzayev, Paweł W. Woźniak, 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* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3173574.3173619

[68] Abhraneel Sarma and Matthew Kay. 2020. Prior Setting in Practice: Strategies and Rationales Used in Choosing Prior Distributions for Bayesian Analysis. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376377

[69] Hande Husniye Telek, Hidayet Erdol, Adem Turk, et al. 2018. The effects of age on pupil diameter at different light amplitudes. *Beyoglu Eye J* 3, 2 (2018), 80–85.

[70] Shangshang Zhu, Xinyu Su, and Yenan Dong. 2021. Effects of the Font Size and Line Spacing of Simplified Chinese Characters on Smartphone Readability. *Interacting with Computers* 33, 2 (08 2021), 177–187. https://doi.org/10.1093/iwc/iwab020 arXiv:https://academic.oup.com/iwc/article-pdf/33/2/177/40358726/iwab020.pdf

Received February 2024; revised May 2024; accepted June 2024