Supplemental Materials for Stimulating the Human Visual System Beyond Real World Performance in Future Augmented Reality Displays

David Dunn
UNC Chapel Hill

Okan Tursun
MPI Informatik, Germany
Università della Svizzera italiana, Switzerland

Hyeonseung Yu
MPI Informatik, Germany

Piotr Didyk
Università della Svizzera italiana, Switzerland

Karol Myszkowski
MPI Informatik, Germany

Henry Fuchs
UNC Chapel Hill

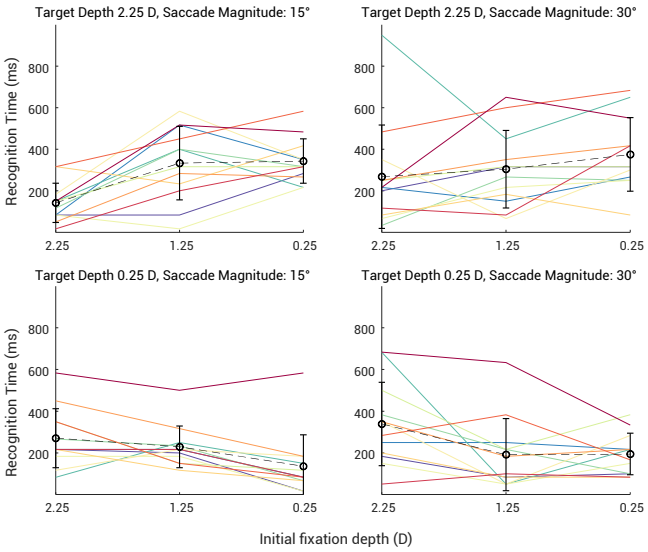


Figure 1: Individual user data from the combined latency experiment as described in Section 4. Each user is represented by a color, and the mean is indicated by the dashed line.

1 Structure

Section numbers correspond with the section they supplement from the main paper.

4 Measuring the Recognition Time During Combined Accommodation, Vergence, and Saccade

In Figure 1, the results for each individual user are reported for the experiment detailed in Section 4 of the main paper.

5 Overdriving Accommodation and Convergence During Saccade

5.1 Early Investigations

5.1.1 Monocular Overdrive Experiment

As reported in our main paper, the goal for this experiment was to see if overdriving the accommodation of a user was possible. For this study, there was no vergence change or saccade.

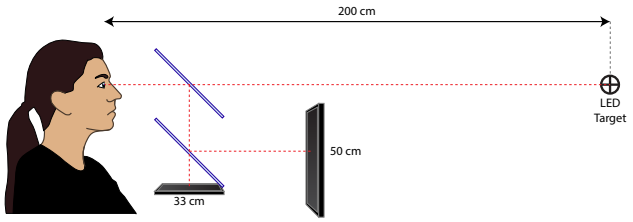


Figure 2: Display hardware configuration at Location B for running monocular focus over-driving experiment described in Section 5.1.1.

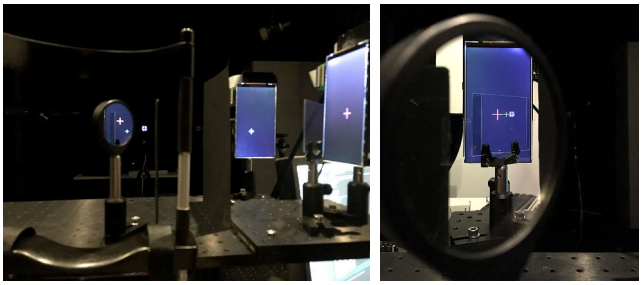


Figure 3: View of physical display hardware for running monocular over-driving focus experiment at Location B described in Section 5.1.1. Left) beamsplitters, LED target, and LCDs are visible. Right) slightly offset view through the display with calibration markers visible.

Hypothesis A user given over-driven focal cues in certain circumstances can perceive virtual content faster than if correct focal cues are provided.

Experimental Configuration The physical configuration for this experiment was at Location B and consisted of a chin and head rest, a controller for input, an LED target positioned 200 cm from the user, and two LCDs positioned 50 cm and 33 cm from the user visible via planar beamsplitters as seen in Figure 2 and Figure 3. The first beamsplitter was positioned in front of the participant's right eye, while the left eye was obstructed by means of a blacked-out barrier.

Stimuli A preparation stimulus was created by back illuminating the cross-hair target with a diffused LED. The main stimulus was comprised of a single Landolt C shape spanning $0°10'$ which sets the gap size to $0°2'$, where a normal 20/20 eye can identify a gap size of $0°1'$. With equal probability, the Landolt C shape was randomly

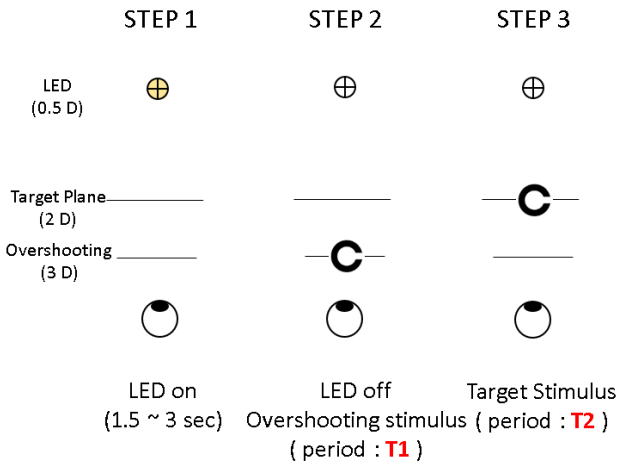


Figure 4: Experimental procedure for running monocular over-driving focus experiment described in Section 5.1.1.

oriented with the gap in one of the four directions, up, right, down, and left.

Participants Four subjects (all M, 25 to 35 years of age) that had a normal or corrected-to-normal vision, took part in the experiment. To keep participants inside the eyebox of our multiplane display, all participants used a chin and forehead rest.

Procedure The goal of the experiment was to find the detection threshold for identifying the orientation of the Landolt C for four independent test cases. For each test case, we ran a virulent PEST staircase procedure as described by [4] and [2]. All four staircase procedures were interleaved to run in parallel.

A single trial — described pictorially in Figure 4 — consisted of 4 phases:

- Preparation phase, this phase was responsible for ensuring correct initial focus at 200 cm (0.5 D(diopter)) by illuminating the LED target at the initial depth. This stimulus was illuminated for a random period between 1.5 s and 3 s.
- An overdrive phase in which the main randomly oriented Landolt C stimulus was displayed on the 33 cm (3 D) screen for a specified number of frames $T1$.
- The main stimulus phase where the same Landolt C stimulus from the overdrive phase was displayed for a variable number of frames $T2$ at the destination focus depth on the 50 cm (2 D) screen.
- A response phase where as soon as the main phase time expired, the stimulus disappeared. This phase continued until the user responded with the orientation of the Landolt C stimulus.

After the response was given, the next trial began. The variation between the four test cases was the length of time the over-driven stimulus was visible, or $T1$. Times corresponding to the four cases are 0 ms, 67 ms, 133 ms, and 200 ms. The 0 ms test case corresponds to a non-over-driven control case where the overdrive phase is skipped entirely. Each vPEST staircase procedure, before ending, independently finds the perceptual threshold for the free variable $T2$, which was the number of frames the stimulus was visible in the main phase before disappearing at the start of the response phase.

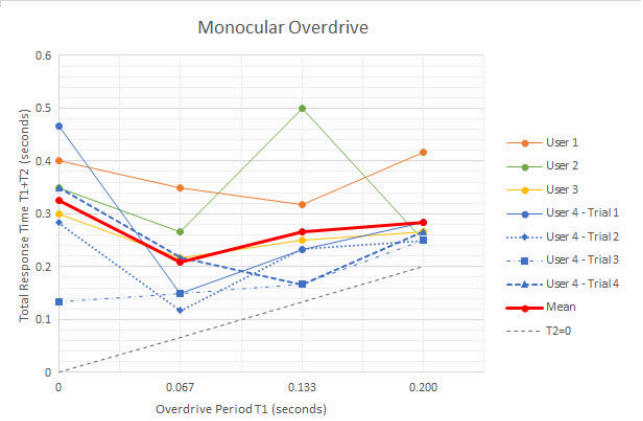


Figure 5: The results for the monoscopic over-driving focus experiment described in Section 5.1.1. Each point represents the threshold detected from an entire staircase procedure. The 4 test cases with differing overdrive periods $T1$ are plotted on the horizontal axis, while the total time the stimulus was visible $T1+T2$ including both the overdrive and main phase times are plotted vertically.

Results A plot of the results can be seen in Figure 5. Each point represents the threshold detected from an entire staircase procedure. The 4 test cases with differing overdrive times $T1$ are plotted on the horizontal axis, while the total time the stimulus was visible $T1+T2$ including both the overdrive and main phase times are plotted vertically.

Our hypothesis can be tested by comparing the $T1=0$ control case on the far left to the other cases with some amount of overdrive time for the same user. As can be seen, for each user, with the exception of user 4 trial 3, there was at least one case where the total time to identify the stimulus was improved with the additional overdrive time. This indicates that not only is over-driving the accommodation response possible, but that it can speed up perception of visual stimuli.

In the single case that did not exhibit a speed up, user 4 trial 3, we speculate that the control response time was so quick that there wasn't very much room for improvement, as can be seen by how close it is to the $T2=0$ boundary line. All the other trials from the same user exhibited the expected behavior, so trial 3 is viewed as an outlier.

Discussion While we didn't run enough users through the experiment to make any generalized claims based on the results, the overall result showed the promise of over-driving focus, which merited further study. In this study we examined the one-eye case which only accounted for speeding up the eye's accommodation. While this result may be useful for users with depleted vision in one eye, such as those with amblyopia, or those with stereo blindness, we wanted to know if the same speed-up would be exhibited in binocular situations where both accommodation and vergence are operating in concert. So we developed and ran the experiment described in Section 5.1.2 below.

5.1.2 Pilot Study

With the promise of enhancement to perception indicated by the previous accommodation-only study for a single eye, we sought to extend the principle to both eyes when vergence and accommodation work together. Vergence has been shown to be accelerated when accompanied by a saccade, thus with this experiment, we investigated if adding a saccade while over-driving the focus and vergence would produce the improved response times.

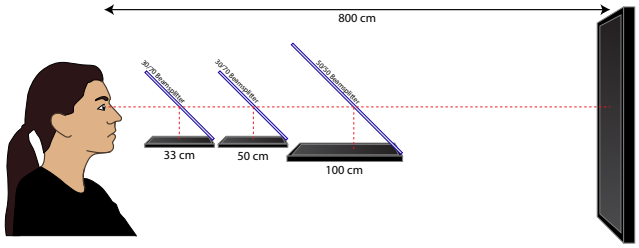


Figure 6: Display hardware configuration at Location A for running binocular accommodation and vergence over-driving experiment described in Section 5.1.2. Four displays at depths of 33 cm (3 D), 50 cm (2 D), 100 cm (1 D), and 800 cm (0.125 D) combined with a series of planar beamsplitters with increasing reflectance ratios.

Hypotheses

1. A user given over-driven focal and vergence cues can perceive virtual content faster than if correct focal and vergence cues are provided.

2. A user given over-driven focal and vergence cues during saccadic movement can perceive virtual content faster than if correct focal and vergence cues are provided.

3. A user given over-driven focal and vergence cues during saccadic movement can perceive virtual content faster than without the saccade

Experimental Configuration The physical configuration for this experiment was at Location A and consisted of a chin and head rest, a controller for input, and a 4-depth multiplane display which employed a combination of LCDs and planar beamsplitters. The 3 beamsplitters and 4 LCD screens were distributed such that the depth of the images were 33 cm, 50 cm, 100 cm, and 800 cm distant from the user, as can be seen in Figure 6 and main paper Figure 3. Stereoscopic views between the 4 display depths were not supported by the display, so virtual objects could only be presented at one of the 4 fixed depths.

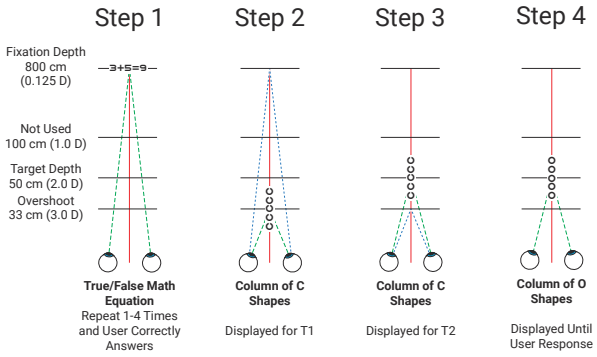
Stimuli For this experiment the preparation stimulus consisted of a series of randomized true/false math equation with the form:

$$x + y = z \tag{1}$$

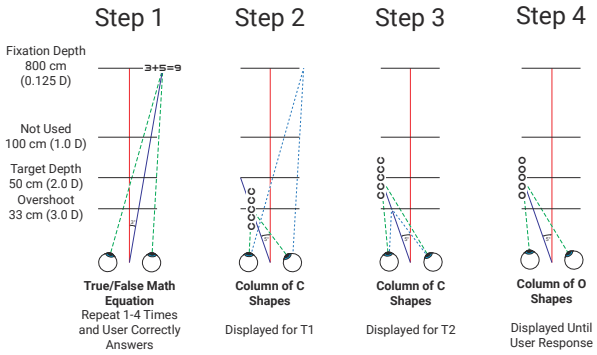
where x and y were single digit integers and z was the answer. When the equation was false, z was altered by adding δ where $\delta \in \{n \in \mathbb{Z} \mid -2 \leq n \leq 2, n \neq 0\}$. The equations had an even probability of being true or false. The equation was centered along the line of sight with the characters subtending $0°9'$ visual angle each and the entire equation subtending a horizontal visual angle of $0°45'$ or $0°54'$.

The main stimulus consisted of a vertical column of 12 Landolt C shapes, all oriented with the gap in the same one of the four directions: up, right, down, and left. Each Landolt C spanned $0°10'$ making the gap size $0°2'$. The entire column of Landolt C shapes subtended a vertical angle of $3°50'$. The choice of presenting more than one shape was made to avoid adding an indeterminate amount of visual scanning time to each trial while the user attempts to locate the relatively small stimulus. Enlarging the stimulus would have decreased the visual acuity required to correctly identify the orientation which would have been detrimental to the results. A vertical column was chosen to avoid any vergence ambiguity that multiple horizontal stimuli would introduce.

During the response phase, an additional response stimulus was displayed allowing the user to continue to fixate at the target depth



(a) Non-saccade procedure



(b) Saccade procedure

Figure 7: Experimental procedure for running binocular over-driving accommodation and vergence experiment with and without saccade described in Section 5.1.2.

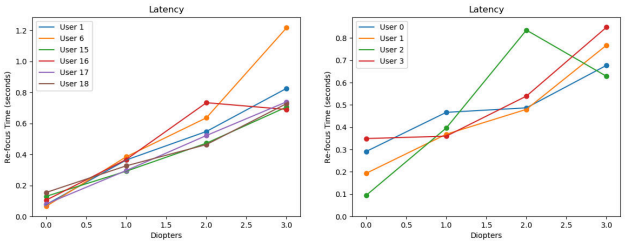
while they selected their response. The response stimulus was composed of a vertical column of 12 circular shapes co-located with the main stimulus. Thus the only difference between the main stimulus and the response stimulus was the gap in each Landolt C, which simply disappeared when the response phase was entered.

For the non-saccade trials, the stimuli were all presented on the medial visual axis. For the saccade trials, the preparation stimulus was located 3° right of the user's central field, the main stimulus wass located 5° left of the user's central field, and the response stimulus was similarly located 5° left of the user's central field, for a saccade magnitude of 8°.

Participants Six subjects (2 F, 4 M, 23 to 44 years of age) participated in the non-saccade trials, and four subjects (all M, 26 to 35 years of age) took part in the saccade trials. To keep participants inside the eyebox of our near-eye display, all participants used a chin and forehead rest.

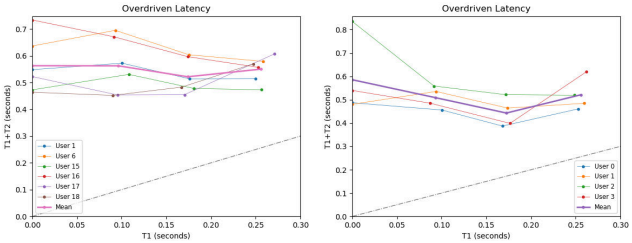
Procedure We began the procedure by performing a one-time-per-subject display alignment calibration, which aligns the subject's eyes vertically with our display and performs an interpupillary distance adjustment. Then, user acuity was tested with a 3-down/1-up staircase procedure for detecting the orientation of a Landolt C shape using the size of the shape as the free variable. Users with acuity better than 20/30 were allowed to continue. At that point a basic training sequence familiarized the subject with the task and procedure and tested the user to ensure correct responses with a series of practice rounds. Three correct responses were required to continue on to the main experiment.

The goal of the experiment was to find the detection threshold for



(a) No saccade trials (b) 8° saccade trials

Figure 8: The time required to recognize a stimulus after the depth change given in diopters both with and without a saccade of 8° during the over-driving focus and vergence experiment.



(a) No saccade trials (b) 8° saccade trials

Figure 9: The time required to recognize a stimulus after a 2 D depth change with and without 8° saccade when the stimulus is over driven for $T1$.

identifying the orientation of the Landolt C on seven independent test cases. For each test case, we ran a virulent PEST staircase procedure as described by [4] and [2]. All seven staircase procedures were interleaved to run in parallel. Four of the test cases were analogous to the monocular experiment, all having the same depth change (0.125 D to 2 D, but with different over-drive periods. The other three test cases had no over-drive periods, but instead varied the amount of depth change between the preparation stimulus and the main stimulus. By randomizing the amount of depth change, these three extra cases worked to prevent a learning effect where the subject would anticipate the focus change.

A single trial — described pictorially in Figure 7 — consisted of 4 phases:

- The preparation phase ensured correct initial focus depth at 800 cm (0.125 D) by presenting a sequence of the preparation stimuli. The initial number of problems in the sequence was determined by a random choice from the distribution $(1,1,1,1,2,2,2,2,3,3,4)$. Each stimulus is shown until the user responds. If a user does not answer the last problem correctly, an additional problem is presented. This randomization was also introduced to prevent a learning effect where the subject would anticipate the focus change.

- Between the preparation phase and main stimulus phase, an overdrive phase would present the main stimulus on the 33 cm (3 D) screen for a specified number of frames $T1$.

- The main stimulus phase where the main stimulus was displayed at the target depth d for a variable number of frames $T2$.

- A response phase where as soon as the main phase time expired, the main stimulus would be replaced by a response stimulus at the same location. This phase continued until the user responded with the orientation of the main stimulus gap.

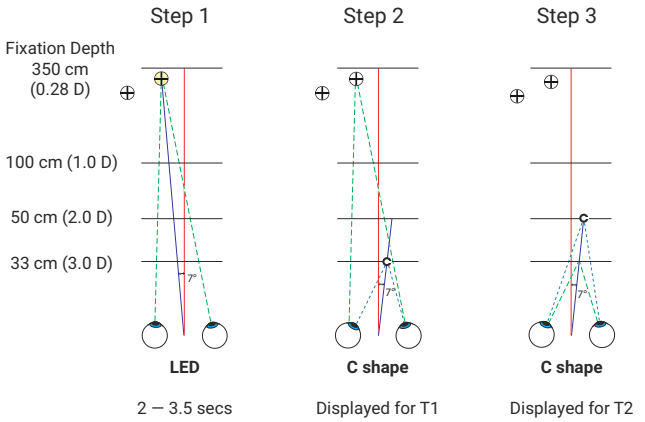


Figure 10: Top view of experimental procedure for running binocular overdriving accommodation and vergence during saccade experiment described in Section 4.2 of the main paper (Experiment Setup 2). The figure shows one test case where a 14° saccade is combined with overdrive and target depth levels at 3 D and 2 D, respectively. Please note that this experiment additionally tests for other saccade magnitudes of 7°, 21°, and 28° and a convergence to 1 D with overdrive at 2 D.

After the response was given, the next trial began. The variation between the four main test cases was the length of time the overdriven stimulus was visible, or $T1$. Frames corresponding to the four cases are 0 frames (0 ms), 5 frames (~83 ms), 10 frames (~167 ms), and 15 frames (~250 ms). For all four main test cases, the target depth d is set to 50 cm (2 D). The 0 ms test case corresponds to a non-over-driven control case where the overdrive phase is skipped entirely.

The additional three test cases also have $T1$ set to zero, meaning no overdrive, however their target depths d are set as: 800 cm (0.125 D), 100 cm (1 D), and 33 cm (3 D). Each vPEST staircase procedure, before ending, independently finds the perceptual threshold for the free variable $T2$, which was the number of frames the stimulus was visible in the main phase before disappearing at the start of the response phase. The entire procedure took about 30 min of time to complete.

Results We can glean two sets of results from these experiments. The first is visualized in Figure 8, which is a plot of the three additional test cases and the control case from the main test set, or all test cases where $T1 = 0$. Again, each point represents the threshold result of an entire staircase. In this plot, the target depth d makes the horizontal axis, while the thresholded amount of time required to correctly identify $T2$ is the vertical axis. Here we expect to see that as the amount of focus change increases between the preparation and main stimuli, that the amount of time also increases, and this is exhibited nicely, in an almost linear manner. These data also show that our measures against subjects anticipating the focus change worked. The 0 D case has the lowest latency for all cases. For trials with the added saccade, times for the 0 D case were longer, and the results have a larger distribution.

The second set of results, similar to the results of the previous experiment, will determine the validity our hypotheses. Although there are no statistically significant results due to the lack of participants, by looking at Figure 9, we see promising results. The no saccade trials did not significantly speed up perception as we thought it would, but the saccade trials did show promising results. With the addition of 10 frames of over-drive, all subjects identified the orientation of the stimuli faster than without the overdrive.

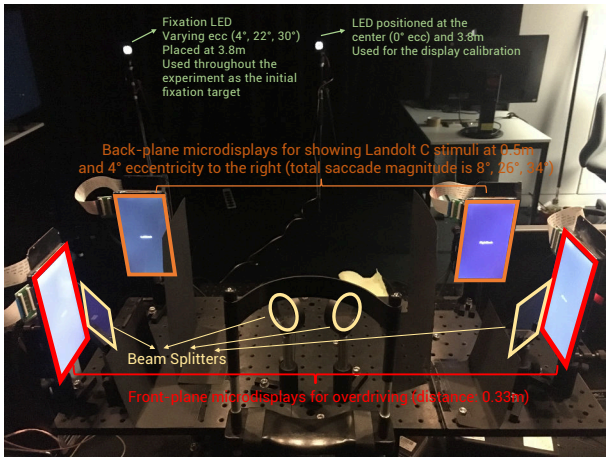


Figure 11: View of Experiment Setup 1 for running the accommodation and convergence overdriving during saccade described in Section 4.2 of the main paper.

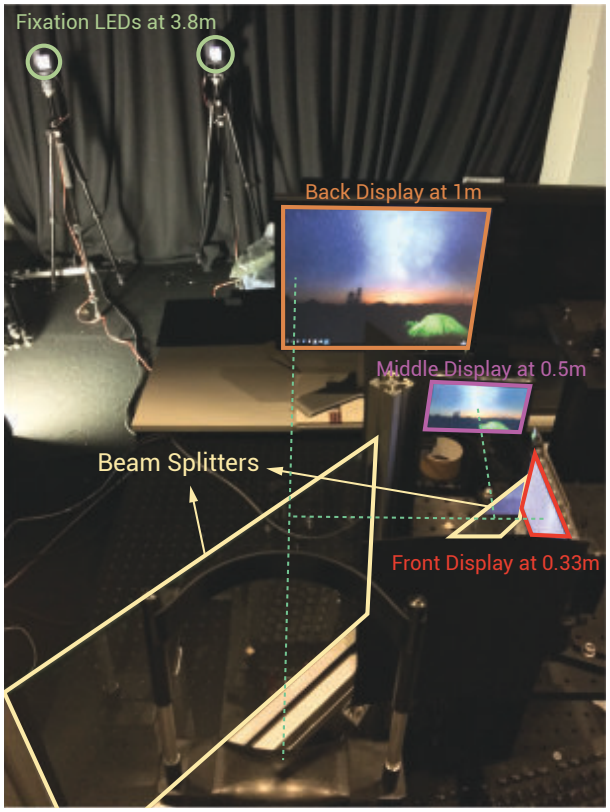


Figure 12: View of Experiment Setup 2 for running the accommodation and convergence overdriving during saccade described in Section 4.2 of the main paper.

5.2 Additional Details for Experiment Setup 1 and 2

Mean recognition times measured in this experiment and associated p-values are provided in Table 1 and 2. On Experiment Setup 1, we observe a speedup in all saccade combinations; however, only the speedup measured with a saccade magnitude of 8° and overdrive duration of T1=83 ms is statistically significant. We also observe that a longer overdrive duration of T1=250 ms usually slows down the recognition compared to the non-overdriven stimulus. A representative image of this experiment setup is provided in Figure 11.

Experiment Setup 2 shows similar speedups with saccade magnitudes of 7° and 14° when target and overdrive depth levels are at 2 D and 3 D, respectively. However, the statistically significant speedup is observed at T1=167 ms for the 7° instead of T1=83 ms. In addition, we see that the recognition times are longer on average for this set of target and overdrive depth levels on Experiment Setup 2. We attribute prolonged recognition times to the optical imperfections caused by the lack of an anti-reflective coating on one of the half mirrors and resulting subtle ghost image of the stimulus. This setup provides us with an extra target at 1 D (Figure 10). We do not observe any meaningful improvement in recognition times for the target depth at 1 D. In that case, the only speedup observed with a saccade magnitude of 28°. This is not surprising as there is not much room for improvement due to shorter durations associated with smaller change in vergence.

Individual timings from each participant are shown in Figure 13 and 14. We observe large variability among the participants, which is also reported by previous studies with instrument-based measurements [1, 3, 5]. The plots in Figure 14 include participants P1 and P7, who are regarded as outliers. We attribute extremely long recognition times of participant P1 at 7° saccade and participant P7 at 28° saccade to lack of attention. Their data is discarded from all of the statistical analysis because the deviation from the mean is approximately 3σ. However, as a reference we provide timings obtained from all participants as well as the outliers in Figure 14.

Table 1: Mean recognition times observed on Experiment Setup 1 for different values of saccade magnitude (SM), overdrive depth (OD) and target depth (TD) settings. p-values are indicated between parentheses (H_0: Mean recognition time of the overdriven stimulus is equal to that of non-overdriven stimulus, where T1=0 ms). The statistically significant speedup is shown in bold ($p \leq 0.05$, Wilcoxon signed-rank test).

		Mean Recognition Time (T1+T2) (ms)			
SM	**OD → TD**	**T1=0**	**T1=83ms**	**T1=167ms**	**T1=250ms**
8°	3 D → 2 D	373.3	**326.7(0.02)**	348.3(0.26)	366.7(0.76)
26°	3 D → 2 D	409.5	378.6(0.35)	371.4(0.67)	495.2(0.06)
34°	3 D → 2 D	356.7	306.7(0.22)	360.0(0.89)	380.0(0.72)

Table 2: Mean recognition times observed on Experiment Setup 2 for different values of saccade magnitude (SM), overdrive depth (OD) and target depth (TD) settings. p-values are indicated between parentheses using the same statistical test as Table 1. Statistically significant speedups are shown in bold ($p \leq 0.05$, Wilcoxon signed-rank test).

		Mean Recognition Time (T1+T2) (ms)		
SM	**OD → TD**	**T1=0**	**T1=83ms**	**T1=167ms**
7°	3 D → 2 D	482.9	454.2(0.39)	**410.5(0.03)**
14°	3 D → 2 D	511.6	**431.6(0.05)**	470.8(0.24)
21°	3 D → 2 D	528.2	529.7(0.92)	549.3(0.66)
28°	3 D → 2 D	523.7	513.1(0.80)	511.6(0.76)
7°	2 D → 1 D	230.9	282.2(< 0.01)	274.7(0.05)
14°	2 D → 1 D	273.1	310.9(0.14)	345.6(< 0.01)
21°	2 D → 1 D	267.1	330.5(0.02)	341.1(0.02)
28°	2 D → 1 D	365.2	354.6(0.35)	384.8(0.35)

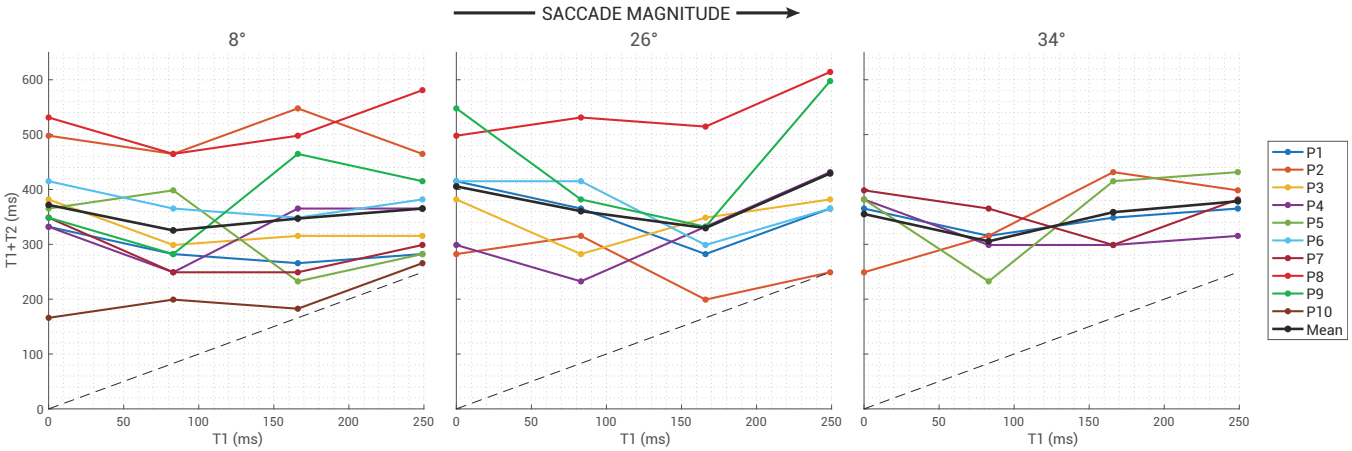


Figure 13: Recognition times measured in the accommodation and convergence overdriving during saccade described in Section 4.2 of the main paper (Experiment Setup 1).

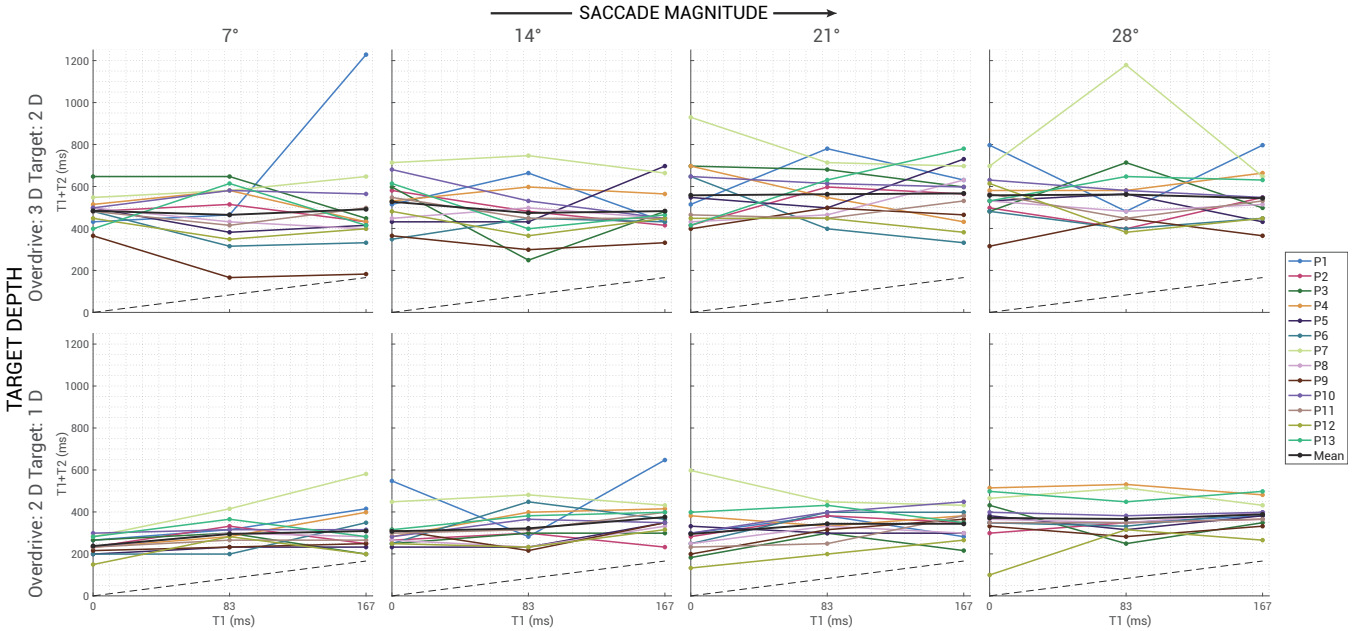


Figure 14: Recognition times measured in the accommodation and convergence overdriving during saccade described in Section 4.2 of the main paper (Experiment Setup 2). Participants P1 and P7 are discarded from further analysis because of their significantly longer recognition times.

REFERENCES

[1] O. A. Coubard and Z. Kapoula. Saccades during symmetrical vergence. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 246(4):521–536, 2008.

[2] J. M. Findlay. Estimates on probability functions: A more virulent PEST. *Perception & Psychophysics*, 23(2):181–185, Mar. 1978. doi: 10.3758/BF03208300

[3] M. Takagi, E. M. Frohman, and D. S. Zee. Gap-overlap effects on latencies of saccades, vergence and combined vergence-saccades in humans. *Vision Research*, 35(23-24):3373–3388, 1995.

[4] M. M. Taylor and C. D. Creelman. PEST: Efficient Estimates on Probability Functions. *The Journal of the Acoustical Society of America*, 41(4A):782–787, Apr. 1967. doi: 10.1121/1.1910407

[5] Q. Yang and Z. Kapoula. Saccade–vergence dynamics and interaction in children and in adults. *Experimental Brain Research*, 156(2):212–223, 2004.