Workload Assessment for Mental Arithmetic Tasks using the Task-Evoked Pupillary Response

G. Marquart
Workload Assessment for Mental Arithmetic Tasks using the Task-Evoked Pupillary Response

By

G. Marquart

in partial fulfilment of the requirements for the degree of

Master of Science
in Mechanical Engineering

at the Delft University of Technology,
to be defended publicly on Wednesday June 3, 2015 at 02:00 PM.

Supervisor: Dr.ir. J.C.F. de Winter, TU Delft
Thesis committee: Dr.ir. R. Happee, TU Delft
Dr. D. Dodou, TU Delft
Ir. P.M. van Leeuwen, TU Delft

An electronic version of this thesis is available at http://repository.tudelft.nl/.
A smooth interaction between human and machine is required in order to provide the optimal joint performance. Humans do not only have to understand the machine they are working with, but the machine also needs to understand the human. The human's mental state is one of those things that need to be understood by machines, and many researchers have tried to find a valid and reliable way to achieve this.

With this MSc thesis, I hope to contribute to the investigation of the pupil diameter as a measure of mental workload. This report is part of the fulfilment for the degree of Master of Science in Mechanical Engineering at the Delft University of Technology and is the result of a year of research under the supervision of Joost de Winter.

This thesis consists of six chapters. Chapter one provides the research paper which describes the investigation to the assessment of mental workload for mental arithmetic tasks using the task-evoked pupillary response. In the three subsequent chapters the results are reported of two preliminary studies and one pilot study. Chapter five includes a literature paper which focusses on the eye-related physiological measures of drivers’ mental workload. A condensed version of this paper (Marquart et al., 2015) will be presented at the 6th International Conference on Applied Human Factors and Ergonomics. The sixth chapter contains all the appendices providing background information, additional results, and supplementary materials.

Lastly, I would like to thank Joost de Winter for supervising me during my graduation process. I could not have written this thesis without him. He provided me with feedback on a regular basis allowing me to develop as a future engineer. Additionally, I would like to thank Christopher Cabrall for reviewing my literature report and preparing it for the conference and Peter van Leeuwen for his support with the eye tracker. Special thanks go to my parents for giving me the opportunity to study in Delft and my girlfriend Britty for her love and support.

G. Marquart
Delft, June 2015
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong> Thesis Report</td>
<td>7</td>
</tr>
<tr>
<td>1.1. Abstract</td>
<td>7</td>
</tr>
<tr>
<td>1.2. Introduction</td>
<td>7</td>
</tr>
<tr>
<td>1.3. Method</td>
<td>9</td>
</tr>
<tr>
<td>1.4. Results</td>
<td>11</td>
</tr>
<tr>
<td>1.5. Discussion</td>
<td>18</td>
</tr>
<tr>
<td>1.6. References</td>
<td>20</td>
</tr>
<tr>
<td><strong>2</strong> Preliminary Study 1</td>
<td>22</td>
</tr>
<tr>
<td>2.1. Introduction</td>
<td>22</td>
</tr>
<tr>
<td>2.2. Setup</td>
<td>22</td>
</tr>
<tr>
<td>2.3. Data Processing</td>
<td>22</td>
</tr>
<tr>
<td>2.4. Results</td>
<td>23</td>
</tr>
<tr>
<td>2.5. Conclusion</td>
<td>25</td>
</tr>
<tr>
<td>2.6. References</td>
<td>25</td>
</tr>
<tr>
<td><strong>3</strong> Preliminary Study 2</td>
<td>26</td>
</tr>
<tr>
<td>3.1. Introduction</td>
<td>26</td>
</tr>
<tr>
<td>3.2. Setup</td>
<td>26</td>
</tr>
<tr>
<td>3.3. Data Processing</td>
<td>27</td>
</tr>
<tr>
<td>3.4. Results</td>
<td>27</td>
</tr>
<tr>
<td>3.5. Conclusion</td>
<td>28</td>
</tr>
<tr>
<td><strong>4</strong> Pilot Study</td>
<td>29</td>
</tr>
<tr>
<td>4.1. Introduction</td>
<td>29</td>
</tr>
<tr>
<td>4.2. Setup</td>
<td>29</td>
</tr>
<tr>
<td>4.3. Data Processing</td>
<td>30</td>
</tr>
<tr>
<td>4.4. Results</td>
<td>30</td>
</tr>
<tr>
<td>4.5. Conclusion</td>
<td>32</td>
</tr>
<tr>
<td>4.6. Discussion</td>
<td>33</td>
</tr>
<tr>
<td>4.7. References</td>
<td>33</td>
</tr>
<tr>
<td><strong>5</strong> Literature Report</td>
<td>34</td>
</tr>
<tr>
<td>5.1. Abstract</td>
<td>34</td>
</tr>
<tr>
<td>5.2. Introduction</td>
<td>34</td>
</tr>
<tr>
<td>5.3. Drivers’ Mental Workload</td>
<td>35</td>
</tr>
<tr>
<td>5.4. Blinks</td>
<td>36</td>
</tr>
<tr>
<td>5.5. Fixations</td>
<td>38</td>
</tr>
<tr>
<td>5.6. Pupillometry</td>
<td>40</td>
</tr>
<tr>
<td>5.7. Discussion</td>
<td>43</td>
</tr>
<tr>
<td>5.8. References</td>
<td>45</td>
</tr>
<tr>
<td><strong>6</strong> Appendices</td>
<td>49</td>
</tr>
<tr>
<td>6.1. Appendix A: Informed Consent Form for Participants</td>
<td>50</td>
</tr>
<tr>
<td>6.2. Appendix B: Classification of Arithmetic Tasks II</td>
<td>52</td>
</tr>
<tr>
<td>6.3. Appendix C: Blink Identification and Removal</td>
<td>53</td>
</tr>
<tr>
<td>6.4. Appendix D: Individual Trials</td>
<td>54</td>
</tr>
</tbody>
</table>
6.5. Appendix E: MPD and MPDC bar graphs .................................................................56
6.6. Appendix F: Eight-point analysis of correct and incorrect responses ...............57
6.7. Appendix G: Results Ahern (1978) and Klingner (2010) .....................................58
6.8. Appendix H: Classification of Arithmetic Tasks I .................................................61
6.9. Appendix I: Additional Results ..............................................................................62
6.10. Appendix J: Measurement Equipment ..................................................................64
6.11. Appendix K: Experiment Interface .......................................................................65
6.12. Appendix L: Data Transfer and Logging ...............................................................67
6.13. Appendix M: Literature Blinks ..............................................................................68
6.15. Appendix O: Literature MPDC/MPDCR .............................................................71
6.16. Appendix P: MATLAB Code .................................................................................72
1.1. Abstract

Pupillometry is a promising method for assessing mental workload and could be helpful in the optimization of systems that involve human-computer interaction. The present study focuses on replicating the pupil diameter study by Ahern (1978) for mental multiplications of varying difficulty, using an automatic remote eye tracker. Our results showed that the findings of Ahern were replicated and that the mean pupil diameter and mean pupil diameter change (MPDC) discriminated just as well between the three difficulty levels as did a self-report questionnaire of mental workload (NASA-TLX). A higher mean blink rate was observed during the multiplication period for the highest level of difficulty in comparison with the other two levels. Moderate to strong correlations were found between the MPDC and the proportion of incorrect responses, indicating that the MPDC was higher for participants with a lower performance. For practical applications, validity could be improved by combining pupillometry with other physiological techniques.

1.2. Introduction

Mental workload is an important psychological construct that is challenging to assess on a continuous basis. A commonly used definition of mental workload is the one proposed by Hart and Staveland (1988). These authors defined workload as the “cost incurred by human operators to achieve a specific level of performance”. A valid and reliable assessment method of workload could be helpful in the optimization of systems that involve human-computer interaction, such as vehicles, computers, and simulators. One promising method for measuring workload is pupillometry, which is the measurement of the pupil diameter (e.g., Granholm & Steinhaeuer, 2004; Marshall, 2007; Schwalm et al., 2008; Klingner et al., 2008; Palinko et al., 2010; Goldinger & Papesh, 2012; Laeng et al., 2012).

Two antagonistic muscles regulate the pupil size, the sphincter and the dilator muscle. Activation of these muscles results in the constriction and dilation of the pupil, respectively. During a mentally demanding task, the pupils have been found to dilate up to 0.5 mm, which is small compared to the maximum dilation of about 6 mm caused by changes in lighting conditions. The involuntary reaction is also called the task-evoked pupillary response (TEPR; Beatty, 1982). In the past, TEPRs were obtained at 1 to 2 Hz by motion picture photography (Hess & Polt, 1964). This required researchers to measure the pupil diameter manually frame by frame (Janisse, 1977).
Nowadays, remote non-obtrusive eye trackers are increasingly being used to automatically measure TEPRs, as these devices are getting more and more accurate.

Over the years, researchers have encountered a few challenges in pupillometry. Reflexes of the pupil to changes in luminance, for example, may undermine the validity of TEPRs. One way to achieve this is by strictly controlling luminance, but this limits the usability of pupillometry. Marshall (2000) reported to have found a valid way to filter out the pupil light reflex using wavelet transform techniques. She patented this method and dubbed it the “index of cognitive activity”. The influence of gaze direction on the measured pupil size is another issue. Whereas Pomplun and Sunkara (2003) reported a systematic dependence of pupil size on gaze direction, Klingner et al. (2008) argued that the ellipse-fitting method for the estimation of the pupil size is not affected by perspective distortion.

In the last few decades many researchers have investigated the pupillary response for different types of tasks. Typically, the dilation was found to be higher for more challenging tasks (Beatty & Kahneman, 1966; Ahern, 1978). Not only task demands have been found to influence the pupil diameter, but also factors like anxiety, stress, and fatigue. Tryon (1975) and Janisse (1977) extensively reviewed known sources regarding variation in pupil size. Back then, he commented about the underexplored area of individual differences in intelligence. Ahern (1978) continued on this topic and discovered that persons scoring higher on intelligence tests showed smaller pupillary dilations on tasks of fixed difficulty. In a more recent study, Van Der Meer et al. (2010) found greater pupil dilations for individuals with high fluid intelligence than with low fluid intelligence during the execution of geometric analogy tasks. Thus, the results are not consistent and demand further investigation.

The present study focuses on replicating the film-based pupil diameter study by Ahern (1978) for mental multiplications of varying difficulty (43 participants, 1376 trials), and is intended as a follow-up study of Klingner (2010). Klingner replicated Ahern’s results with an automatic remote eye tracker and found a clear difficulty effect, with the more difficult multiplications showing a greater dilation. With more participants (30 vs. 7) and trials (1350 vs. 431) than Klingner, the present study aims to analyze the TEPRs for three levels of difficulty in high temporal detail, attempts to provide new insights into individual differences, and compare the effect sizes between the pupil diameter and a classic subjective measurement method of workload, the NASA-TLX. Additionally, the mean pupil diameter change rate (MPDCR) will be examined, which is a new measure introduced by Palinko et al. (2010). He expected it to be useful in assessing moment-to-moment changes in mental workload. Lastly, this study discusses the feasibility of using the pupil diameter in practical applications. One example of such an application is adaptive automation, which is “an approach to automation design where tasks are dynamically allocated between the human operator and computer systems” (Byrne & Parasuraman, 1996). As mentioned above, reliability and validity are crucial in this.

The digits in the task in this study were presented visually, in contrast to the experiment conducted by Ahern, where the digits were presented aurally. This was done to gain more temporal consistency in the presentation duration of the numbers. Like Klingner (2010), the pupil diameter was recorded with an automatic remote eye tracker.
1.3. Method

*Ethics statement.* The research was approved by the Human Research Ethics Committee (HREC) of the Delft University of Technology (TU Delft). All participants provided written informed consent.

*Participants.* Thirty participants (2 women and 28 men), aged between 19 and 38 years (mean = 23, SD = 4.1 years) were recruited to volunteer in this experiment (25 MSc/BSc, 3 PhD, and 2 graduate students). Individuals wearing glasses or lenses were excluded from participation. All participants read and signed an informed consent form (see Appendix A), explaining the purpose and procedures of the experiment and received €5 in compensation for their time.

*Equipment.* The SmartEye DR120 remote eye tracker, with a sampling rate of 120 Hz, was used to record the participant’s pupil diameter, eyelid opening, and gaze direction while sitting behind a desktop computer (see Fig. 1, left). The pupil diameter was estimated by averaging the five longest lines found in the pupil (Wilhelm, 2010). This method is comparable to the ellipse-fitting method, since they are both unaffected by perspective distortion. In order to obtain accurate measurements, a headrest was used to avoid head displacements. The eye tracker was equipped with a 24-inch screen, which was positioned approximately 65 cm in front of the sitting participant and was used to display task-relevant information. The outcome of a task had to be entered using the numeric keypad of a keyboard. The experiment took place in a room where there was office lighting and where daylight could not enter. A screen background with variable brightness was used, which was designed to minimize the pupillary light reflex in case a participant looked away from the center of the screen (see Fig. 1, right; Marquart, 2015).

![Figure 1. Experimental equipment. Left: eye tracker, monitor, table, headrest, chair, keyboard. Right: task display.](image)

*Procedure.* The participants were requested to perform 50 trials of mental arithmetic tasks (multiplications of two numbers), five of which were used as a short training. The remaining 45 trials were sorted by the outcome of their multiplication and evenly divided into 3 sessions of varying difficulty (easy, medium, and hard; see Appendix B). Level 1 contained the 15 easiest multiplications (outcomes ranging between 72 and 117), Level 2 contained 15 multiplications of intermediate difficulty (outcomes between 119 and 192), and Level 3 contained the 15 hardest multiplications (outcomes between 196 and 324).

The sequence of the three sessions was counterbalanced across the participants. Each trial was initiated by the participant with a button press and started with a 4 second accommodation period, followed by a 1 second visual presentation of two numbers (multiplicand and multiplier) between 6 and 18, with a 1.5 second pause in between. The participants were asked to multiply the two numbers and type their answer on the numeric keypad 10 seconds after the multiplier disappeared (see Table 1). Thus, the total duration of one trial was 17.5 seconds (4 + 1 + 1.5 + 1...
+ 10). When the numbers were not presented, a double “X” was shown to avoid pupillary reflexes caused by changes in brightness or contrast.

After each session, participants were asked to fill out a NASA-TLX questionnaire to assess their subjective workload on six facets: mental demand, physical demand, temporal demand, performance, effort, and frustration (Hart & Staveland, 1988). All questions were answered on a scale from 0 % (very low) to 100 % (very high). For the performance question, 0 % meant perfect and 100 % was failure. The participants’ overall subjective workload was obtained by averaging the scores across the six items. The total duration of the experiment was approximately 30 minutes.

Table 1. Timeline of an individual trial.

<table>
<thead>
<tr>
<th>Period</th>
<th>Start time (s)</th>
<th>End time (s)</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accommodation</td>
<td>0.0</td>
<td>4.0</td>
<td>X X</td>
</tr>
<tr>
<td>Baseline</td>
<td>3.6</td>
<td>4.0</td>
<td>X X</td>
</tr>
<tr>
<td>Multiplicand</td>
<td>4.0</td>
<td>5.0</td>
<td>0 8</td>
</tr>
<tr>
<td>Pause</td>
<td>5.0</td>
<td>6.5</td>
<td>X X</td>
</tr>
<tr>
<td>Multiplier</td>
<td>6.5</td>
<td>7.5</td>
<td>1 6</td>
</tr>
<tr>
<td>Calculation</td>
<td>7.5</td>
<td>17.5</td>
<td>X X</td>
</tr>
<tr>
<td>Response</td>
<td>17.5</td>
<td>when pressing enter key</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Instructions to participants. Before the experiment started, the participants were requested to position themselves in front of the monitor with their chin leaning on the headrest. They were instructed to stay still and keep their gaze fixed and focus (not stare) at the center of the screen throughout a trial. In addition, participants were asked to blink as little as possible, obviously without causing irritation, and to start each trial with ‘a clear mind’ (i.e., not thinking about the previous trial). If the participants could not complete the multiplication in time, they were instructed to enter zero as their answer.

Data Processing. The data were processed in two steps. In the first step, the missing values in the pupil diameter data (lost during recording) were removed and the signals were repaired with linear interpolation (see Fig 2, left, for an illustration). On average, 1.2% of the data were lost, so this processing step did not significantly influence the results. Step two included the removal of the blinks and the poor-quality data. During a blink, the eyelid opening rapidly diminishes to zero and then increases in a few tenths of a second until it is fully open again. It is impossible to track the pupil diameter while blinking. These instances in time were removed from the data (for a detailed description of how the blinks were identified and removed, see Appendix C and D). The pupil diameter quality signal (provided by SmartEye software) was used to filter out the poor quality data. This signal ranges from 0 to 1, with values close to 1 indicating a good quality (SmartEye AB, 2013). All data points with a pupil diameter quality below 0.75 were removed. Trials containing less than 70% of the original data were excluded from the analysis. Of the initial 1350 trials from 30 participants, 1110 trials spread of 29 participants passed these criteria. The results of one participant (45 trials) were discarded completely. The gaps in the remaining trials were again filled using linear interpolation (see Fig 2, right), a process that does not substantially alter the data according to Beatty and Lucero-Wagoner (2000).

The last 0.4 seconds of the accommodation period (3.6–4 s) were defined as the pupillary baseline, as was done by Klingner (2010). The mean pupil diameter of the baseline period of each trial was subtracted from each trial to accommodate for any possible shifts or drifts. The mean pupil diameter change (MPDC) for each participant was then obtained by averaging all trials per level of difficulty. Similarly, the mean pupil diameter (MPD) for each participant was obtained but then without subtracting the mean pupil diameter of the baseline period. The MPDCR was
calculated for each participant as the average velocity (mm/sample) or change in MPD between two points in time. In order to compare the three difficulty levels, the MPD and MPDC were analyzed at eight fixed points in time from the multiplier and calculation periods. Both measures were reported such that a complete picture of the pupillary behavior could be given (Beatty & Lucero-Wagoner, 2000). The MPDCR was assessed across the seven interim periods.

![Figure 2](image.png)

**Figure 2.** Example of the data processing steps. Left: Pupil diameter (PD) before and after linear interpolation for missing values. Right: PD before and after blink and poor quality data removal and linear interpolation.

In addition to these analyses, the mean blink rate (MBR) for two different periods in time was calculated and Pearson’s *r* correlation coefficients were obtained between the MPDC and the NASA-TLX and responses. Cohen’s *d* effect size (see Eq. 1) was calculated to determine at which points in time the differences in MPDC between the three levels of difficulty were largest.

\[
Cohen's\ d_i = \frac{|M_i - M_j|}{\sqrt{SD_i^2 + SD_j^2 - 2*r*SD_i*SD_j}}
\]  

**Statistical Analyses.** The pupil diameter measures (MPD, MPDC, and MPDCR), the blink rates (MBR), and the results of the NASA-TLX questionnaire were analyzed with a one-way repeated measures ANOVA. Tukey’s honest significant difference test was used with a significance level of 0.05 to determine whether pairs of conditions were significantly different from each other. To determine whether the Pearson correlation coefficients were significantly different from zero, a Bonferroni correction was applied. Thus, because 24 correlation coefficients were calculated (8 points in time * 3 levels of difficulty), the significance level was reduced to 0.002 (0.05/24).

1.4. **Results**

**MPD.** The mean pupillary response during the mental multiplication task of 29 participants is shown in Figure 3a. It can be seen that the MPD was higher for the higher levels of difficulty at all points in time. The pattern of the MPD was similar for all levels during the first ten seconds. Hereafter, the response seems different for each level and was split for further analysis in seven periods with eight points (see Fig. 3b). The points are indicated by a ‘P’ and the numbers of the periods are shown in parentheses.
Figure 3a. Mean pupil diameter (MPD) during the mental multiplication task of 29 participants, for the three levels of difficulty. The grey bars represent the periods where the multiplicand and multiplier were shown on the screen. The numbers were masked by an “XX” during the rest of the trial.

Figure 3b. Mean pupil diameter (MPD) during the presentation of the multiplier and the calculation period of 29 participants, for the three levels of difficulty. The seven periods are indicated in parenthesis.

The means and standard deviations of the MPD for the eight points in time and three levels of difficulty are shown in Table 2, together with the effect sizes and the p-values of the one-way repeated measures ANOVA and the pairwise comparisons. The results confirm that the MPD was significantly higher for the more difficult levels at all points in time and between most of the conditions.

Table 2. Mean Pupil Diameter Change (MPDC), Mean Pupil Diameter Change Rate (MPDCR), NASA-TLX, and Mean Blink Rate (MBR). The means (M) and standard deviations (SD) of 29 participants are shown per level of difficulty of the multiplications. P1-P8 refers to the eight points in time, while (1)-(7) refers to the seven periods. Statistically significant differences are indicated in boldface. The bar graphs of the MPD and MPDC can be found in Appendix E.

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>p-value</th>
<th>Effect size</th>
<th>Pairwise comparison of levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>(SD)</td>
<td>M</td>
<td>(SD)</td>
<td>M</td>
<td>(SD)</td>
</tr>
<tr>
<td>MPD (mm) (N = 29)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>3.770</td>
<td>3.804</td>
<td>3.881</td>
<td>0.004</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>(0.456)</td>
<td>(0.467)</td>
<td>(0.490)</td>
<td></td>
<td>0.555</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>P2</td>
<td>3.814</td>
<td>3.865</td>
<td>3.954</td>
<td>1.94*10^-4</td>
<td>0.26</td>
</tr>
<tr>
<td>P3</td>
<td>3.919</td>
<td>3.979</td>
<td>4.061</td>
<td>0.001</td>
<td>0.22</td>
</tr>
<tr>
<td>P4</td>
<td>3.902</td>
<td>4.003</td>
<td>4.116</td>
<td>2.02*10^-5</td>
<td>0.32</td>
</tr>
<tr>
<td>P5</td>
<td>3.836</td>
<td>3.949</td>
<td>4.140</td>
<td>7.14*10^-9</td>
<td>0.49</td>
</tr>
<tr>
<td>P6</td>
<td>3.767</td>
<td>3.894</td>
<td>4.127</td>
<td>1.98*10^-9</td>
<td>0.51</td>
</tr>
<tr>
<td>P7</td>
<td>3.720</td>
<td>3.815</td>
<td>4.130</td>
<td>3.50*10^-12</td>
<td>0.61</td>
</tr>
<tr>
<td>P8</td>
<td>3.693</td>
<td>3.781</td>
<td>4.114</td>
<td>1.03*10^-12</td>
<td>0.63</td>
</tr>
</tbody>
</table>

**MPDC (mm) (N = 29)**

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>-0.001</td>
<td>0.004</td>
<td>0.024</td>
<td>0.474</td>
<td>0.03</td>
<td>0.977</td>
</tr>
<tr>
<td>P2</td>
<td>0.043</td>
<td>0.065</td>
<td>0.097</td>
<td>0.064</td>
<td>0.09</td>
<td>0.583</td>
</tr>
<tr>
<td>P3</td>
<td>0.148</td>
<td>0.179</td>
<td>0.203</td>
<td>0.178</td>
<td>0.06</td>
<td>0.548</td>
</tr>
<tr>
<td>P4</td>
<td>0.131</td>
<td>0.203</td>
<td>0.259</td>
<td>0.001</td>
<td>0.21</td>
<td>0.085</td>
</tr>
<tr>
<td>P5</td>
<td>0.064</td>
<td>0.148</td>
<td>0.282</td>
<td>7.26*10^-8</td>
<td>0.44</td>
<td>0.036</td>
</tr>
<tr>
<td>P6</td>
<td>-0.005</td>
<td>0.094</td>
<td>0.270</td>
<td>1.54*10^-9</td>
<td>0.52</td>
<td>0.022</td>
</tr>
<tr>
<td>P7</td>
<td>-0.051</td>
<td>0.015</td>
<td>0.273</td>
<td>6.52*10^-14</td>
<td>0.66</td>
<td>0.116</td>
</tr>
<tr>
<td>P8</td>
<td>-0.078</td>
<td>-0.018</td>
<td>0.259</td>
<td>1.72*10^-12</td>
<td>0.62</td>
<td>0.251</td>
</tr>
</tbody>
</table>

**MPDCR (mm/sample) (N = 29) (x 1*10**^0**)**

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>0.361</td>
<td>0.513</td>
<td>0.611</td>
<td>0.143</td>
<td>0.07</td>
<td>0.450</td>
</tr>
<tr>
<td>(2)</td>
<td>0.586</td>
<td>0.632</td>
<td>0.592</td>
<td>0.902</td>
<td>0.00</td>
<td>0.909</td>
</tr>
<tr>
<td>(3)</td>
<td>-0.094</td>
<td>0.134</td>
<td>0.309</td>
<td>0.006</td>
<td>0.17</td>
<td>0.150</td>
</tr>
<tr>
<td>(4)</td>
<td>-0.371</td>
<td>-0.305</td>
<td>0.130</td>
<td>3.67*10^-5</td>
<td>0.31</td>
<td>0.820</td>
</tr>
<tr>
<td>(5)</td>
<td>-0.383</td>
<td>-0.302</td>
<td>-0.070</td>
<td>0.044</td>
<td>0.11</td>
<td>0.797</td>
</tr>
<tr>
<td>(6)</td>
<td>-0.257</td>
<td>-0.433</td>
<td>0.017</td>
<td>4.96*10^-4</td>
<td>0.24</td>
<td>0.235</td>
</tr>
<tr>
<td>(7)</td>
<td>-0.152</td>
<td>-0.184</td>
<td>-0.080</td>
<td>0.694</td>
<td>0.01</td>
<td>0.964</td>
</tr>
</tbody>
</table>

**NASA-TLX (%) (N = 30)**

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>20.744</td>
<td>30.883</td>
<td>48.658</td>
<td>1.65*10^-16</td>
<td>0.71</td>
<td>1.80*10^-4</td>
</tr>
<tr>
<td>Mental</td>
<td>33.833</td>
<td>46.833</td>
<td>70.180</td>
<td>1.67*10^-12</td>
<td>0.61</td>
<td>0.004</td>
</tr>
<tr>
<td>Physical</td>
<td>16.000</td>
<td>19.000</td>
<td>19.833</td>
<td>0.152</td>
<td>0.06</td>
<td>0.314</td>
</tr>
<tr>
<td>Temporal</td>
<td>18.667</td>
<td>29.167</td>
<td>53.167</td>
<td>3.67*10^-12</td>
<td>0.60</td>
<td>0.021</td>
</tr>
<tr>
<td>Performance</td>
<td>10.033</td>
<td>20.667</td>
<td>40.433</td>
<td>8.09*10^-11</td>
<td>0.55</td>
<td>0.014</td>
</tr>
<tr>
<td>Effort</td>
<td>28.000</td>
<td>43.133</td>
<td>63.500</td>
<td>9.89*10^-12</td>
<td>0.58</td>
<td>9.37*10^-4</td>
</tr>
<tr>
<td>Frustration</td>
<td>17.933</td>
<td>26.500</td>
<td>44.833</td>
<td>2.51*10^-9</td>
<td>0.49</td>
<td>0.057</td>
</tr>
</tbody>
</table>
MPDC. Figures 4a shows the MPDC of 29 participants as a function of the level of difficulty. As mentioned above, this measure takes into account the shift of the baseline by subtracting the mean of the baseline period of each trial. The difference between the three pupillary responses during the calculation period can now be seen more clearly. Again, the multiplier and calculation were split into seven periods by eight points (see Fig. 4b).

![Graph showing MPDC for different levels of difficulty](image)

**Figure 4a.** Mean pupil diameter change (MPDC) during the mental multiplication task of 29 participants, for the three levels of difficulty. The grey bars represent the periods where the multiplicand and multiplier were shown on the screen. The numbers were masked by an “XX” during the rest of the trial.

![Graph showing MPDC during multiplier and calculation period](image)

**Figure 4b.** Mean pupil diameter change (MPDC) during the presentation of the multiplier and the calculation period of 29 participants, for the three levels of difficulty.

The results of the analysis of the MPDC at the eight points in time and three levels of difficulty are shown in Table 2. It shows that a significant difference occurred at points 4 to 8 and that the effect size was largest at point 7.
A scatterplot of the MPDC at points 1, 5 and 8 of Level 1 versus Level 3 gives insight into the differences between individuals (see Fig. 5). The MPDC of Level 3 lies above the unity line for 16, 28, and 29 of the 29 participants for the three points respectively, and has a range of about 1 mm.

Figure 5. Scatterplot of the mean pupil diameter change (MPDC; blue dots) of 29 participants at point 5 of Levels 1 and 3. Also depicted is the unity line (solid black).

**MPDCR.** Figure 6 shows the MPDCR of the 29 participants as a function of the difficulty level, for the seven periods. A positive value indicates overall pupil dilation during that period and a negative value means overall contraction of the pupil diameter. In the first two periods, the diameter increased with approximately equal velocity for the three levels. During the other periods, the velocities decreased and became negative. Significant differences were found between the three conditions (see also Table 2).

Figure 6. Mean pupil diameter change rate (MPDCR) of 29 participants as a function of difficulty level, for seven periods in time during the presentation of the multiplier and the calculation period. The asterisks indicate significant differences between the three levels of difficulty.

**NASA-TLX.** The results of the NASA-TLX questionnaire are shown in Figure 7. For almost all items, the TLX score was significantly higher for the more difficult multiplications (see also Table 2). Only the subjective physical workload did not differ significantly across the levels of difficulty.
Responses. The percentage correct responses for Levels 1, 2, and 3 were respectively 94.2%, 93.8%, and 69.2%. Figure 8 shows the MPD of 29 participants for Level 3 of all trials, and separated for correct and incorrect responses. Too few incorrect answers were given for the other two levels and the results for these levels are therefore not reported. The MPD of the incorrect responses shows the same pattern as the one of the correct responses for the first twelve seconds. From this moment onward, the MPD belonging to the trials with incorrect responses was higher. A significant difference was observed at point 2 and 8 between the two lines when the same eight-point analysis was used (see Appendix F).

Effect size. The effect size estimate Cohen’s $d_z$ was calculated for the MPDC between pairs of difficulty levels for every point in time. Figure 9 shows the results. Large effect sizes arose after approximately 11 seconds since the start of the trial, especially between Levels 1 and 3.
Correlations. The results of the correlation analyses between the MPDC, NASA-TLX, and proportion of incorrect responses are shown in Table 3. For the MPDC, the table shows overall positive correlations, for the eight points in time and for the three different levels of difficulty. Between the MPDC and the percentage of incorrect responses, two statistically significant positive correlation coefficients were observed at points 1 and 2. Furthermore, Table 3 shows that people who experienced higher subjective workload (i.e., a higher NASA-TLX score) generally gave more incorrect responses.

Table 3. Pearson’s r correlations between the mean pupil diameter change (MPDC), percentage of incorrect responses, and the overall NASA-TLX scores, for the three levels of difficulty. Statistically significant correlations are indicated in boldface.

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r (p-value)</td>
<td>r (p-value)</td>
<td>r (p-value)</td>
<td>r (p-value)</td>
</tr>
<tr>
<td><strong>MPDC vs. Overall NASA-TLX (N = 29)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>-0.009 (0.961)</td>
<td>0.195 (0.310)</td>
<td>0.201 (0.296)</td>
<td>0.355 (0.059)</td>
</tr>
<tr>
<td>P2</td>
<td>-0.131 (0.498)</td>
<td>0.288 (0.130)</td>
<td>0.079 (0.685)</td>
<td>0.247 (0.195)</td>
</tr>
<tr>
<td>P3</td>
<td>-0.035 (0.857)</td>
<td>0.045 (0.818)</td>
<td>0.009 (0.964)</td>
<td>0.040 (0.836)</td>
</tr>
<tr>
<td>P4</td>
<td>0.303 (0.109)</td>
<td>0.066 (0.733)</td>
<td>0.030 (0.878)</td>
<td>0.272 (0.153)</td>
</tr>
<tr>
<td>P5</td>
<td>0.243 (0.204)</td>
<td>0.115 (0.554)</td>
<td>0.010 (0.956)</td>
<td>0.168 (0.384)</td>
</tr>
<tr>
<td>P6</td>
<td>0.211 (0.272)</td>
<td>0.196 (0.307)</td>
<td>-0.016 (0.934)</td>
<td>0.139 (0.472)</td>
</tr>
<tr>
<td>P7</td>
<td>0.175 (0.363)</td>
<td>0.203 (0.290)</td>
<td>0.163 (0.397)</td>
<td>0.226 (0.238)</td>
</tr>
<tr>
<td>P8</td>
<td>0.056 (0.766)</td>
<td>0.258 (0.176)</td>
<td>0.163 (0.399)</td>
<td>0.215 (0.262)</td>
</tr>
<tr>
<td><strong>MPDC vs. % Incorrect responses (N = 29)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>0.353 (0.060)</td>
<td>0.438 (0.017)</td>
<td>0.349 (0.063)</td>
<td>0.643 (1.70*10⁻⁴)</td>
</tr>
<tr>
<td>P2</td>
<td>0.228 (0.233)</td>
<td>0.505 (0.005)</td>
<td>0.264 (0.166)</td>
<td>0.561 (0.002)</td>
</tr>
<tr>
<td>P3</td>
<td>0.069 (0.722)</td>
<td>0.256 (0.180)</td>
<td>0.130 (0.500)</td>
<td>0.196 (0.309)</td>
</tr>
<tr>
<td>P4</td>
<td>0.306 (0.106)</td>
<td>0.254 (0.183)</td>
<td>0.122 (0.528)</td>
<td>0.312 (0.099)</td>
</tr>
<tr>
<td>P5</td>
<td>0.232 (0.224)</td>
<td>0.159 (0.409)</td>
<td>0.027 (0.887)</td>
<td>0.199 (0.302)</td>
</tr>
<tr>
<td>P6</td>
<td>0.064 (0.740)</td>
<td>0.205 (0.285)</td>
<td>0.016 (0.932)</td>
<td>0.123 (0.525)</td>
</tr>
<tr>
<td>P7</td>
<td>0.048 (0.803)</td>
<td>0.321 (0.090)</td>
<td>0.087 (0.653)</td>
<td>0.226 (0.238)</td>
</tr>
<tr>
<td>P8</td>
<td>0.063 (0.744)</td>
<td>0.249 (0.193)</td>
<td>0.137 (0.477)</td>
<td>0.218 (0.255)</td>
</tr>
</tbody>
</table>

Overall NASA-TLX vs. % Incorrect responses (N = 30)
Blinks. Figure 10 shows the MBR of all participants and sorted per level of difficulty during a period with low (2–6.5 s) and high (6.5–13 s) mental demands. The MBR of Level 3 during the second period was significantly higher than those of Level 1 and 2. More details can be found in Table 2.

Figure 10. Mean blink rate (MBR) of 30 participants during a period with low and high mental demands, for three levels of difficulty.

1.5. Discussion

Main results. The results showed that the overall MPD was higher for the higher levels of difficulty. Points 7 and 8 showed the largest differences. These findings demonstrate that the mean or baseline of the pupil diameter can shift during mental activity. If the pupil was given more time to recover from the previous trial, by increasing the length of the accommodation period, the difference of the MPD between the three levels of difficulty in the first period would probably have been smaller.

A remarkable finding is the behavior of the MPD during the first three seconds of the accommodation period (0–3 s). Where a clear decline from the start or a low horizontal line might be expected, the MPD starts to decline only after three seconds. This unexpected effect may have been caused by the fact that participants looked away from the center of the screen when their outcome to the multiplication had to be entered. Although the responses were not given during the accommodation period, the fluctuation could be an aftereffect because the trials came in relatively quick succession. During the presentation of the multiplicand and the pause (4–6.5 s) the MPD decreased further, at a slower pace however, which seems to indicate memory load (cf. Kahneman & Beatty, 1966). This small increase of the pupil diameter after the presentation of the first number was also observed by Ahern (1978) and Klingner (2010).

What is notable in the MPDC figure (Fig. 4) is that the pupillary behavior among the three difficulty levels was highly similar during the first few seconds after the presentation of the multiplier (6.5–9 s). This might be due to the strategy that the participants used. One can imagine that the first step in each multiplication, regardless of its difficulty, is similar. For example, the first step for many people of the Level 1 multiplication 7x14 would probably be 7x10. This is comparable to the first step of the Level 3 multiplication 14x18, which would then be 14x10. These observations are in line with the TEPRs obtained by Ahern (1978). She also observed a similar response among the three levels of difficulty at the beginning of the calculation. The MPDC
during the other periods was found to differ significantly between the three levels, particularly when Levels 1 and 2 were compared to Level 3. This finding is in accordance with the results in the scatterplot (Fig. 5), where 28 and 29 of the 29 participants had a higher MPDC for Level 3 than for Level 1, for points 5 and 8, respectively.

The results of the MPDCR illustrate that the effect sizes are smaller when compared to the results of the MPDC measure. It does provide, however, a clear understanding of when the muscles of the pupil relax and hence when the mental workload decreases.

According to the results of the NASA-TLX questionnaire, the classification of the arithmetic tasks was done properly, since a statistically significant difference was found in the subjective mental workload across all three levels. The big contrast between the subjective mental and physical workload underlines that the task was predominantly mentally demanding. Not to be overlooked are the roles of the subjective temporal demand and frustration. Looking at the increase of the MPD of the incorrect responses after 12 seconds for Level 3 (Fig. 8), it is plausible that, although only one significant difference was found, this increase was caused by the time pressure of the task or the frustration of not having solved the multiplication yet, instead of increased task demands.

Another interesting question related to Figure 8 showing the trials with the correct versus incorrect responses is: were the participants really trying to complete the task or did they give up on the task because it was too difficult? If the latter were the case, one would expect an early decline of the MPD. But the opposite is true, instead. A small increase of the MPD was measured, suggesting that the participants were trying hard to complete the task.

At the first two points in particular, moderate to strong correlations were found between the MPDC and the proportion of incorrect responses. A similar but weaker effect was obtained between the MPDC and the NASA-TLX. It may not be surprising that the strongest correlations were found at points 1 and 2, considering the fact that at these points in time probably all participants were still calculating. Once the task has been completed, the pupil diameter decreases again (cf. Kahneman, 1966 for similar findings in a memory paradigm). Since this decline does not occur at the same time for each trial, this causes higher variability and lower correlation coefficients. Apart from that, the results seem to indicate that the MPDC was higher for participants who gave more incorrect responses and experienced a higher workload. This could help in determining the feasibility of using the pupil diameter in adaptive automation. Combining the pupil diameter with other assessment methods could help increase validity and robustness. Correlations of similar size between the pupil diameter and proportion of incorrect responses and NASA-TLX were respectively found by Payne et al. (1986) and Recarte et al. (2008).

The relation between mental workload and blink rate has been unclear (Kramer, 1990; Recarte et al., 2008; Marquart et al., 2015). The results in the present study show that the MBR is significantly higher for Level 3 than for Level 1 and 2 during mentally demanding periods. However, the differences between Level 1 and 2 and the two periods in time are small. The MBR therefore appears to be less sensitive than the MPDC and more suited for the detection of a task’s overall mental workload, because of its low temporal resolution.

Conclusions and recommendations. It is concluded that the results of Ahern (1978) and Klingner (2010) have been accurately replicated with the SmartEye DR120 remote eye tracker. The partial eta squared effect sizes ($\eta_p^2$) for point 7 and 8 of the MPD, MPDC, and NASA-TLX are approximately the same (~0.6), which demonstrates that pupil diameter measurements can be just as valid as the NASA-TLX. An attempt was made to provide more insight into the individual
differences of TEPRs by means of a correlation analysis. Results showed a few moderate to strong correlations at the beginning of the calculation period between the MPDC and the NASA-TLX, on the one hand, and the ratio of incorrect responses, on the other.

Thus, it seems possible to assess workload by tracking the pupil diameter. However, the validity of pupil diameter measurements may need improvement before it could be implemented in practice. One possible way to do this is by combining pupillometry with other physiological measures, such as blink and heart rate (Kahneman et al., 1969; Molen et al., 1989; Just et al., 2003; Satterthwaite et al., 2007; Haapalainen et al., 2010). Additionally, future research could focus on improving signal analysis techniques that filter out effects other than mental workload, such as the light reflex.

The supplementary materials provide all measurement data, software, and scripts that would allow others to reproduce these results.

1.6. References


SmartEye AB. (2013). Programmer’s Guide, Revision 1.3. support@smarteye.se.


Wilhelm, T. (2010). Accuracy and precision of the pupil size measured with SmartEye Pro 5.6.0.
2

Preliminary Study 1
Mental Multiplication

2.1. Introduction

The goal of this experiment was to find out whether or not the SmartEye DR120 remote eye tracker can detect the task-evoked pupillary response (TEPR).

2.2. Setup

For this experiment, I recorded my own pupil diameter and eyelid opening during 20 trials of mental multiplication with the SmartEye DR120 remote eye tracker, which has a sampling rate of 120 Hz. The experiment took place in a room where there was office lighting and where daylight could not enter. The brightness of the screen was kept constant during the experiment.

Each trial started with a five second baseline condition which consisted of looking at a small black bar (5 x 2 mm) in the center of the screen. I was then presented with two random numbers between 12 and 19 for one second with a one-second pause in between. During the five seconds after the second number was presented, I had to calculate the product of the two numbers. Hereafter, the right answer was shown on the screen for one second. So in total, it took 14 seconds to complete one trial. The next trial started immediately after the previous one had been completed.

2.3. Data Processing

During a blink, the eyelid opening rapidly diminishes to zero and then increases in a few tenths of a second until it is fully open again (see Figure 1, solid blue line). It is impossible to track the pupil’s diameter while blinking. These instances in time should therefore be removed from the data. The recordings of the eyelid opening were used to identify the blinks in the pupil diameter data. A threshold of 12 mm was used to make a clear distinction between blinks and no blinks as depicted in Figure 1 by the dashed red line.
As can be seen in Figure 1, it takes some time to cross the threshold and the blink has not been completed after the eyelid opening signal crossed the threshold line for the second time. That is why 12 additional data points (~0.1 s) are removed from the data before the blink and 36 additional data points (~0.3 s) after the blink. The gaps in the data that are left by removing the blinks were restored by linear interpolation (see Figure 2, for an example), a process that does not substantially alter the data according to Beatty & Lucero-Wagoner (2000).

2.4. Results

The mean TEPR was obtained by averaging all 20 trials. The result (red) was plotted in Figure 3a together with all the trials (black) and the overall mean of the baseline period (blue), which was plotted for the entire 14 seconds in order to make the increase in pupil diameter clearer. Right after the multiplicand was presented, the pupil size increased a little (5–6.5 s) and even more when the multiplier was given and the multiplication has to be computed (7–13 s). A similar experiment of mental multiplication was conducted by Ahern and Beatty (1979) and Klingner (2008, 2010) and they found similar results (see Appendix G).
A visual inspection of the quality signal of the eyelid opening data showed some remarkable behavior after approximately 200 seconds (see Figure 4). This quality signal was therefore used to filter out the poor data points. It ranges from 0 to 1, with numbers close to 1 indicating a good quality. All data points with an eyelid opening quality below 0.95 were removed from the data. Additionally, trials with less than 75% of good data were excluded from the analysis. The mean of the baseline of each trial was subtracted from each trial to accommodate for any possible shifts or drifts of the baseline. The first second of the baseline period was excluded from the calculation of the mean to allow the pupil to recover from the previous trial.

The TEPR becomes more apparent and the range of the data is clearly much smaller after using the above data processing steps (see Figure 5). Sixteen trials have passed the processing criteria.
The vertical blue lines reflect the maximum mean amplitude in the baseline, memory and calculation period and are equal to 0.02, 0.08 and 0.19 mm respectively.

**Figure 5.** TEPR for the first twelve mental multiplication trials. With the recovery (0-1 s), new baseline (2-5 s), multiplicand (5-6 s), memory (6-7 s), multiplier (7-8 s), calculation (8-13 s), and correct answer (13-14 s).

### 2.5. Conclusion

Comparing the TEPR of figure 5 with the results obtained by Klingner (2010) as mentioned earlier shows a lot of similarities in timing, duration, shape and amplitude (see Appendix G). The only difference between the two experiments is that Klingner needed 165 trials, while in this experiment only 16 trials have shown to give a similar result. It can therefore be concluded that TEPR’s can be measured using the SmartEye DR120 remote eye tracker.

### 2.6. References


Preliminary Study 2
Pupil Light Reflex Suppression by Variable Screen Brightness

3.1. Introduction

The goal of this experiment was to quantify the effect of gaze direction and light on the pupil diameter and create a screen background with a variable brightness that could suppress these effects.

3.2. Setup

For this experiment, I recorded my own pupil diameter and eyelid opening during 10 trials per experimental condition with the SmartEye DR120 remote eye tracker, which has a sampling rate of 120 Hz. The experiment took place in a room where there was office lighting and where daylight could not enter.

During each trial I focused on a black dot on the screen. This dot shifted to the next position (nine in total) every 1 or 5 seconds. I conducted this experiment with two different backgrounds (equal brightness and variable brightness, see Figure 1a and b). The order in which the stimuli were presented is indicated in the figure by the numbers in the white circles (The circles and numbers were not present during the experiment).

Figure 1a. Background condition I.
3.3. Data Processing

The blinks were removed from the raw data in the same way as in the first preliminary study: mental multiplication.

The eyelid opening quality (EOQ) signal turned out to be a good quality indicator and was therefore used to filter out the poor data points. This signal ranges from 0 to 1, with numbers close to 1 indicating a good quality. All data points with an EOQ below 0.95 were removed from the data. Additionally, trials with less than 75% of good data were excluded from the analysis. The mean of each trial was subtracted from each trial to accommodate for possible shifts or drifts of the pupil diameter baseline.

3.4. Results

The results of the experiment during condition I with a fixation time per target of 1s are plotted in Figure 2. Nine trials have passed the processing criteria. The maximum absolute mean pupil dilation of 0.22 mm can be found at fixation point 4.
The results of the experiment during condition II with a fixation time per target of 1s are plotted in Figure 3. Eight trials have passed the processing criteria. The maximum absolute mean dilation is 0.12 mm at fixation point 2.

![Figure 3](image3.png)

*Figure 3. Results condition II with fixation time of 1s.*

The results of the experiment during condition II with a fixation time per target of 5s are plotted in Figure 4. Eight trials have passed the processing criteria. The maximum absolute mean dilation is 0.29 mm at fixation point 2.

![Figure 4](image4.png)

*Figure 4. Results condition II with fixation time of 5s.*

3.5. **Conclusion**

Comparing Figure 2 and 3 shows a large reduction of the pupillary light reflex for the two different conditions. So, using a variable screen brightness seems to be a way to suppress the effect of light and gaze direction on the pupil diameter. However, the results also show that the pupillary response for longer fixation durations is still substantial, as shown in Figure 4. Of course, more research is required to be able to link the responses to a specific cause. But this at least gives an idea of the magnitude of the different effects and should therefore be taken into account during the design of the full-scale experiment.
4.1. Introduction

The goal of this pilot study was to test and improve the experimental design of the full-scale experiment. The purpose of the full-scale study is to investigate whether the pupil diameter is a good indicator of mental workload. Klingner (2010) and Ahern & Beatty (1979) investigated this question for mental arithmetic tasks and found that the task-evoked pupillary response (TEPR) can be used to distinguish between different levels of mental workload.

4.2. Setup

Six individuals participated in this experiment and were requested to do a series of mental arithmetic tasks (multiplications) while sitting behind a desktop computer and looking at a fixed point on the screen. Their pupil diameter and eyelid opening were recorded with the SmartEye DR120 remote eye tracker, which has a sampling rate of 120 Hz. The experiment took place in a room where there was office lighting and where daylight could not enter. The brightness of the screen was kept constant during the experiment, except during the administration of the NASA-TLX questionnaires.

The experiment consisted of 10 sessions (2 sessions x 5 levels of difficulty). Each session was of random difficulty and contained 5 trials. So in total each participant made 50 calculations. Each trial started with a 5 second accommodation period, which gave the pupils time to recover from the previous trial. The last 0.4 seconds of the accommodation period were defined as the pupillary baseline, as was used by Klingner (2010). The accommodation period was followed by an auditory presentation of two numbers between 5 and 24, with a 1.5 second pause in between. A more detailed description of how the arithmetic tasks were classified into five levels of difficulty can be found in Appendix H. These auditory presentations were computer-controlled and varied in duration, ranging from 0.52 (eleven) to 1.09 (twenty-three) seconds with a mean of 0.79 seconds. The participants were asked to multiply the two numbers as quickly and accurately as possible and type their answer on the numeric keypad of the keyboard. No time limit was imposed on the participants during the calculation. After each session, the participants were asked to fill out a NASA-TLX questionnaire to assess their subjective mental workload.

In addition to the eye tracking recordings and the questionnaire, the participants’ response times were registered. The response time was defined as the elapsed time between the presentation of
the last stimuli and the first button press of the participant. It gives an indication of the difficulty of the mental arithmetic task, and could therefore be used to approve the design of the five levels of difficulty.

4.3. Data Processing

During a blink, the eyelid opening rapidly diminishes to zero and then increases in a few tenths of a second until it is fully open again. It is impossible to track the pupil’s diameter while blinking. These instances in time should therefore be removed from the data. This was done in the same way as was done in the chapter “Preliminary Study: Mental Multiplication Experiment”. The gaps were then filled using linear interpolation.

The mean pupil diameter of the baseline period of each trial was subtracted from each trial to accommodate for any possible shifts or drifts of the baseline.

4.4. Results

Pupil Diameter. The mean pupillary response during the mental multiplication task for all trials and all levels of difficulty is shown in Figure 1. The grey bars represent the moments in time in which the multiplicand and the multiplier were aurally presented. The accommodation period (0-5 s) shows a clear drop of the diameter, as was expected and also accounted for to allow the pupil to recover from the previous trial. In between the presentation of the two numbers, the pupil diameter increases a little (~0.05 mm) and even more when the multiplication is calculated (~0.35 mm).

![Figure 1](image1.png)

*Figure 1.* The TEPR during the mental multiplication task for all trials for all levels of difficulty.

Figure 2 shows the mean TEPR during the mental multiplication task but now plotted per level of difficulty. All levels showed similar pupillary behaviour during the first eight seconds: a decline in the accommodation period and a small increase of the pupil diameter after the presentation of the multiplicand. The pupil dilation in the calculation period was less consistent and a lot noisier. It
was expected that the largest dilation, or the highest mental workload, would be found for the most difficult multiplications (Level 5). However, Figure 2 shows that this was clearly not the case.

![Figure 2. The TEPRs during the mental multiplication tasks for all trials sorted per level of difficulty](image)

**NASA-TLX.** The results of the NASA-TLX questionnaire were analysed with a one-way repeated measures ANOVA. It shows a significant difference between the five conditions (F(4,20) = 16.558, p = 3.91×10⁻⁶). A multiple comparisons test shows that the subjective workload during Level 1 and 2 was significantly lower than during Level 3, 4, and 5 (see Table 1a and b).

**Table 1a. NASA-TLX subjective workload - Means and standard deviations (SD) of all trials sorted per level of difficulty.**

<table>
<thead>
<tr>
<th>Level</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>17.164</td>
<td>8.649</td>
</tr>
<tr>
<td>Level 2</td>
<td>25.903</td>
<td>14.612</td>
</tr>
<tr>
<td>Level 3</td>
<td>40.972</td>
<td>14.020</td>
</tr>
<tr>
<td>Level 4</td>
<td>37.917</td>
<td>17.027</td>
</tr>
<tr>
<td>Level 5</td>
<td>41.014</td>
<td>14.776</td>
</tr>
</tbody>
</table>

**Table 1b. NASA-TLX multiple comparisons test results.**

<table>
<thead>
<tr>
<th>Level (i)</th>
<th>Level (j)</th>
<th>Mean Diff. (i-j)</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>-8.736</td>
<td>.166</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>-23.806*</td>
<td>.000</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>-20.750*</td>
<td>.000</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>-23.847*</td>
<td>.000</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>-15.069*</td>
<td>.004</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>-12.014*</td>
<td>.029</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-15.111*</td>
<td>.004</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>3.056</td>
<td>.919</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>-0.042</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>-3.097</td>
<td>.915</td>
</tr>
</tbody>
</table>

*. The mean difference is significant at the .05 level.

**Response Time.** For the analysis of the response times, only the trials with the correct answers to the multiplications were included, which was done to ensure that the trials which the participants
gave up on were filtered out. One participant gave wrong answers to all multiplications of Level 5, so his results were completely excluded from this analysis. The response times were analysed with a one-way repeated measures ANOVA. It shows a significant difference between the five conditions ($F(4,16) = 20.332, p = 4.10 \times 10^{-6}$). A multiple comparisons test shows that the response time for Level 1 was significantly lower than for Level 3, 4, and 5 and the response time for Level 2 was significantly lower than for Level 4 and 5 (see Table 2a and b).

Table 2a. Response time (s) means and standard deviations (SD) of all trials sorted per level of difficulty.

<table>
<thead>
<tr>
<th>Level</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.294</td>
<td>0.936</td>
</tr>
<tr>
<td>2</td>
<td>4.375</td>
<td>1.219</td>
</tr>
<tr>
<td>3</td>
<td>7.898</td>
<td>2.707</td>
</tr>
<tr>
<td>4</td>
<td>10.744</td>
<td>3.846</td>
</tr>
<tr>
<td>5</td>
<td>10.200</td>
<td>2.962</td>
</tr>
</tbody>
</table>

Table 2b. Response time (s) multiple comparisons test results.

<table>
<thead>
<tr>
<th>Level (i)</th>
<th>Level (j)</th>
<th>Mean Diff. (i-j)</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>-2.081</td>
<td>.404</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-5.604*</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>-8.450*</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-7.906*</td>
<td>.000</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>-3.523</td>
<td>.050</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>-6.369*</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-5.825*</td>
<td>.001</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>-2.846</td>
<td>.147</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-2.301</td>
<td>.311</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>0.544</td>
<td>.988</td>
</tr>
</tbody>
</table>

* The mean difference is significant at the .05 level.

4.5. Conclusion

Pupil Diameter. The overall TEPR has shown its typical characteristics during each period (decline during accommodation, small increase during memory, and a large increase during the calculation). If the TEPR was split and sorted per level of difficulty, the dilation was less consistent. The dilation amplitude was not higher for more difficult calculations and the graphs appeared to contain more noise. This remarkable behaviour may have various causes besides the fact that only a limited number of individuals participated in this pilot study. First, two participants wore glasses and one participant had brown eyes. This made eye tracking and measuring the pupil diameter more difficult and therefore the outcome less reliable. Second, a good calibration and steady head position are key in remote pupillometry. Although the participants were requested not to move during the experiment, some of them may have displaced themselves making the results of the eye tracker less reliable. Third, since there was no time pressure in the execution of the task, the participants may have adjusted their strategy for the more difficult tasks to obtain a more comfortable level of mental workload. Where the first two causes could explain the increased level of noise, the third cause could certainly explain the lack of a monotonic trend.

NASA-TLX. The results of the NASA-TLX questionnaire show that Level 1 and 2 significantly provoked a lower subjective mental workload than Level 3, 4 and 5.

Response Time. The response time for Level 1 was significantly lower than for Level 3, 4, and 5 and the response time for Level 2 is significantly lower than for Level 4 and 5.
4.6. **Discussion**

Before the full-scale experiment can be conducted, the issues described above need to be resolved. For future measurements with the DR120 eye tracker, it is recommended to exclude individuals with glasses or dark coloured eyes from participation, as their eye tracking results are less reliable. The reliability can also be increased by using a headrest to avoid head displacements. In order to prevent participants from adjusting their strategy or controlling their own level of mental workload, the time pressure component can be added to the experiment at the expense of more errors in their performance and the loss the response time measure.

Looking at the means of the NASA-TLX questionnaire (Table 1a), it appears that only three levels of mental workload can be distinguished with this questionnaire for mental multiplication tasks ranging between 5 x 5 and 24 x 24. The response times also seem to give evidence that at least the first three levels are of increasing difficulty. On top of that, after the experiment one participant commented that he experienced only three levels of difficulty. So, if the pupil diameter measurements are to be compared with the response times and the results of the questionnaire, it is advised to reduce the number of levels of difficulty to three instead of five.

4.7. **References**


5.1. Abstract

Human errors in traffic accidents are often caused by mental underload or overload. The assessment of mental workload could be helpful in optimising driving tasks and, hence, to reduce the number of accidents. This review summarizes the knowledge obtained in older studies and discusses the results of more recently conducted studies on the relation between eye-related physiological parameters and drivers’ mental workload. Various different eye activity measures including blinks, fixations, and pupillometry have shown to be useful in estimating a driver’s mental workload, but there is no single method that works in all cases, since driving is a multi-dimensional task and all metrics have specific drawbacks. Advanced algorithms, such as the ‘index of cognitive activity’ applied on pupil diameter recordings, or a combination of multiple assessment methods, have been shown to increase validity.

5.2. Introduction

Human errors, such as misperceptions, information processing errors, and poor decision making, are involved in over 90% of all traffic accidents (Treat et al., 1977). These errors are often caused by hurry, low vigilance, stress, or other types of driver states that resemble an inappropriate (i.e., either too low or too high) ‘mental workload’. It is therefore of great importance for car manufacturers and in-vehicle technology developers to improve and optimise for levels of mental workload before they take their product to market. For example, advanced cruise control and navigation systems may require a lot of human-computer interaction and therefore have major effects on mental workload and thus impact driver behaviour.

Driving is not only a physical (e.g., applying force on steering wheel and pedals) but also a visual and a mental task. The visual and mental demands can both be high (e.g., heavy traffic on crossroads) and low (e.g., empty country road). The eyes of a driver are indispensable in performing visual tasks such as scanning the road, communicating with other road users, and monitoring in-vehicle devices. Not only visual tasks, but also mental tasks are important during driving, including such factors as understanding vehicle dynamics, making situation-dependent decisions, and judging time/space relationships. Visual and mental tasks, such as scanning a busy crossroad while simultaneously judging the time/space relationships of all other road users, are closely linked to each other. It is therefore of no surprise that many researchers use or study eye-
related parameters to assess drivers’ mental workload (Beatty & Lucero-Wagoner, 2000; Brookhuis & De Waard, 2010; De Waard, 1996; Kramer, 1990; Recarte & Nunes, 2000).

Probably the first extensive literature review on the topic of measuring drivers’ mental workload was conducted by De Waard (1996). Since then, much research has been done in this field, resulting in new insights and the development of new measurement techniques. Borghini et al. (2012) reviewed the literature papers related to the neurophysiological signal measurements of mental workload, fatigue, and drowsiness of aircraft pilots and car drivers. However, this particular review was limited to the electrooculography (EOG) measurement method and did not cover any other eye-related techniques. It may therefore be stated that there is a gap in the scientific literature regarding the most recent findings, results, and theories from recent eye-related workload research.

The aim of this review is to summarize the knowledge obtained in older studies and to discuss the results of more recently conducted studies about the relation between eye-related parameters and driver mental workload. The heart of the review consists of a description of the concept of driver mental workload followed by an analysis of this concept in relation with the eye-related parameters found in the literature. The results will be discussed and recommendations will be made at the end of this paper.

5.3. Drivers’ Mental Workload

Mental workload is a broad concept that has various definitions. The mental workload definitions by O’Donnel & Eggemeier (1986) and De Waard (1996) are well known and used by many researchers (Borghini et al., 2012; Brookhuis & De Waard, 2010; Di Stasi et al., 2009; Palinko et al., 2010). O’Donnel and Eggemeier (1986) define mental workload as the ratio between the task required mental demand and the operator's total mental capacity and De Waard (1996) defines it as the portion of the human’s limited information processing capacity that is used for task performance. The general agreement between the various definitions according to Recarte et al. (2008) is that “mental workload is seen as the result of an interaction between task demands and human characteristics, that is, separately, neither the task properties nor the human operator characteristics can explain mental workload”.

An operator’s mental workload can vary between very low (underload), resulting in boredom or fatigue, and very high (overload), resulting in stress. Both extreme levels of workload can adversely affect an operator’s performance. Mental workload has also been found to dynamically vary during task execution, which could help assistance systems to find appropriate moments for interruption (Iqbal et al., 2005). One solution to mental overload, if properly designed, is automation. Poorly designed automated systems could lead to situations characterized by the phrase “99% boredom, 1% terror”. This phrase refers to supervisory control situations in which an operator experiences boredom due to mental underload for most of the time and terror due to mental overload in case of emergencies. Mental underload should be of much greater concern than overload according to Young and Stanton (2002, see for a more extensive discussion on this topic), since it may be more difficult to detect in the sense that mental underload shows more variable physiological behaviour (e.g., stress, fatigue, boredom; Hancock & Parasuraman, 1992; Hancock & Verwey, 1997).

The mental workload in driving is more specifically defined by Boer (2005) as the effort required to maintain the driving state within a subjective safety zone. In this definition, the subjective safety zone depends on the driver’s needs in terms of costs and benefits. The assessment of
drivers’ mental workload could be helpful in improving and optimising driving tasks to reduce the number of accidents, which are largely attributable to the drivers themselves.

In the literature, researchers usually distinguish three categories of mental workload measures: performance, physiological, and subjective measures (O'Donnel & Eggemeier, 1986). Performance measures are used to determine how well an operator, or driver in this case, is performing one or more tasks. If driving a vehicle would be the primary task, then a secondary task could be used to determine the driver’s remaining mental capacity, assuming that both tasks use the same mental capacity pool. Primary task performance measures are often related to lateral and longitudinal vehicle control tasks. Examples of these measures are: standard deviation of the vehicle’s lateral position on the road, time delay in car following, and the number of lane departures. Secondary task performance measures can either be superficially or closely related to the driving task itself (see De Winter et al., 2014, for a conceptualization of the driver’s task). Examples of superficial secondary tasks are: peripheral detection, mental arithmetic, and memory tasks. On the other hand, tasks that are normally part of the driving task may be affected with workload changes, such as checking the mirror less often if the workload increases.

Subjective measures capture the driver’s own assessment of mental workload. Frequently used assessment tools are the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988) and the Rating Scale Mental Effort (RSME; Zijlstra & Van Doorn, 1985) questionnaires (De Winter et al., 2014). These assessment tools can easily be administered after an experiment and are therefore particularly useful to capture a task’s overall subjective mental workload. Although these tools seem to be not well suited to on-line measurements, Tattersall and Foord (1996) adopted the instantaneous self-assessment (ISA) method, originally designed for air traffic control, which can be used for on-line measurements. Statistically significant correlations were found in this study between the ISA scale and other mental workload ratings, such as heart-rate variability. However, the results in this study also indicated that the ISA method interferes with the driver’s performance, making it less useful.

Physiological measures record the driver’s physiological state during a drive and are therefore a direct measure of mental workload. Examples of these measures are: eye activity (e.g., pupil dilation, blink rate, and eye fixations), head motion, brain activity, heart rate, blood pressure, muscular activity, body movement and posture, endocrine response, and galvanic skin response (e.g., Desmond, 1997). Besides mental workload, these physiological measures are also influenced by other aspects of the driver’s mental and physical state (e.g., physical fatigue and circadian rhythm) and by environmental variables (e.g., illumination and temperature; Palinko et al., 2010). Most physiological measurements also require specialized equipment and technical expertise (Kramer, 1990). De Waard (1996) mentions as advantages of physiological measures that most of these measures can be collected continuously and relatively unobtrusively, due to miniaturisation of the measurement equipment. The focus in the remainder of this review will be laid on eye-related physiological measures.

5.4. Blinks

There are three fundamental types of eye blinks: reflexive, voluntary, and endogenous. Endogenous eye blinks are distinguished from other blinks by the absence of an identifiable eliciting stimulus (Stern et al., 1984). Stern et al. (1984) argued that endogenous blinks can reflect the mental workload induced by task demands. Different parameters characterising the endogenous eye blinks have been studied in relation with driver’s mental workload. The three most commonly used parameters are: blink rate, blink duration, and blink latency. The former two
parameters are self-explanatory and task-independent, whereas blink latency refers to the length of time between the occurrence of task relevant information and the next subsequent blink initiation. Lastly, another blink related parameter worth discussing is that of PERCLOS (percentage of eye closure). Results from the literature regarding these parameters will be discussed below.

Blink rate has been investigated in a series of studies and their results seem inconsistent (Heger, 1998; Kramer, 1990; Recarte et al., 2008). More than two decades ago, Kramer (1990) stated in his review that the results related to blink rate were mixed and that more research was required before blink rate could be used as a measure of mental workload. In a real car-driving study by Heger (1998) twelve individuals navigated through twelve test courses alternating between two directions of travel. In this experiment the mental workload was captured with a number of physiological measures, including blink rate. The results in this study showed that the eye blink rate decreased as the curves of the road became sharper. Ten years later Recarte et al. (2008) analysed, among other things, the eye blink rate in single- and dual-task (cognitive task plus visual search) conditions in a within-subjects experiment with 29 participants. It was found that the blink rate increased for all cognitive tasks (listening, talking, and calculating) when compared to the control condition. The researchers also found a decrease in eye blink rate for more visually demanding tasks when compared to less visually demanding tasks (see Appendix M, figure M1). Recarte et al. (2008) therefore concluded that “according to blink rate, visual and mental workload produce opposite effects: Blink inhibition for higher visual demand and increased blink rate for higher mental workload”. This theory might explain the mixed results obtained in an earlier study where visual and mental workload were not clearly differentiated, as in Heger’s (1998) study where the results show a blink rate decrease as the curves become sharper. The driver’s task in this experiment might be more visually demanding (scanning the curved road ahead) than mentally demanding (controlling the vehicle), which could explain the decrease in blink rate. Blink rate has also been investigated in relation with highly automated driving, a system that automates a vehicle’s motion in the lateral as well as the longitudinal direction. Systems like these are designed, inter alia, to decrease a driver’s mental workload. Remarkably enough however, several studies found a blink rate increase during highly automated driving compared to normal driving (Cha, 2003; Damböck et al., 2013; Merat et al., 2012). Another completely different approach with respect to blink rate was given by Stern et al. (1994). These authors stated that the increase in blink rate might be a better indicator of fatigue. Like blink rate, the results found in literature related to blink latency and blink duration also appear explainable in a consistent manner across various experimental research. Kramer (1990) provided multiple examples of studies in his review in which the latency of blinks increases with increasing mental task demands. The author stated that it is likely that eye blinks are postponed until sufficient visual information is extracted from the environment to perform the task as well as possible. Similar results of increasing blink latency for increasing mental task demands were found in studies conducted by Eggemeier et al. (1990) and Carmody (1994). Blink duration has also been found to be an indicator of both visual and mental workload. The studies mentioned in Kramer’s (1990) review all found a decrease in the participant’s blink duration for increasing task demands. The author therefore concluded that when operators are faced with either high visual or mental processing demands the blink duration decreases such that the available time could be used as efficient as possible to obtain and process the task-relevant information. In flight simulator studies researchers also report a decrease in blink duration as visual workload increases (Ahlstrom & Friedman-Berg, 2006; Veltman & Gaillard, 1996). In a more recent driving simulator study conducted by Benedetto et al. (2011), the authors examined the effects of in-vehicle information systems (IVIS) on eye blinks during driving. Fifteen individuals participated in this single- and dual-task study and performed as the primary task a lane change test (LCT) which was designed by Mattes (2003) and the International Organization for Standardization (ISO/DIS
This test allowed the researchers to estimate the driver’s secondary task demand. The IVIS task was used as a secondary task and required visual perception and manual response. Benedetto et al. (2011) assumed that this task uses the same mental capacity pool and will therefore increase the visual and mental processing demands. The results showed a blink duration inhibition for the dual-task (LCT and IVIS task) compared with the single-task condition (LCT). According to the authors this blink duration inhibition may occur to avoid visual information loss. This thought fits well with Kramer’s (1990) statements mentioned earlier that blinks are postponed until sufficient visual information is obtained.

PERCLOS is defined as the percentage of time that the eyelid covers 80% or more of the pupil (Wierwille et al., 1994). PERCLOS values have been found to positively correlate with increasing subjective sleepiness, performance decrements, and the number of lapses in a visual reaction time task (Dinges & Grace, 1998; Friedrichs & Yang, 2010; Kozak et al., 2005; Mallis, 1999). The results obtained by Mallis (1999) showed that PERCLOS correlated even better with performance lapses in the reaction task than the participants’ own ratings of sleepiness. Halverson et al. (2012) compared the use of several eye metrics, including PERCLOS, for classifying mental workload. The authors knew that PERCLOS was a popular well-verified metric of fatigue but thought it surprising that it had not yet been studied in relation with mental workload. Ten extensively trained individuals participated in his study by performing visual searches on simulated radar images while driving an autonomous vehicle in a simulated environment. The results showed that PERCLOS can be used to distinguish between two different levels of workload. Halverson et al. (2012) also state that this metric and the pupil diameter complement each other for the assessment of mental workload.

One disadvantage of measuring the endogenous eye blink activity is its sensitivity to factors other than the mental or visual processing demands. Annoying luminance and poor air quality may contaminate the results and reduce their validity. These factors might have a lesser impact in highly controlled environments, such as laboratories, but they will affect the measurements on real roads.

5.5. Fixations

Measures of eye fixations can be classified as either performance or physiological measures (De Waard, 1996). The task performance measures are often divided into two types: primary and secondary. For highly visual tasks eye fixations can be used as primary task performance measures. Measuring eye fixations during driving while occasionally looking at a navigation system is an example of a secondary task performance measure. But still, due to its measuring techniques (eye-tracking and EOG) measures of eye fixations are mostly considered as physiological measures.

Over the years, many parameters characterising eye fixations have been studied in relation with driver’s mental workload including: number of fixations, fixation duration, number of saccades, saccadic duration, saccadic amplitude, peak of saccadic velocity, and gaze distribution (Di Stasi et al., 2009; Fu et al., 2011; Recarte & Nunes, 2000, 2003; Reimer, 2009; Reimer et al., 2010; Underwood et al., 2011; Victor et al., 2005; Young & Mahfoud, 2008). Another eye fixation parameter that has been linked to mental workload is ‘dwell time’ and is defined as the total time of all saccades and fixations spent in one area (Maltz & Shinar, 1999). The dwell time has been observed to increase for increasing mental workload (May et al., 1990; Miura, 1990). However, this parameter will not be further discussed since no literature was found connecting it to drivers’ mental workload.
Fixation duration is the most extensively used metric and is generally believed to increase with increasing visual and mental task demands (Rayner & Morris, 1990; Recarte & Nunes, 2000). In the years after, it has also been investigated in a series of studies in relation with hazard perception as described in Underwood et al.’s (2011) review (Chapman & Underwood, 1998a, 1998b; Chapman et al., 2007; Crundall & Underwood, 1998; Konstantopoulos et al., 2010; Underwood et al., 2005). All these studies showed increased fixation durations during hazardous moments, indicating increased mental workload. Underwood et al. (2011) therefore stated that “Since long fixation durations are typically associated with high processing load, it makes sense to think that during these hazards viewers are spending longer extracting information from their point of gaze”.

Previous research demonstrated that increased attentional demand (sum of visual and mental demand) produces attentional focus narrowing, also called spatial gaze concentration/dispersion (Miura, 1986, 1990; Radach, 1998). Recarte and Nunes (2000) were one of the first to investigate the effect of mental instead of visual tasks on eye fixation parameters during driving in real traffic. They hypothesised that “if the narrowing effect is attentional, then it should also occur when the increment of attentional demands is due to concurrent cognitive tasks instead of to an increase of visual scene complexity”. Horizontal and vertical variability’s of fixation positions were measured in order to detect a narrowing effect. For this experiment, they used two types of secondary tasks: spatial-imagery (both mentally and visually demanding) and verbal (only mentally demanding). The results showed lower variability’s in both horizontal and vertical spatial gaze direction for both secondary tasks compared to the condition without a secondary task. This confirmed their hypothesis that verbal cognitive tasks produce attentional focus narrowing. In addition to these results, the authors found that the spatial-imagery task increased mean fixation duration and reduced saccadic amplitude during visual search behaviour more so than the verbal tasks. The increment in fixation duration is described by the authors as an eye freezing effect: long fixations caused by mental processes rather than external events.

Similar results of spatial gaze dispersions due to added cognitive workload were found in later studies (Reimer, 2009; Reimer et al., 2010; Victor et al., 2005). In the on-road study conducted by Reimer (2009) the horizontal and vertical eye gaze of 26 participants during driving were measured. The subjects were asked to perform several auditory cognitive secondary tasks of varying difficulty. The drivers’ gaze distributions were found to be significantly smaller while performing the additional secondary tasks, especially during the most difficult task. It is interesting to note that in this study the drivers seemed to change their visual search behaviour (smaller gaze distributions) before vehicle control suffered. More recent studies discussed (Reimer et al., 2012), and investigated, whether combining the horizontal and vertical gaze dispersion would produce a more sensitive measure of drivers’ mental workload (Wang et al., 2014). But this was not the case and so it was concluded that the horizontal gaze dispersion is the most sensitive measure. This result was not surprising given the fact that drivers more frequently scan the roadway environment than the dashboard or the rear-view mirror (Wang et al., 2014).

This effect on visual search behaviour was further investigated in another study by Recarte and Nunes (2003) in real traffic with 12 participants. One of the aims of this study was to figure out whether visual-detection impairment and spatial gaze concentration (induced by increased mental workload) equally affects the entire visual field or the peripheral areas in particular. The former concept of gaze dispersion is called general interference, and the latter is called tunnel vision. Eight different mental tasks were used to increase the driver’s mental workload. As expected, the tasks in this study also produced spatial gaze concentration and visual-detection impairment. This impairment was, according to the results, independent of target eccentricity and the authors
therefore concluded that a general interference effect was produced instead of a tunnel vision effect.

In a driving simulator study carried out by Young and Mahfoud (2008), the effects of roadside advertising on driver attention and performance were investigated. It was found that the mental workload and the total number of fixations increased when billboards were placed at the side of road. Despite finding these noticeable relations, no hard conclusions were drawn. Concerning the increase of the number of fixations, Young and Mahfoud (2008) argued that the drivers altered their visual attention strategies towards more but shorter eye fixations.

Di Stasi et al. (2009) studied the relation between motorcycle drivers’ risk behaviour and mental workload in a static motorcycle riding simulator. During the experiment, eye activity was recorded to estimate the mental workload. Sixty individuals participated in this study, which were assigned by the researchers either to a high-risk or a low-risk group. The results showed a fairly strong negative correlation of -0.61 in the high-risk group between the peak of saccadic velocity and the subjective mental workload. This means that the results suggest that a higher driver’s mental workload lowers the peak of saccadic velocity for high-risk individuals.

Fu et al. (2011) also studied the correlations between eye fixation parameters and driver’s mental workload. The researchers analysed 8 experienced and 8 novice drivers on a real urban segmented road (27 segments, differing in complexity). They collected four different eye fixation parameters (saccadic amplitude, average saccadic velocity, peak saccadic velocity, and fixation duration) together with three heart rate indicators of driver’s mental workload. Strong correlations were found between two of the saccade parameters and the workload parameters. The authors therefore concluded that the driver’s mental workload, saccadic velocity, and saccadic amplitude increase for increasing road complexity.

5.6. Pupillometry

Pupillometry is the measurement of the pupil diameter and is considered to be a reliable physiological measure of mental workload (Beatty, 1982; Beatty & Lucero-Wagoner, 2000; Kahneman, 1973). The size of the pupil is regulated by two antagonistic muscles in the iris, the sphincter and the dilator muscles (See Appendix N, figure N1). The sphincter muscle constricts the iris and is controlled by the parasympathetic nervous system (PSNS) located in the midbrain. The dilator muscle opens the iris and is innervated by the sympathetic nervous system (SNS) located in the hypothalamus. The PSNS and the SNS are both part of the autonomic nervous system (ANS), which acts like a control system and functions involuntary. The activation of the two muscles is determined primarily by light and accommodation reflexes, but the ANS is also related to emotional behaviour and therefore to mental workload as well.

The pupil diameter typically ranges from 2 mm to 8 mm for respectively bright and dark lighting. As a result of mental activity the pupils slightly dilate (up to 0.5 mm) and return to their previous size within a few seconds after a mental task is completed. The diameter during mental activity is not smooth and constant but has a rather irregular and impulsive character. These small and involuntary fluctuations of the pupil diameter as a reaction to a discrete cognitive processing event are called the task-evoked pupillary response (TEPR). Examples of such mental tasks are: short and long-term memory access, mental arithmetic, sentence comprehension, vigilance, and visual and auditory perception tasks (Beatty & Lucero-Wagoner, 2000). See Appendix N, figure N2 and N3 for their results on pupil dilation during mental arithmetic and short-term memory tasks. Beatty and Lucero-Wagoner (2000) think that the TEPR is biologically useless and has not been
erased by evolution, possibly because it has no evolutionary cost. Pupillary dilations or constrictions evoked by changes in lighting conditions can be much larger (up to 6 mm), where the dilations themselves can take a few seconds and the constrictions about 1 second. During constant ambient luminance the pupils also respond with a continual but irregular response oscillation, which is called the light reflex. However, the constrictions and dilations are slower and smaller for light reflexes than for cognitive processing reflexes (Marshall, 2000; See Appendix N, figure N4 for three typical pupillary responses).

There are several ways of measuring mental workload using pupillometry. Most of the measures make use of TEPR averaging of the mean pupil diameter. A different approach is the index of cognitive activity (ICA), which will be discussed later since this technique is much younger. The first researcher that measured pupil dilation evoked by mental workload was Heinrich (1896) and his results were confirmed by Hess and Polt (1964) many years later. Beatty and Lucero-Wagoner (2000) provide detailed descriptions in their review of how to measure mental workload using TEPR averaging techniques and how to analyse the data. It is important to note that since TEPR’s do not occur reliably, multiple pupillary responses must be recorded and averaged to come up with a good estimate of an individual’s mental workload. But before the data can be averaged, all disturbances (e.g. eye blinks, movement artefacts, light reflexes) should be removed in order to minimise the noise, which can be a time consuming process.

The relation between pupil dilation and driver’s mental workload has been researched in a series of studies (Demberg, 2013; Dlugosch et al., 2013; Kun et al., 2013; Palinko et al., 2010; Razumenic et al., 2012; Recarte & Nunes, 2000, 2003; Schwalm et al., 2008). In two real driving studies, Recarte and Nunes (2000, 2003) used secondary cognitive tasks to induce mental workload during driving and recorded the pupil’s dilation of 12 individuals. The results were consistent and showed that the pupil diameter is sensitive to changes in mental processing task demands. In a follow-up study Recarte et al. (2008) studied the validity of pupil dilation in single- and dual-task conditions. For the dual-task condition it was found that, unlike blink rate, pupil dilation does not discriminate between mental and visual workload. The researchers therefore think that pupil dilation measures the average arousal underlying the cognitive tasks.

Klingner et al. (2008) performed three experiments (mental multiplication, short-term memory, and aural vigilance) in which they measured the TEPR with a remote eye tracker to estimate cognitive load. The lighting was carefully controlled by covering all the room’s windows. All three experiments showed that the pupil diameter is a reliable measure of mental workload. Elaborating on these results, Palinko et al. (2010) used a remote eye tracker in a driving simulator study to estimate driver’s mental workload. The 16 drivers that participated in this study were instructed to follow a vehicle while driving responsibly. An ongoing spoken task was used as a secondary task and was interrupted by another spoken task once in all corners and straight road segments in order to vary the mental workload. The lighting conditions were not strictly controlled, but the brightness of the simulator screen was limited to ±5% of the mean brightness. Again, it was found that the pupil diameter is a reliable way of measuring the driver’s mental workload, especially in simulated driving environments. Palinko et al. (2010) also introduced a new measuring variant of the pupil diameter called the mean pupil diameter change rate (MPDCR). They suggested that this new measure might be particularly useful when rapid changes in the mental workload should be detected (See Appendix O, figure O1).

Subsequent studies explored the effects of mental workload and illumination on the pupil diameter in more detail (Palinko & Kun, 2011, 2012; Razumenic et al., 2012). Palinko and Kun (2011) used three different tasks to investigate these effects: illumination, aural vigilance, and combined. For the illumination task subjects were instructed to fixate their gaze on one of three differently
coloured trucks (black, grey, and white). The aural vigilance task, originally designed by Klingner et al. (2008), consisted of repeatedly listening to a sequence of numbers. It was the subjects’ task to press a button if they detected an error. The results showed that it is possible to separate the effects of lighting and mental workload on the pupil diameter using a light reflex predictor, which tries to predict the reflex of the pupil to illumination. The prediction is than subtracted from the pupil signal, leaving an estimation of the TEPR. However, the authors also mentioned that due to the averaging process this method cannot be used in real time. One year later, Palinko and Kun (2012) studied the interference between two different visual tasks (illumination and visual vigilance) in alternating lighting conditions and their effect on the pupil diameter. Again, the researchers confirmed that it is possible to separate mental workload and illumination effects on pupil diameter. These results seem promising; however, the authors note that the tasks that were used in these experiments were slow-paced and highly constrained, while driving a car in a simulator is not.

Driver’s mental workload has also been investigated by Kun et al. (2013) in relation with in-vehicle spoken dialogues. The 16 individuals that participated in this driving simulator study were instructed to follow a vehicle at a comfortable distance. During the ride, the drivers played a series of word games with a remote conversant with the intention to vary the mental workload. The driver’s TEPR was analysed for two points in time: 1) just before the first contribution of the remote conversant and 2) just before the first contribution of the driver’s response. It was found that the pupil diameter was significantly larger at point 2 than at point 1 in 69% of all word games. Several possible explanations are given by the authors for this fairly weak effect, including the possibility that in some cases the TEPR may have been masked by the pupillary light reflex. So once again, researchers warn against the negative effects of light reflexes on pupil dilation.

As mentioned before, a different method to measure mental workload using pupillometry is the index of cognitive activity (ICA), which is patented by Marshall (2000). The ICA is an algorithm that uses the different reflex properties to separate the effects of illumination and mental workload on the pupil’s size. Advanced signal processing techniques are being used to remove noise and the large and slow oscillations from the raw data in order to retain the very short and rapid oscillations. The resultant ICA values are then obtained by transforming the number of increases in pupil size per second using the hyperbolic tangent. This leaves a number between zero and one, with values close to zero indicating a low cognitive workload and close to one reflecting a high cognitive workload. Just like in the previously described pupillary measurements, blinks and other eye movement artefacts are removed from the raw data to obtain a continuous measure.

The ICA method has been used in a number of studies and has shown to be a promising method of measuring driver’s mental workload (Demberg, 2013; Demberg et al., 2013; Dlugosch et al., 2013; Schwalm, 2009; Schwalm et al., 2008). Schwalm (2009) was one of the first to show that the ICA may be a reliable method to identify driver’s mental demands in both highly controlled driving simulators as well as in real environments. The author emphasises in his report that this index is particularly useful for studies in which a higher temporal resolution is required than could be obtained by the traditional performance or subjective measures.

Demberg et al. (2013) used the ICA method together with the continuous tracking and reaction (ConTRe) task to estimate mental workload during driving. The ConTRe task (Mahr et al., 2012) was designed for driving simulator studies to enhance the investigation of mental workload and consists of two parts: tracking and reaction. The tracking part requires continuous driver attention to trace a reference cylinder with a steering cylinder (projected on the simulator screen at a constant longitudinal distance in front of the vehicle) using the steering wheel. The reaction part
of the ConTRe task requires subjects to react to a red or a green light stimulus by pushing down the brake or accelerator pedal respectively, although Demberg et al. (2013) did not use this part in their experiment. While driving, participants performed a speech comprehension task: listening to sentences containing subject and object relative clauses followed by a comprehension question. Since object relative clauses are known to be more mentally demanding than subject relative clauses, the researchers expected to find a decreased tracking performance and increased resultant ICA values for object compared to subject relative clauses. The results confirmed this hypothesis and the authors therefore conclude that the ICA method is a reliable method to measure language-induced mental workload. The ConTRe task played an important role in this investigation allowing for sensitive and continuous measurements of mental workload in a driving simulator.

The ICA method was also used by Dlugosch et al. (2013) in a static simulator study. Driver’s mental workload in this study was assessed using a head mounted eye tracker. The primary task was to follow a leading car at a constant distance while driving safely. During the experiment, participants performed a detection response task (DRT) every 3-4 seconds. In addition, participants were asked to do several secondary tasks, including three conversation tasks and a surrogate reference (visual search and manual response) task (SuRT). Significantly higher resultant ICA values were found for the SuRT conditions compared to all other conditions, indicating a higher mental workload. The resultant ICA values for the driving experiments with the three conversation tasks were also significantly higher than the values for driving only. This clearly reflects the added mental workload through conversation.

5.7. Discussion

The aim of this study was to summarize the knowledge obtained in older studies and to discuss the results of the more recently conducted studies about the relation between eye-related parameters and drivers’ mental workload.

As described in this review, there are four blink-related parameters that are used in the literature to assess a driver’s mental workload: blink rate, blink duration, blink latency, and PERCLOS. Blink rate has shown to be sensitive to visual and mental workload, and fatigue. Both types of workload produce opposite effects according to Recarte et al. (2008): “Blink inhibition for higher visual demand and increased blink rate for higher mental workload”. However, several studies found a blink rate increase during highly automated driving compared to normal driving (Cha, 2003; Damböck et al., 2013; Merat et al., 2012). This could suggest that highly automated driving relieves a driver more from the visual tasks than from the mental tasks. Blink latency and blink duration have shown to be more straightforward or less nuanced indicators of workload (i.e. non-discriminatory between a mental or visual workload component). The results seem to be consistent: blink duration decreases and blink latency increases for both increases in mental and visual processing demands. Kramer (1990) argued that blink durations decrease and latencies increase to extract task-relevant information from the environment as efficient as possible.

Although PERCLOS is mainly used in the literature as a measure of fatigue, Halverson et al. (2012) demonstrated that it could also be used to distinguish between two different levels of workload. The authors encouraged researchers to combine different metrics when assessing mental workload to increase accuracy and robustness. This suggestion is sensible, because mental workload is a broad concept and might not best be described by a single one-dimensional metric, especially in driving where visual workload plays a major role as well. Additionally, individual
differences like age, risk-proneness, and driving experience have shown to affect drivers’ mental workload, which makes assessing it even more challenging.

This literature review also pointed out the major drawback of using eye blinks to assess a driver’s mental workload: unwanted blinks due to for instance annoying luminance or poor air quality. Factors like these are hard to control which makes measuring blinks on real roads very complicated.

Fixation duration is found to be sensitive to workload and is generally believed to increase with increasing visual and mental processing demands (Rayner & Morris, 1990; Recarte & Nunes, 2000). A series of studies later reconfirmed this finding by proving that fixation duration also increases when approaching potential hazards during driving (Chapman & Underwood, 1998a, 1998b; Chapman et al., 2007; Crundall & Underwood, 1998; Konstantopoulos et al., 2010; Underwood et al., 2005).

Several other studies have demonstrated that visual and mental processing demands produce visual-detection impairment and attentional focus narrowing (Miura, 1986, 1990; Radach, 1998; Recarte & Nunes, 2000; Reimer, 2009; Reimer et al., 2010; Victor et al., 2005). A recent study showed that horizontal gaze dispersion is more sensitive to drivers’ mental workload than a combined metric between horizontal and vertical gaze dispersion (Wang et al., 2014).

Pupillometry is the measurement of the pupil diameter and is considered to be a reliable physiological measure of mental workload. This measurement method has also been used in driving studies to measure drivers’ mental workload. In highly controllable environments, such as driving simulators, the results are consistent and show that the pupil diameter is sensitive to changes in mental workload. However, it is believed that pupil dilation, unlike blink rate, does not discriminate between mental and visual workload.

Palinko et al. (2010) introduced a new measuring variant of the pupil diameter called the mean pupil diameter change rate. They suggested that this new measure might be particularly useful when rapid changes in the mental workload have to be detected.

More recent studies explored the effects of mental workload and illumination on the pupil diameter in more detail. Palinko and Kun (2011, 2012) showed that it is possible to separate mental workload and illumination effects on pupil diameter. However, it should be noted that the tasks that were used in these studies were slow paced and highly constrained, while driving a car in a simulator is not.

The index of cognitive activity (ICA; patented by Marshall, 2000) is a promising high temporal resolution method and can be used to identify drivers’ mental workload in both highly controlled driving simulators as well as in real environments due to its ability to account for illumination effects (Demberg, 2013; Demberg et al., 2013; Dlugosch et al., 2013; Schwalm, 2009). This index is calculated through an algorithm that uses pupillometry, but is unfortunately patented and therefore cannot easily be used in research.

In summary, there are many different ways to measure a driver’s mental workload but there is no single method that works in all cases, since driving is a multi-dimensional task and all metrics have specific drawbacks. The focus for future research should be on combining multiple assessment methods to increase validity and robustness, which is particularly important when applied in real cars. Advanced measures, like the ICA method, should play a major in this ongoing
development of measuring methods towards a complete understanding of the driver’s psychophysiological state.

5.8. References


6.1. Appendix A: Informed Consent Form for Participants

This appendix contains the form which allowed participants to provide written informed consent. It describes, inter alia, the purpose and procedures of the experiment.

Research title: “Detecting changes in pupil diameter while completing mental arithmetic tasks”

Researchers:
Gerhard Marquart – MSc student
Email: g.marquart@student.tudelft.nl

Dr.ir. Joost C.F. de Winter – Supervisor
Email: j.c.f.dewinter@tudelft.nl

Location of the experiment:
Driving simulator lab; room 34 G-0-210
Faculty of Mechanical, Maritime and Materials Engineering
Delft University of Technology
Mekelweg 2, 2628 CD Delft

Introduction: Before agreeing to participate in this study, you are asked to read this document carefully. This document describes the purpose, procedures, risks, and discomforts of this study. Your signature is required for participation. You have the right to withdraw at any time. If you desire a copy of this consent form, you may request one and we will provide it.

Purpose of the study: The purpose of this study is to investigate whether the pupil diameter is a good indicator of mental workload. You will be requested to do a series of mental arithmetic tasks (multiplications) while sitting behind a desktop computer with your chin on a chinrest and looking at a fixed point on the screen. The results of this experiment will be published in a Master’s thesis and research paper.

Duration: Your participation in this study will last approximately 30 minutes. You will be compensated for your time with €5 and have the right to withdraw from the experiment at any time, without losing the compensation.

Procedures: Before the experiment starts, you will be asked to position yourself at approximately 65 cm from the monitor, after which the eye tracker will be calibrated.

You will be requested to do 50 trials of mental arithmetic tasks (multiplications), five of which will be used as a short training. The remaining 45 trials are evenly divided into 3 sessions of varying difficulty (easy, medium, and hard). Each trial is initiated by you and starts with a 4 seconds accommodation period, followed by a 1 second visual presentation of two numbers (multiplicand and multiplier) between 5 and 18, with a 1.5 seconds pause in between. It is your task to multiply these two numbers and type your answer 10 seconds after the multiplier is shown. So the total duration of one trial is 17.5 seconds (4+1+1.5+1+10). When the numbers are not presented, they will be masked by an “X”.

After each session, you will be asked to fill out a questionnaire to assess your subjective mental workload. In order to obtain the most reliable eye tracking results, it is important for you to stay
still and focus (not stare) at the black letters or numbers in the middle of the screen during an entire session.

**Risks and discomforts:** There are no known risks for you in this study. You may feel uncomfortable focusing your eyes on a fixed point on the screen for a longer period of time.

**Confidentiality:** All data collected in this study will be kept confidential and will be used for research purposes only.

**Right to refuse or withdraw:** Your participation in this study is voluntary. You have the right to refuse or withdraw from this experiment at any time, without negative consequences.

**Questions:** Contact Gerhard Marquart (contact details are included at the top) in case you have any questions or concerns about this study or your rights as a research participant.

I have read and understood the information provided above.
I give permission to process the data for the purposes described above.
I voluntarily agree to participate in this study.

Name:

...........................................

..................................................

Signature of participant: Date:

....................................................................................  ....../ ......../ .......

..........................................................  ....../ ....../ .......
6.2. Appendix B: Classification of Arithmetic Tasks II

Three levels of arithmetic task difficulty were used for the full-scale experiment. Each task consisted of calculating the multiplication between two digits ranging from 5 to 18. The tasks were sorted from easy to hard by the outcome of their multiplication. It was assumed that multiplications with a lower outcome were easier than those with a higher outcome. So in this case the easiest task was 5x12 and the hardest was 18x18. The digits 10, 11 were excluded in this method, since they were considered to be too easy. This left 63 possible multiplications, with the assumption that AxB and BxA were equally difficult.

The multiplications were then distributed over three different levels of difficulty (easy, medium and hard), all containing 21 possible multiplications. In order to make a clear distinction between the three levels of difficulty, the first six multiplications were removed from each level. Table B1 shows the removed and selected multiplications of the three levels. Note that the smallest digit of a pair is put down first, but during the experiment they were presented to the participant in randomized order.

*Table B1.* All possible multiplications between 6 and 18 (10, 11 and 15 are excluded), sorted by difficulty and classified into three different levels (Level 1 being the easiest and Level 3 being the hardest).

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Removed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 12</td>
<td>7 16</td>
<td>13 15</td>
</tr>
<tr>
<td>5 13</td>
<td>8 14</td>
<td>14 14</td>
</tr>
<tr>
<td>5 14</td>
<td>9 13</td>
<td>12 17</td>
</tr>
<tr>
<td>6 12</td>
<td>7 17</td>
<td>13 16</td>
</tr>
<tr>
<td>5 15</td>
<td>8 15</td>
<td>14 15</td>
</tr>
<tr>
<td>6 13</td>
<td>7 18</td>
<td>12 18</td>
</tr>
<tr>
<td>Selected</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 16</td>
<td>9 14</td>
<td>13 17</td>
</tr>
<tr>
<td>6 14</td>
<td>8 16</td>
<td>14 16</td>
</tr>
<tr>
<td>7 12</td>
<td>9 15</td>
<td>15 15</td>
</tr>
<tr>
<td>5 17</td>
<td>8 17</td>
<td>13 18</td>
</tr>
<tr>
<td>5 18</td>
<td>8 18</td>
<td>14 17</td>
</tr>
<tr>
<td>6 15</td>
<td>9 16</td>
<td>15 16</td>
</tr>
<tr>
<td>7 13</td>
<td>12 12</td>
<td>14 18</td>
</tr>
<tr>
<td>6 16</td>
<td>9 17</td>
<td>15 17</td>
</tr>
<tr>
<td>8 12</td>
<td>12 13</td>
<td>16 16</td>
</tr>
<tr>
<td>7 14</td>
<td>9 18</td>
<td>15 18</td>
</tr>
<tr>
<td>6 17</td>
<td>12 14</td>
<td>16 17</td>
</tr>
<tr>
<td>8 13</td>
<td>13 13</td>
<td>16 18</td>
</tr>
<tr>
<td>7 15</td>
<td>12 15</td>
<td>17 17</td>
</tr>
<tr>
<td>6 18</td>
<td>13 14</td>
<td>17 18</td>
</tr>
<tr>
<td>9 12</td>
<td>12 16</td>
<td>18 18</td>
</tr>
</tbody>
</table>
6.3 Appendix C: Blink Identification and Removal

During a blink, the eyelid opening rapidly diminishes to zero and then increases in a few tenths of a second until it is fully open again (see Fig. C1, solid blue line). It is impossible to track the pupil’s diameter while blinking. These instances in time should therefore be removed from the data. The recordings of the eyelid opening were used to identify the blinks in the pupil diameter data. A threshold of 75% of the mean eyelid opening was used to make a clear distinction between blinks and no blinks as depicted in the figure by the dashed red line.

![Figure C1. Sample of the recordings of the eyelid opening showing a typical blink (blue) and the threshold (red) used to identify it.](image)

As can be seen in the figure, it takes some time to cross the threshold and the blink has not been completed after the eyelid opening signal crossed the threshold line for the second time. That is why 12 additional data points (~0.1 s) were removed from the data before the blink and 36 additional data points (~0.3 s) after the blink.
6.4. **Appendix D: Individual Trials**

This appendix shows the processed and unprocessed pupil diameter (PD), pupil diameter quality (PDQ), and eyelid opening for four individual trials (see Figure D1 and D2).

*Figure D1.* Left: Level 1, participant 2, trial 7. Right: Level 2, participant 11, trial 9.
Figure D2. Left: Level 3, participant 13, trial 1. Right: Level 3, participant 24, trial 15.
6.5. Appendix E: MPD and MPDC bar graphs

Figure E1 and E2 respectively show the mean pupil diameter (MPD) and mean pupil diameter change (MPDC), for six points in time and three levels of difficulty. The bar graphs are a visualisation of the results found in Table 2 of Chapter 1.

*Figure E1.* Mean pupil diameter (MPD) of 29 participants, for three levels of difficulty and eight points in time during the presentation of the multiplier and the calculation period. The asterisks indicate significant differences between the levels of difficulty.

*Figure E2.* Mean pupil diameter change (MPDC) of 29 participants, for three levels of difficulty and eight points in time during the presentation of the multiplier and the calculation period.
6.6. Appendix F: Eight-point analysis of correct and incorrect responses

The results of the eight-point analysis for the correct and incorrect responses of difficulty Level 3 are shown in Table F1 and Figure F1.

Table F1. Mean Pupil Diameter (MPD). The means (M) and standard deviations (SD) of 25 participants are shown for Level 3 of the multiplications, and separated for correct and incorrect responses. P1-P8 refers to the eight points in time. Statistically significant differences are indicated in boldface.

<table>
<thead>
<tr>
<th>Level 3</th>
<th>Level 3 Correct</th>
<th>Level 3 Incorrect</th>
<th>p-value</th>
<th>Effect size</th>
<th>Pairwise comparison of conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MPD (mm) (N = 25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>3.915 (0.490)</td>
<td>3.919 (0.508)</td>
<td>0.140</td>
<td>0.08 (0.00)</td>
<td>0.991  0.222  0.178</td>
</tr>
<tr>
<td>P2</td>
<td>3.985 (0.516)</td>
<td>4.002 (0.524)</td>
<td>0.027</td>
<td>0.14 (0.00)</td>
<td>0.803  0.112  <strong>0.027</strong></td>
</tr>
<tr>
<td>P3</td>
<td>4.080 (0.531)</td>
<td>4.079 (0.534)</td>
<td>0.642</td>
<td>0.02 (0.00)</td>
<td>1.000  0.685  0.703</td>
</tr>
<tr>
<td>P4</td>
<td>4.138 (0.522)</td>
<td>4.132 (0.526)</td>
<td>0.638</td>
<td>0.02 (0.00)</td>
<td>0.975  0.636  0.767</td>
</tr>
<tr>
<td>P5</td>
<td>4.157 (0.521)</td>
<td>4.135 (0.534)</td>
<td>0.662</td>
<td>0.02 (0.00)</td>
<td>0.711  1.000  0.709</td>
</tr>
<tr>
<td>P6</td>
<td>4.135 (0.518)</td>
<td>4.109 (0.529)</td>
<td>0.063</td>
<td>0.11 (0.00)</td>
<td>0.732  0.250  0.056</td>
</tr>
<tr>
<td>P7</td>
<td>4.144 (0.500)</td>
<td>4.126 (0.517)</td>
<td>0.224</td>
<td>0.06 (0.00)</td>
<td>0.906  0.421  0.220</td>
</tr>
<tr>
<td>P8</td>
<td>4.125 (0.493)</td>
<td>4.084 (0.516)</td>
<td><strong>0.049</strong></td>
<td>0.12 (0.01)</td>
<td>0.672  0.240  <strong>0.042</strong></td>
</tr>
</tbody>
</table>

Figure F1. Mean pupil diameter (MPD) of 25 participants for Level 3, and separated for correct and incorrect responses.
This appendix shows the task-evoked pupillary responses obtained by Ahern (see Figure G1) and Klingner (see Figure G2, G3, and G4) during mental multiplication. In the first three figures, the multiplication digits were presented one at a time with a short pause in between, while in the last figure the digits were presented simultaneously. In all figures, clear pupil dilations can be found during the multiplication period and, where applicable, larger dilations for more difficult calculations.

Figure G1. “Averaged task-evoked pupillary responses for correctly solved problems at three levels of difficulty for subject in the high and low groups of psychometrically measured intelligence. At all difficulty levels, larger pupillary responses are observed for subjects in the low group”. From “Pupillary Responses During Information Processing Vary with Scholastic Aptitude Test Scores,” by S.K. Ahern & J. Beatty, 1979, Science, 205, 1289-92.
Figure G2. “Pupillary response during the mental multiplication task”. From “Measuring cognitive load during visual tasks by combining pupillometry and eye tracking,” by J. Klingner, 2010, Ph.D. dissertation, Stanford University Computer Science Department.

Figure G3. “Average pupil dilation evoked by visually and aurally presented mental multiplication problems. The two presentation modes elicited dilations with similar timing, duration, and shape, but different magnitude. Vertical lines show the time during which the two numbers were spoken or displayed and the time during which the two numbers were spoken or displayed and the time during which the participants responded”. From “Measuring cognitive load during visual tasks by combining pupillometry and eye tracking,” by J. Klingner, 2010, Ph.D. dissertation, Stanford University Computer Science Department.
Figure G4. “Difficulty effect on pupil dilation evoked by mental multiplication of two numbers displayed together for eight seconds, separated by difficulty”. From “Measuring cognitive load during visual tasks by combining pupillometry and eye tracking,” by J. Klingner, 2010, Ph.D. dissertation, Stanford University Computer Science Department.
6.8. Appendix H: Classification of Arithmetic Tasks I

Five levels of arithmetic task difficulty were used in the pilot study. Each task consisted of calculating the multiplication between two digits ranging from 5 to 24. The tasks were sorted from easy to hard by the outcome of their multiplication. It was assumed that multiplications with a lower outcome were easier than those with a higher outcome. So in this case the easiest task was 5x5 and the hardest was 24x24. The digits 10 and 20 were excluded in this method, since they were considered to be too easy. This left 171 possible multiplications, with the assumption that AxB and BxA were equally difficult.

The multiplications were then distributed over five different levels of difficulty (easy to hard), all containing 34 possible multiplications (except Level 1 which has 35; see Table H1). Note that the smallest digit of a pair was put down first, but this does not mean that it was also presented first to the participant.

*Table H1.* All possible multiplications between 5 and 24 (10 and 20 are excluded), sorted by difficulty and classified into five different levels (Level 1 being the easiest and Level 5 being the hardest).

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 5</td>
<td>7 14</td>
<td>8 19</td>
<td>14 16</td>
<td>17 19</td>
</tr>
<tr>
<td>5 6</td>
<td>9 11</td>
<td>9 17</td>
<td>15 15</td>
<td>18 18</td>
</tr>
<tr>
<td>5 7</td>
<td>6 17</td>
<td>7 22</td>
<td>12 19</td>
<td>15 22</td>
</tr>
<tr>
<td>6 6</td>
<td>8 13</td>
<td>11 14</td>
<td>11 21</td>
<td>14 24</td>
</tr>
<tr>
<td>5 8</td>
<td>5 21</td>
<td>12 13</td>
<td>13 18</td>
<td>16 21</td>
</tr>
<tr>
<td>6 7</td>
<td>7 15</td>
<td>7 23</td>
<td>14 17</td>
<td>18 19</td>
</tr>
<tr>
<td>5 9</td>
<td>6 18</td>
<td>9 18</td>
<td>15 16</td>
<td>15 23</td>
</tr>
<tr>
<td>6 8</td>
<td>9 12</td>
<td>11 15</td>
<td>11 22</td>
<td>16 22</td>
</tr>
<tr>
<td>7 7</td>
<td>5 22</td>
<td>7 24</td>
<td>13 19</td>
<td>17 21</td>
</tr>
<tr>
<td>6 9</td>
<td>7 16</td>
<td>8 21</td>
<td>12 21</td>
<td>15 24</td>
</tr>
<tr>
<td>5 11</td>
<td>8 14</td>
<td>12 14</td>
<td>14 18</td>
<td>19 19</td>
</tr>
<tr>
<td>7 8</td>
<td>6 19</td>
<td>13 13</td>
<td>11 23</td>
<td>16 23</td>
</tr>
<tr>
<td>5 12</td>
<td>5 23</td>
<td>9 19</td>
<td>15 17</td>
<td>17 22</td>
</tr>
<tr>
<td>7 9</td>
<td>9 13</td>
<td>8 22</td>
<td>16 16</td>
<td>18 21</td>
</tr>
<tr>
<td>8 8</td>
<td>7 17</td>
<td>11 16</td>
<td>11 24</td>
<td>16 24</td>
</tr>
<tr>
<td>5 13</td>
<td>5 24</td>
<td>12 15</td>
<td>12 22</td>
<td>17 23</td>
</tr>
<tr>
<td>6 11</td>
<td>8 15</td>
<td>13 14</td>
<td>14 19</td>
<td>18 22</td>
</tr>
<tr>
<td>5 14</td>
<td>11 11</td>
<td>8 23</td>
<td>15 18</td>
<td>19 21</td>
</tr>
<tr>
<td>6 12</td>
<td>6 21</td>
<td>11 17</td>
<td>16 17</td>
<td>17 24</td>
</tr>
<tr>
<td>8 9</td>
<td>7 18</td>
<td>9 21</td>
<td>13 21</td>
<td>18 23</td>
</tr>
<tr>
<td>5 15</td>
<td>9 14</td>
<td>8 24</td>
<td>12 23</td>
<td>19 22</td>
</tr>
<tr>
<td>7 11</td>
<td>8 16</td>
<td>12 16</td>
<td>15 19</td>
<td>18 24</td>
</tr>
<tr>
<td>6 13</td>
<td>6 22</td>
<td>13 15</td>
<td>13 22</td>
<td>19 23</td>
</tr>
<tr>
<td>5 16</td>
<td>11 12</td>
<td>14 14</td>
<td>12 24</td>
<td>21 21</td>
</tr>
<tr>
<td>9 9</td>
<td>7 19</td>
<td>9 22</td>
<td>16 18</td>
<td>19 24</td>
</tr>
<tr>
<td>6 14</td>
<td>9 15</td>
<td>11 18</td>
<td>17 17</td>
<td>21 22</td>
</tr>
<tr>
<td>7 12</td>
<td>8 17</td>
<td>12 17</td>
<td>14 21</td>
<td>21 23</td>
</tr>
<tr>
<td>5 17</td>
<td>6 23</td>
<td>9 23</td>
<td>13 23</td>
<td>22 22</td>
</tr>
<tr>
<td>8 11</td>
<td>11 13</td>
<td>13 16</td>
<td>16 19</td>
<td>21 24</td>
</tr>
<tr>
<td>5 18</td>
<td>6 24</td>
<td>11 19</td>
<td>17 18</td>
<td>22 23</td>
</tr>
<tr>
<td>6 15</td>
<td>8 18</td>
<td>14 15</td>
<td>14 22</td>
<td>22 24</td>
</tr>
<tr>
<td>7 13</td>
<td>9 16</td>
<td>9 24</td>
<td>13 24</td>
<td>23 23</td>
</tr>
<tr>
<td>5 19</td>
<td>12 12</td>
<td>12 18</td>
<td>15 21</td>
<td>23 24</td>
</tr>
<tr>
<td>6 16</td>
<td>7 21</td>
<td>13 17</td>
<td>14 23</td>
<td>24 24</td>
</tr>
<tr>
<td>8 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 6.9. Appendix I: Additional Results

In addition to the analyses conducted as described in Chapter 1, the mean pupil diameter velocity (MPDV) and its absolute form was analysed. First, all the MPD trials were filtered in order to remove the unrealistically sharp edges and to avoid peaks in the velocity. Figure I1 visualises this processing step for one trial. The MPDV was then obtained by differentiating the filtered individual MPDC trials (see Figure I2). The absolute MPDV was calculated by taking the absolute difference between two data points (see Figure I3). Since no remarkable differences were found between the three levels, no statistical tests were performed. It should be noted, however, that a big increase of the MPVD for all three levels was observed just after the multiplier was presented.

![Figure I1. Filtered and unfiltered version of the mean pupil diameter change of one trial.](image)

![Figure I2. Mean pupil diameter velocity during the mental multiplication task of 29 participants, for three levels of difficulty.](image)
Figure I3. Absolute mean pupil diameter velocity during the mental multiplication task of 29 participants for three levels of difficulty.
6.10. Appendix J: Measurement Equipment

Figure J1 and J2 give an overview of the physical arrangement of the equipment. The experimenter’s desk was positioned behind the participant such that the participant could completely focus on the task, while the experimenter could monitor the experiment. Each device is labelled with a number and its corresponding function description can be found in Table J1.

![Figure J1. Arrangement of equipment on participant’s desk.](image1)

![Figure J2. Arrangement of equipment on experimenter’s desk.](image2)

<table>
<thead>
<tr>
<th>#</th>
<th>Device</th>
<th>Function description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Headrest</td>
<td>Adjustable head support stabilises the head</td>
</tr>
<tr>
<td>2</td>
<td>Keyboard</td>
<td>Numpad functions as input device for the experiment</td>
</tr>
<tr>
<td>3</td>
<td>Eye tracker</td>
<td>SmartEye DR120 eye tracker tracks the pupil diameter</td>
</tr>
<tr>
<td>4</td>
<td>Monitor</td>
<td>SmartEye DR120 monitor displays task-relevant information</td>
</tr>
<tr>
<td>5</td>
<td>Experiment computer</td>
<td>Laptop runs MATLAB with experiment and logs data</td>
</tr>
<tr>
<td>6</td>
<td>SmartEye computer</td>
<td>Laptop runs eye tracking software: SmartEyePro 6.0</td>
</tr>
<tr>
<td>7</td>
<td>Data storage</td>
<td>Portable storage is used to back-up data</td>
</tr>
<tr>
<td>8</td>
<td>Administration</td>
<td>Forms standardise experimental procedures</td>
</tr>
</tbody>
</table>
6.11. Appendix K: Experiment Interface

Prior to the experiment, the participants were asked to enter their participant number, age, and gender in the Information panel (see Figure K1, left panel), and hit the ‘OK’ button (by pressing enter) when ready. The experiment was then started with by hitting the two consecutive buttons ‘Next round’ and ‘Next trial’ (see Figure K1, bottom). After each trial, the participants entered their outcome to the task using the Answer panel (see Figure K1, right), and again hit the ‘OK’ button when they were ready.

![Information panel](image)

![Answer panel](image)

*Figure K1.* Left: Information panel with the participant information related questions. Right: Answer panel. Bottom: Next round and trial buttons, used to initiate a round or trial.

As explained in the method section of Chapter 1, in each trial the participants were visually presented with an “XX” during the accommodation, pause, and calculation period (see Figure K2, left). During the presentation of the multiplicand and multiplier the mask was replaced by two digits (see Figure K2, right).

![XX](image)

![15](image)

*Figure K2.* Left: Information panel with the participant information related questions. Right: Answer panel. Bottom: Next round and trial buttons, used to initiate a round or trial.
After each round, the participants were asked to fill out the NASA-TLX questionnaire in the NASA-TLX panel shown in Figure K3, and again hit the ‘OK’ button when ready.

All panels and buttons were presented one at a time and programmed such that a simple press of the enter key automatically placed the cursor at the text box or button. The tab key could be used to navigate through a panel.

Figure K3. NASA-TLX panel, used to assess a participant’s subjective workload.
6.12. Appendix L: Data Transfer and Logging

Two laptops were used during this experiment, as was shown in Appendix J. One laptop was used to run the eye tracking software (SmartEyePro 6.0), and one to run the experiment and store the data (MATLAB). The data from the first laptop was transferred real-time to the other via a UDP wireless network. A simple Simulink model was used to store the data in the MATLAB workspace (see Figure L1). This model consisted of four blocks: UDP Receive, Simulation Pace, To Workspace, and Terminator. The UDP Receive block brought in the UDP packets from a given IP port and emitted them as a one-dimensional vector to the workspace. The Simulation Pace block ensured that the pace of the simulation was real time, meaning 1s of simulation time was exactly equal to 1s in real time.

*Figure L1. Simulink model receives and stores the data in the MATLAB Workspace.*
6.13. **Appendix M: Literature Blinks**

Appendix N: Literature Pupillometry

Figure N1. "Muscles of the iris. Two opposing muscle groups within the iris of the human eye determine the aperture of the pupil. The sphincter muscles of the iris constrict the pupil when they contract, whereas constriction of the dilator muscles increases the pupillary diameter." From "The pupillary system," by J. Beatty & B. Lucero-Wagoner, 2000, Handbook of Psychophysiology, Cambridge: Cambridge University Press.

Figure N2. "Pupillary dilations during mental multiplication: the original Hess and Polt (1964) finding. A simple but convincing demonstration that pupillary movements reflect an essential physiological aspect of cognitive processing was provided by Hess and Polt, who measured pupillary dilations as subjects mentally calculated the product of two numbers provided by the experimenter. They reported their results as percentage pupillary dilation. Today, pupillary movements are conventionally reported as changes in diameter (in millimeters), since this measure appears to be independent of baseline pupillary diameter for a wide range of baseline values (Beatty 1982b)." From "The pupillary system," by J. Beatty & B. Lucero-Wagoner, 2000, Handbook of Psychophysiology, Cambridge: Cambridge University Press.

Figure N3. "Task-evoked pupillary response amplitude during a short-term memory task. In a paced auditory digit-span task, pupillary diameter increases as each successive digit is presented to the subject and reaches maximum dilation between the end of digit presentation and the beginning of digit report. Peak pupillary dilation varies directly with the number of digits held in short-term memory for immediate report. During the report itself, pupillary diameter decreases with each digit spoken returning to baseline diameter at the completion of the trial. (Data from Kahneman & Beatty, 1966)." From "The pupillary system," by J. Beatty & B. Lucero-Wagoner, 2000, Handbook of Psychophysiology, Cambridge: Cambridge University Press.
Figure N4. "Three different test conditions for one subject, in which (a) depicts the ordinary pupillary light reflex under typical experimental lighting, (b) depicts the pupil signal for the same subject when the room was darkened and the monitor turned off, and (c) depicts the pupil signal in the darkened room and dark screen condition when the subject engaged in active cognitive processing." From "Method and apparatus for eye tracking and monitoring pupil dilation to evaluate cognitive activity," by S.P. Marshall, 2000, U.S. patent no. 6,090,051.
Appendix P: MATLAB Code

This Appendix contains six m-files. “EXPERIMENT.m” was used to run the experiment and save the data. “EXPERIMENT_GUI.m” starts the graphical user interface. “ANALYSIS_MPDC.m”, “ANALYSIS_MPD.m”, and “ANALYSIS_BLINKS.m” contain respectively the analysis for the MPDC, MPD, and MBR. The function “blinkremoval.m” contains the code that identifies and removes the blinks from the data.

EXPERIMENT.m

%% Mental Multiplication Experiment
% Gerhard Marquart
% #1538446

clear all; close all; clc;
i = 0; % First round (Can be started at round x, fill in x-1)
j = 0; % First trial (Always starts at j=1)

% Training
rounds = 1; % Total number of rounds
trials = 5; % Total number of trials per round

% Experiment
rounds = 3; % Total number of rounds
trials = 15; % Total number of trials per round

Ran_D = [1 2 3];
% Ran_D = datasample([1 2 3],3,'Replace',false);

PN = []; Age = []; Gender = [];
N1 = nan(1,rounds); N2 = nan(1,rounds); N3 = nan(1,rounds); N4 = nan(1,rounds); N5 = nan(1,rounds); N6 = nan(1,rounds);
A = nan(rounds,trials);
DATA(trials,rounds).RTC = [];
DATA(trials,rounds).EO = [];
DATA(trials,rounds).EOQ = [];
DATA(trials,rounds).PD = [];
DATA(trials,rounds).PDQ = [];
DATA(trials,rounds).EPX = [];
DATA(trials,rounds).EPY = [];
DATA(trials,rounds).EPZ = [];
DATA(trials,rounds).GDX = [];
DATA(trials,rounds).GDY = [];
DATA(trials,rounds).GDZ = [];
DATA(trials,rounds).GDQ = [];

% Load different levels of difficulty
load 'Level_1';
load 'Level_2';
load 'Level_3';
Level = {Level_1 Level_2 Level_3};

na1 = randi([1,2],trials,1);
na2 = randi([1,2],trials,1);
na3 = randi([1,2],trials,1);
na = [na1 na2 na3];

Y1 = datasample(Level{1},trials,1,'Replace',false);
Y2 = datasample(Level{2},trials,1,'Replace',false);
Y3 = datasample(Level{3},trials,1,'Replace',false);
Y = {Y1 Y2 Y3};
open_system('UDPSimulinkModel')
EXPERIMENT_GUI

%% Run after experiment to save all data
% C1 = Y{Ran_D(1)}(:,1).*Y{Ran_D(1)}(:,2);
% C2 = Y{Ran_D(2)}(:,1).*Y{Ran_D(2)}(:,2);
% C3 = Y{Ran_D(3)}(:,1).*Y{Ran_D(3)}(:,2);
%
% NASA = [N1; N2; N3; N4; N5; N6];
% INFO = [PN Age Gender];
% TASK = [[Y{Ran_D(1)} C1 A(1,:)'];[Y{Ran_D(2)} C2 A(2,:)'];[Y{Ran_D(3)} C3 A(3,:)']];
%
% filename1 = ['RESULTS\DATA\DATA',num2str(PN),'.mat'];
% filename2 = ['RESULTS\NASA\NASA',num2str(PN),'.mat'];
% filename3 = ['RESULTS\INFO\INFO',num2str(PN),'.mat'];
% filename4 = ['RESULTS\TASK_ANSW\TASK',num2str(PN),'.mat'];
%
% save(filename1,'DATA');
% save(filename2,'NASA');
% save(filename3,'INFO');
% save(filename4,'TASK');
function varargout = EXPERIMENT_GUI(varargin)
% EXPERIMENT_GUI MATLAB code for EXPERIMENT_GUI.fig
% EXPERIMENT_GUI, by itself, creates a new EXPERIMENT_GUI or raises the existing
% singleton*.
% 
% H = EXPERIMENT_GUI returns the handle to a new EXPERIMENT_GUI or the handle to
% the existing singleton*.
% 
% EXPERIMENT_GUI('CALLBACK',hObject,eventData,handles,...) calls the local
% function named CALLBACK in EXPERIMENT_GUI.M with the given input arguments.
% 
% EXPERIMENT_GUI('Property','Value',...) creates a new EXPERIMENT_GUI or raises the
% existing singleton*. Starting from the left, property value pairs are
% applied to the GUI before EXPERIMENT_GUI_OpeningFcn gets called. An
% unrecognized property name or invalid value makes property application
% stop. All inputs are passed to EXPERIMENT_GUI_OpeningFcn via varargin.
% 
% *See GUI Options on GUIDE's Tools menu. Choose "GUI allows only one
% instance to run (singleton)".
% 
% See also: GUIDE, GUIDATA, GUIHANDLES
%
% Edit the above text to modify the response to help EXPERIMENT_GUI

% Last Modified by GUIDE v2.5 10-Feb-2015 11:08:20

% Begin initialization code - DO NOT EDIT

% Begin initialization code - DO NOT EDIT

% Begin initialization code - DO NOT EDIT

% Begin initialization code - DO NOT EDIT
% --- Executes just before EXPERIMENT_GUI is made visible.
function EXPERIMENT_GUI_OpeningFcn(hObject, eventdata, handles, varargin)
% This function has no output args, see OutputFcn.
% hObject    handle to figure
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
% varargin   command line arguments to EXPERIMENT_GUI (see VARARGIN)

% Choose default command line output for EXPERIMENT_GUI
handles.output = hObject;

% Update handles structure
guidata(hObject, handles);
uicontrol(handles.info1)

% UIWAIT makes EXPERIMENT_GUI wait for user response (see UIRESUME)
% uistack(hObject, 'top');
% --- Outputs from this function are returned to the command line.

function varargout = EXPERIMENT_GUI_OutputFcn(hObject, eventdata, handles)
% varargout  cell array for returning output args (see VARARGOUT);
% hObject    handle to figure
% eventdata  reserved - to be defined in a future version of MATLAB
% handles   structure with handles and user data (see GUIDATA)

% Get default command line output from handles structure
varargout{1} = handles.output;
set(hObject, 'units', 'normalized', 'outerposition', [0 0 1 1]);

function info1_Callback(hObject, eventdata, handles)
% hObject    handle to info1 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles   structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of info1 as text
%        str2double(get(hObject,'String')) returns contents of info1 as a double
PN = str2double(get(hObject, 'string'));
assignin('base', 'PN', PN)
uicontrol(handles.info2)

% --- Executes during object creation, after setting all properties.
function info1_CreateFcn(hObject, eventdata, handles)
% hObject    handle to info1 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
    get(0,'defaultUicontrolBackgroundColor'))
    set(hObject, 'BackgroundColor', 'white');
end

function info2_Callback(hObject, eventdata, handles)
% hObject    handle to info2 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of info2 as text
%        str2double(get(hObject,'String')) returns contents of info2 as a double
Age = str2double(get(hObject, 'string'));
assignin('base', 'Age', Age)
uicontrol(handles.info3)

% --- Executes during object creation, after setting all properties.
function info2_CreateFcn(hObject, eventdata, handles)
% hObject    handle to info2 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
    get(0,'defaultUicontrolBackgroundColor'))
    set(hObject, 'BackgroundColor', 'white');
end

function info3_Callback(hObject, eventdata, handles)
% hObject    handle to info3 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
% Hints: get(hObject,'String') returns contents of info3 as text
% str2double(get(hObject,'String')) returns contents of info3 as a double
Gender = str2double(get(hObject,'string'));
assignin('base', 'Gender', Gender)
uicontrol(handles.OK1)

% --- Executes during object creation, after setting all properties.
function info3_CreateFcn(hObject, eventdata, handles)
    % hObject    handle to info3 (see GCBO)
    % eventdata  reserved - to be defined in a future version of MATLAB
    % handles    empty - handles not created until after all CreateFcns called

    % Hint: edit controls usually have a white background on Windows.
    % See ISPC and COMPUTER.
    if ispc && isequal(get(hObject,'BackgroundColor'),
    get(0,'defaultUicontrolBackgroundColor'))
        set(hObject,'BackgroundColor','white');
    end

% --- Executes on button press in OK1.
function OK1_Callback(hObject, eventdata, handles)
    % hObject    handle to OK1 (see GCBO)
    % eventdata  reserved - to be defined in a future version of MATLAB
    % handles    structure with handles and user data (see GUIDATA)

    % --- Executes on key press with focus on OK1 and none of its controls.
    function OK1_KeyPressFcn(hObject, eventdata, handles)
        % hObject    handle to OK1 (see GCBO)
        % eventdata  structure with the following fields (see MATLAB.UI.CONTROL.UICONTROL)
        %   Key: name of the key that was pressed, in lower case
        %   Character: character interpretation of the key(s) that was pressed
        %   Modifier: name(s) of the modifier key(s) (i.e., control, shift) pressed
        % handles    structure with handles and user data (see GUIDATA)
        a = eventdata.Key;
        if strcmp(a,'return') == 1
            set(handles.uipanel1,'visible','off')
            set(handles.NextRound,'visible','on')
nicontrol(handles.NextRound)
        end

% Hints: get(hObject,'String') returns contents of nasal as text
% str2double(get(hObject,'String')) returns contents of nasal as a double
N1t = str2double(get(hObject,'string'));
N1 = evalin('base','N1');
i = evalin('base','i');
N1(i) = N1t;
assignin('base', 'N1', N1)
uicontrol(handles.nasa2)

% --- Executes during object creation, after setting all properties.
function nasal_CreateFcn(hObject, eventdata, handles)
    % hObject    handle to nasal (see GCBO)
    % eventdata  reserved - to be defined in a future version of MATLAB
    % handles    structure with handles and user data (see GUIDATA)

    % Hint: edit controls usually have a white background on Windows.
    % See ISPC and COMPUTER.
    if ispc && isequal(get(hObject,'BackgroundColor'),
    get(0,'defaultUicontrolBackgroundColor'))
        set(hObject,'BackgroundColor','white');
function nasa2_Callback(hObject, eventdata, handles)
% hObject handle to nasa2 (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of nasa2 as text
% str2double(get(hObject,'string')) returns contents of nasa2 as a double
N2t = str2double(get(hObject,'string'));
N2 = evalin('base','N2');
i = evalin('base','i');
N2(i) = N2t;
assignin('base', 'N2', N2)
uicontrol(handles.nasa3)

% --- Executes during object creation, after setting all properties.
function nasa2_CreateFcn(hObject, eventdata, handles)
% hObject handle to nasa2 (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
% See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function nasa3_Callback(hObject, eventdata, handles)
% hObject handle to nasa3 (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of nasa3 as text
% str2double(get(hObject,'string')) returns contents of nasa3 as a double
N3t = str2double(get(hObject,'string'));
N3 = evalin('base','N3');
i = evalin('base','i');
N3(i) = N3t;
assignin('base', 'N3', N3)
uicontrol(handles.nasa4)

% --- Executes during object creation, after setting all properties.
function nasa3_CreateFcn(hObject, eventdata, handles)
% hObject handle to nasa3 (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
% See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function nasa4_Callback(hObject, eventdata, handles)
% hObject handle to nasa4 (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of nasa4 as text
% str2double(get(hObject,'string')) returns contents of nasa4 as a double
N4t = str2double(get(hObject,'string'));
N4 = evalin('base','N4');
i = evalin('base','i');
N4(i) = N4t;
assignin('base', 'N4', N4)
uicontrol(handles.nasa5)

% --- Executes during object creation, after setting all properties.
function nasa4_CreateFcn(hObject, eventdata, handles)

% hObject    handle to nasa4 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
% See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function nasa5_Callback(hObject, eventdata, handles)
% hObject    handle to nasa5 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of nasa5 as text
%        str2double(get(hObject,'String')) returns contents of nasa5 as a double
N5t = str2double(get(hObject,'string'));
N5 = evalin('base','N5');
i = evalin('base','i');
N5(i) = N5t;
assignin('base', 'N5', N5)
uicontrol(handles.nasa6)

% --- Executes during object creation, after setting all properties.
function nasa5_CreateFcn(hObject, eventdata, handles)

% hObject    handle to nasa5 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: edit controls usually have a white background on Windows.
% See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

function nasa6_Callback(hObject, eventdata, handles)
% hObject    handle to nasa6 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'String') returns contents of nasa6 as text
%        str2double(get(hObject,'String')) returns contents of nasa6 as a double
N6t = str2double(get(hObject,'string'));
N6 = evalin('base','N6');
i = evalin('base','i');
N6(i) = N6t;
assignin('base', 'N6', N6)
uicontrol(handle.OK2)

% --- Executes during object creation, after setting all properties.
function nasa6_CreateFcn(hObject, eventdata, handles)

% hObject    handle to nasa6 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called
% Hint: edit controls usually have a white background on Windows.
% See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end

% --- Executes on button press in OK2.
function OK2_Callback(hObject, eventdata, handles)
% hObject handle to OK2 (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)
% --- Executes on key press with focus on OK2 and none of its controls.
function OK2_KeyPressFcn(hObject, eventdata, handles)
% hObject handle to OK2 (see GCBO)
% eventdata structure with the following fields (see MATLAB.UI.CONTROL.UICONTROL)
%   Key: name of the key that was pressed, in lower case
%   Character: character interpretation of the key(s) that was pressed
%   Modifier: name(s) of the modifier key(s) (i.e., control, shift) pressed
% handles structure with handles and user data (see GUIDATA)
b = eventdata.Key;
if strcmp(b,'return') == 1
    set(handles.uipanel2,'visible','off');
    set(handles.nasa1,'String','');
    set(handles.nasa2,'String','');
    set(handles.nasa3,'String','');
    set(handles.nasa4,'String','');
    set(handles.nasa5,'String','');
    set(handles.nasa6,'String','');
    i = evalin('base','i');
    rounds = evalin('base','rounds');
    if i < rounds
        set(handles.NextRound,'visible','on');
        uicontrol(handles.NextRound)
    else
        close(EXPERIMENT_GUI)
    end
end

function answer1_Callback(hObject, eventdata, handles)
% hObject handle to answer1 (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)
% Hints: get(hObject,'String') returns contents of answer1 as text
% str2double(get(hObject,'String')) returns contents of answer1 as a double
At = str2double(get(hObject,'string'));
i = evalin('base','i');
j = evalin('base','j');
A = evalin('base','A');
A(i,j) = At;
assignin('base', 'A', A);
uicontrol(handles.OK3)

% --- Executes during object creation, after setting all properties.
function answer1_CreateFcn(hObject, eventdata, handles)
% hObject handle to answer1 (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles empty - handles not created until after all CreateFcns called
% Hint: edit controls usually have a white background on Windows.
% See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
set(hObject,'BackgroundColor','white');
end

% --- Executes on button press in OK3.
function OK3_Callback(hObject, eventdata, handles)
    % hObject    handle to OK3 (see GCBO)
    % eventdata  reserved - to be defined in a future version of MATLAB
    % handles    structure with handles and user data (see GUIDATA)

    % --- Executes on key press with focus on OK3 and none of its controls.
    function OK3_KeyPressFcn(hObject, eventdata, handles)
        % hObject    handle to OK3 (see GCBO)
        % eventdata  structure with the following fields (see MATLAB.UI.CONTROL.UICONTROL)
        %    Key: name of the key that was pressed, in lower case
        %    Character: character interpretation of the key(s) that was pressed
        %    Modifier: name(s) of the modifier key(s) (i.e., control, shift) pressed
        % handles    structure with handles and user data (see GUIDATA)
        b = eventdata.Key;
        if strcmp(b,'return') == 1
            set(handles.uipanel3,'visible','off')
            set(handles.answer1,'String','');
            j = evalin('base','j');
            j = j+1;
            trials = evalin('base','trials');
            if j <= trials
                set(handles.NextTrial,'visible','on')
                assignin('base','j',j)
                uicontrol(handles.NextTrial)
            else
                set(handles.uipanel2,'visible','on')
                uicontrol(handles.nasa1)
            end
        end
    end

function NextRound_Callback(hObject, eventdata, handles)
    % hObject    handle to OK3 (see GCBO)
    % eventdata  reserved - to be defined in a future version of MATLAB
    % handles    structure with handles and user data (see GUIDATA)

function NextTrial_Callback(hObject, eventdata, handles)
    % hObject    handle to OK3 (see GCBO)
    % eventdata  reserved - to be defined in a future version of MATLAB
    % handles    structure with handles and user data (see GUIDATA)

    % --- Executes on key press with focus on NextRound and none of its controls.
    function NextRound_KeyPressFcn(hObject, eventdata, handles)
        % hObject    handle to NextRound (see GCBO)
        % eventdata  structure with the following fields (see MATLAB.UI.CONTROL.UICONTROL)
        %    Key: name of the key that was pressed, in lower case
        %    Character: character interpretation of the key(s) that was pressed
        %    Modifier: name(s) of the modifier key(s) (i.e., control, shift) pressed
        % handles    structure with handles and user data (see GUIDATA)
        b = eventdata.Key;
        if strcmp(b,'return') == 1
            set(handles.NextRound,'visible','off')
            set(handles.NextTrial,'visible','on')
            i = evalin('base','i'); % Trial number
            i = i+1;
            j = 1;
            assignin('base','i',i)
            assignin('base','j',j)
        end
        uicontrol(handles.NextTrial)
    end

    % --- Executes on key press with focus on NextTrial and none of its controls.

function NextTrial_KeyPressFcn(hObject, eventdata, handles)
% hObject    handle to NextTrial (see GCBO)
% eventdata  structure with the following fields (see MATLAB.UI.CONTROL.UICONTROL)
%   Key: name of the key that was pressed, in lower case
%   Character: character interpretation of the key(s) that was pressed
%   Modifier: name(s) of the modifier key(s) (i.e., control, shift) pressed
% handles    structure with handles and user data (see GUIDATA)

b = eventdata.Key;
if strcmp(b,'return') == 1
    set(handles.NextTrial,'visible','off')
    i = evalin('base','i'); % Round number
    j = evalin('base','j'); % Trial number
    Y = evalin('base','Y');
    na = evalin('base','na');
    Ran_D = evalin('base','Ran_D');
    a = Y{Ran_D(i)}(j,na{i}(j));
    if na{i}(j) == 1
        b = Y{Ran_D(i)}(j,2);
    else
        b = Y{Ran_D(i)}(j,1);
    end

    AccommodationTime = 4.0; % Duration of baseline pause
    PresentationTime  = 1.0; % Presentation duration of the numbers
    InterPause        = 1.5; % Duration of intermediate pause
    CalculationTime   = 10.0; % Calculation Time

    clear EO EOQ PD

    handles.textxx = text(0.482,0.5,'XX','FontSize',40,'units','normalized');

    % set_param('UDPSimulinkModel','SimulationCommand','Start')
    pause(AccommodationTime)
    delete(handles.textxx);

    if a >= 10
        handles.textxx = text(0.482,0.5,num2str(a),'FontSize',40,'units','normalized');
    else
        handles.textxx = text(0.482,0.5,['0' num2str(a)],'FontSize',40,'units','normalized');
    end
    pause(PresentationTime) % Number 1
    delete(handles.textxx);

    handles.textxx = text(0.482,0.5,'XX','FontSize',36,'units','normalized');
    pause(InterPause)
    delete(handles.textxx);

    if b >= 10
        handles.textxx = text(0.482,0.5,num2str(b),'FontSize',40,'units','normalized');
    else
        handles.textxx = text(0.482,0.5,['0' num2str(b)],'FontSize',40,'units','normalized');
    end
    pause(PresentationTime) % Number 2
    delete(handles.textxx);

    handles.textxx = text(0.482,0.5,'XX','FontSize',36,'units','normalized');
    pause(CalculationTime)
    % set_param('UDPSimulinkModel','SimulationCommand','Stop')
    delete(handles.textxx);

    % OUT = evalin('base','OUT');
OUT2 = double(squeeze(OUT.Data));
for k = 1:size(OUT2,2)
    RTC(k) = swapbytes(typecast(uint8(OUT2(65:72,k)),'uint64'));
    EO(k) = swapbytes(typecast(uint8(OUT2(145:152,k)),'double'));
    EOQ(k) = swapbytes(typecast(uint8(OUT2(157:164,k)),'double'));
    PD(k) = swapbytes(typecast(uint8(OUT2(169:176,k)),'double'));
    PDQ(k) = swapbytes(typecast(uint8(OUT2(181:188,k)),'double'));
    EFX(k) = swapbytes(typecast(uint8(OUT2(77:84,k)),'double'));
    EY(k) = swapbytes(typecast(uint8(OUT2(85:92,k)),'double'));
    EPX(k) = swapbytes(typecast(uint8(OUT2(93:100,k)),'double'));
    EPY(k) = swapbytes(typecast(uint8(OUT2(105:112,k)),'double'));
    EPZ(k) = swapbytes(typecast(uint8(OUT2(113:120,k)),'double'));
    GDX(k) = swapbytes(typecast(uint8(OUT2(121:128,k)),'double'));
    GDY(k) = swapbytes(typecast(uint8(OUT2(133:140,k)),'double'));
    GDZ(k) = swapbytes(typecast(uint8(OUT2(141:148,k)),'double'));
end

DATA = evalin('base','DATA');
DATA(j,i).RTC = RTC;
DATA(j,i).EO  = EO;
DATA(j,i).EOQ = EOQ;
DATA(j,i).PD  = PD;
DATA(j,i).PDQ = PDQ;
DATA(j,i).EPX = EFX;
DATA(j,i).EY  = EY;
DATA(j,i).EP2 = EP2;
DATA(j,i).GDX = GDX;
DATA(j,i).GDY = GDY;
DATA(j,i).GDZ = GDZ;
DATA(j,i).GDQ = GDQ;
assignin('base', 'DATA', DATA)

set(handles.uipanel3,'visible','on')
uicontrol(handles.answer1)
end

--- Executes during object creation, after setting all properties.
function axes4_CreateFcn(hObject, eventdata, handles)
    % hObject    handle to axes4 (see GCBO)
    % eventdata  reserved - to be defined in a future version of MATLAB
    % handles    empty - handles not created until after all CreateFcns called

    % Hint: place code in OpeningFcn to populate axes4
    axes(hObject)
imshow('Backgrounds/VarBr7_2.png')
%% ANALYSIS MENTAL MULTIPLICATION EXPERIMENT
% Gerhard Marquart
% #1538446

clear all; close all; clc;

%% LOAD DATA
n = 30; % Number of participants
for i = 1:n
    load(['RESULTS\DATA\DATA' num2str(i) '.mat']); Data(i).D = DATA; clear DATA
    load(['RESULTS\INFO\INFO' num2str(i) '.mat']); Info(i).I = INFO; clear INFO
    load(['RESULTS\NASA\NASA' num2str(i) '.mat']); Nasa(i).N = NASA; clear NASA
    load(['RESULTS\TASK\TASK' num2str(i) '.mat']); Task(i).T = TASK; clear TASK
end

load('Diff.mat');
Color = {{[0.0 0.45 0.74] [0.85 0.33 0.10] [0.93 0.69 0.13]};

% X coordinates of # points
XPoints = 120.*[6.5 6.5 6.5; 7.5 7.5 7.5; 9.0 9.0 9.0; 10.5 10.5 10.5; 12.0 12.0 12.0; 13.5 13.5 13.5; 15.0 15.0 15.0; 16.5 16.5 16.5];
NPoints = size(XPoints,1);

%% DATA PROCESSING
L1 = nan(15*n,2200); L2 = nan(15*n,2200); L3 = nan(15*n,2200);
L = {L1 L2 L3}; LSmooth = {L1 L2 L3}; LDiff1 = {L1 L2 L3}; LDiff2 = {L1 L2 L3}; LDiff3 = {L1 L2 L3}; LCorrect = {L1 L2 L3}; LIncorrect = {L1 L2 L3}; LEO = {L1 L2 L3}; LPDQ = {L1 L2 L3};
LS1 = {L1 L2 L3}; LS2 = {L1 L2 L3}; LS3 = {L1 L2 L3}; LS4 = {L1 L2 L3}; LS5 = {L1 L2 L3}; LS6 = {L1 L2 L3};
clear L1 L2 L3
l1 = 1; l2 = 1; l3 = 1;
l = {l1 l2 l3}; clear l1 l2 l3
Freq = 120;
Time = 17.5;
coeff = ones(1, Freq/4)/(Freq/4);
Delay = (length(coeff)-1)/2;
PERC = nan(30,15,3);
for i = 1:30
    for j = 1:15
        for k = 1:3
            clear PD PD1 EO PDQ
            EO = Data(i).D(j,k).EO;
            EO(EO==0)=NaN;
            PD = Data(i).D(j,k).PD;
            LS1(Diff(i,k)){1} = Data(i).D(j,k).PD;
            PD(PD==0)=NaN;
            LS2(Diff(i,k)){1} = Data(i).D(j,k).PD;
            PDQ = Data(i).D(j,k).PDQ;
            if sum(~isnan(PD)) >= 0.95*17.5*120
84

EO(:, isnan(EO)) = interp1(find(~isnan(EO)), EO(:, ~isnan(EO)), find(isnan(EO)), 'linear');
PD(:, isnan(PD)) = interp1(find(~isnan(PD)), PD(:, ~isnan(PD)), find(isnan(PD)), 'linear');
PDQ(:, isnan(PDQ)) = interp1(find(~isnan(PDQ)), PDQ(:, ~isnan(PDQ)), find(isnan(PDQ)), 'linear');
end
LS3{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
eolim = 0.75*nanmean(EO);
PD = blinkremoval(PD,EO,PDQ,0.75,eolim);
LS4{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
if sum(~isnan(PD)) >= 0.70*17.5*120
PD(:, isnan(PD)) = interp1(find(~isnan(PD)), PD(:, ~isnan(PD)), find(isnan(PD)), 'linear');
LS5{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
PD = PD - nanmean(PD(3.6*120:4*120));
LS6{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
L{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
LEO{Diff(i,k)}(l{Diff(i,k)},1:length(EO)) = EO;
LPDQ{Diff(i,k)}(l{Diff(i,k)},1:length(PDQ)) = PDQ;
TEMP_Smooth = filter(coeff,1,PD);
LSmooth{Diff(i,k)}(l{Diff(i,k)},1:(length(PD)-(Freq/8)+1)) = TEMP_Smooth(Freq/8:end);
LDiff1{Diff(i,k)}(l{Diff(i,k)},1:length(PD)-(Freq/8:end))./(1/120);
if Task(i).T(((k-1)*15+j),3) == Task(i).T(((k-1)*15+j),4)
LCorrect{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
else
LIncorrect{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
end
PD3 = LS2{Diff(i,k)}(l{Diff(i,k)},1:length(PD));
PERC(i,j,k) = (sum(isnan(PD3(1:2088)))/2088)*1e2;
end
l{Diff(i,k)} = l{Diff(i,k)}+1;
end
end
Per_Level_1 = size(LCorrect{1}(any(~isnan(LCorrect{1}),2),:),1) / size(L{1}(any(~isnan(L{1}),2),:),1);
Per_Level_2 = size(LCorrect{2}(any(~isnan(LCorrect{2}),2),:),1) / size(L{2}(any(~isnan(L{2}),2),:),1);
Per_Level_3 = size(LCorrect{3}(any(~isnan(LCorrect{3}),2),:),1) / size(L{3}(any(~isnan(L{3}),2),:),1);

%% Show Data Processing Steps
figure(12); hold on;
subplot(1,2,1);
plot(linspace(0,length(LS1{1}(22,:))/120,length(LS1{1}(22,:))),(LS1{1}(22,:)).*1000,'Color',Color{1},'LineWidth',1); xlim([12 15]); ylim([3.2 4.6]); hold on;
plot(linspace(0,length(LS3{1}(22,:))/120,length(LS3{1}(22,:))),(LS3{1}(22,:)).*1000,'Color',Color{2},'LineWidth',1); hold on;
xlabel('Time (s)'); ylabel('PD (mm)'); legend('Before', 'After'); %title('Linear Interpolation for Missing Values');
subplot(1,2,2);
plot(linspace(0,length(LS3{1}(22,:))/120,length(LS3{1}(22,:))),(LS3{1}(22,:)).*1000,'Color',Color{1},'LineWidth',1); xlim([12 15]); ylim([3.2 4.6]); hold on;
plot(linspace(0,length(LS5{1}(22,:))/120,length(LS5{1}(22,:))),(LS5{1}(22,:)).*1000,'Color',Color{2},'LineWidth',1); hold on;

84
xlabel('Time (s)'); ylabel('PD (mm)'); legend('Before', 'After'); %title('Linear Interpolation for Blinks and Poor Quality Data');

%% Show Data Processing Steps invidiual trials
close all; figure(13); hold on;
IT = 155; %22 %156 %181 360
LT = 3; %1 %2 %3 %3
subplot(3,1,1);
rectangle('Position', [4, -1, 1, 10], 'EdgeColor', [0.6 0.6 0.6], 'FaceColor', [0.9 0.9 0.9]); hold on;
rectangle('Position', [6.5, -1, 1, 10], 'EdgeColor', [0.6 0.6 0.6], 'FaceColor', [0.9 0.9 0.9]); hold on;
plot(linspace(0,length(LS1{LT}(IT,:))/120,length(LS1{LT}(IT,:))),(LS1{LT}(IT,:).*1000, 'Color', Color{1}, 'LineWidth', 1); hold on;
plot(linspace(0,length(LS5{LT}(IT,:))/120,length(LS5{LT}(IT,:))),(LS5{LT}(IT,:).*1000, 'Color', Color{2}, 'LineWidth', 1);
xlabel('Time (s)'); ylabel('PD (mm)'); legend('Before processing', 'After processing', 'Location', 'northwest'); %title('Linear Interpolation for Blinks and Poor Quality Data');
xlim([0 17.5]); ylim([3.0 5.0]); hold on;

subplot(3,1,2);
rectangle('Position', [4, -1, 1, 10], 'EdgeColor', [0.6 0.6 0.6], 'FaceColor', [0.9 0.9 0.9]); hold on;
rectangle('Position', [6.5, -1, 1, 10], 'EdgeColor', [0.6 0.6 0.6], 'FaceColor', [0.9 0.9 0.9]); hold on;
plot(linspace(0,length(LPDQ{LT}(IT,:))/120,length(LPDQ{LT}(IT,:))),(LPDQ{LT}(IT,:), 'Color', Color{1}, 'LineWidth', 1);
xlabel('Time (s)'); ylabel('PD Quality (mm)');
xlim([0 17.5]); ylim([0 1]); hold on;

subplot(3,1,3);
rectangle('Position', [4, -1, 1, 100], 'EdgeColor', [0.6 0.6 0.6], 'FaceColor', [0.9 0.9 0.9]); hold on;
rectangle('Position', [6.5, -1, 1, 100], 'EdgeColor', [0.6 0.6 0.6], 'FaceColor', [0.9 0.9 0.9]); hold on;
plot(linspace(0,length(LEO{LT}(IT,:))/120,length(LEO{LT}(IT,:))),(LEO{LT}(IT,:).*1000, 'Color', Color{1}, 'LineWidth', 1);
xlabel('Time (s)'); ylabel('Eyelid Opening (mm)');
xlim([0 17.5]); ylim([0 25]); hold on;
h = line([0 17.5], [0.75*nanmean((LEO{LT}(IT,:)).*1000) 0.75*nanmean((LEO{LT}(IT,:)).*1000)], 'Color', 'r', 'LineStyle', '--');
legend('Eyelid opening', 'Threshold', 'Location', 'northwest');

%% Cohen’s Effect Size
for i = 0:n-1
    for k = 1:3
        Ln{k}(i+1,:) = nanmean(L{k}(i*15+1:(i+1)*15,:),1);
        LCorrectn{k}(i+1,:) = nanmean(LCorrect{k}(i*15+1:(i+1)*15,:),1);
        LIncorrectn{k}(i+1,:) = nanmean(LIncorrect{k}(i*15+1:(i+1)*15,:),1);
    end
end

for k = 1:3
    Lm{k} = Ln{k}([1:8 10:30],:).*1000; % If neccessary, remove participant here
    M_Lm{k} = nanmean(Lm{k});
    SD_Lm{k} = nanstd(Lm{k});
end
for k = 1:3
    for i = 1:size(Lm{1},2)
        if k <= 2
            r_Lm{k}(i) = corr2(Lm{k}(:,i),Lm{k+1}(:,i));
        else
            k == 3
            break
        end
    end
end
\[
\text{r}_{\text{Lm}}(i) = \text{corr2}(L_{\text{m}2}(:,i), L_{\text{m}1}(i,:));
\]

\[
\text{Cohensdz}_{\text{Lm}1}(i) = \frac{|\text{M}_{\text{Lm}1}(i) - \text{M}_{\text{Lm}2}(i)|}{\sqrt{\text{SD}_{\text{Lm}1}(i)^2 + \text{SD}_{\text{Lm}2}(i)^2 - 2 \text{r}_{\text{Lm}1}(i) \text{SD}_{\text{Lm}1}(i) \text{SD}_{\text{Lm}2}(i)});
\]

\[
\text{Cohensdz}_{\text{Lm}2}(i) = \frac{|\text{M}_{\text{Lm}2}(i) - \text{M}_{\text{Lm}3}(i)|}{\sqrt{\text{SD}_{\text{Lm}2}(i)^2 + \text{SD}_{\text{Lm}3}(i)^2 - 2 \text{r}_{\text{Lm}2}(i) \text{SD}_{\text{Lm}2}(i) \text{SD}_{\text{Lm}3}(i)});
\]

\[
\text{Cohensdz}_{\text{Lm}3}(i) = \frac{|\text{M}_{\text{Lm}1}(i) - \text{M}_{\text{Lm}3}(i)|}{\sqrt{\text{SD}_{\text{Lm}1}(i)^2 + \text{SD}_{\text{Lm}3}(i)^2 - 2 \text{r}_{\text{Lm}3}(i) \text{SD}_{\text{Lm}1}(i) \text{SD}_{\text{Lm}3}(i)});
\]

for \(i = 1: \text{size}(L_{\text{m}1},2)
\]

\[
\text{Cohensdz}_{\text{Lm}1}(i) = \frac{|\text{M}_{\text{Lm}1}(i) - \text{M}_{\text{Lm}2}(i)|}{\sqrt{\text{SD}_{\text{Lm}1}(i)^2 + \text{SD}_{\text{Lm}2}(i)^2 - 2 \text{r}_{\text{Lm}1}(i) \text{SD}_{\text{Lm}1}(i) \text{SD}_{\text{Lm}2}(i)});
\]

\[
\text{Cohensdz}_{\text{Lm}2}(i) = \frac{|\text{M}_{\text{Lm}2}(i) - \text{M}_{\text{Lm}3}(i)|}{\sqrt{\text{SD}_{\text{Lm}2}(i)^2 + \text{SD}_{\text{Lm}3}(i)^2 - 2 \text{r}_{\text{Lm}2}(i) \text{SD}_{\text{Lm}2}(i) \text{SD}_{\text{Lm}3}(i)});
\]

\[
\text{Cohensdz}_{\text{Lm}3}(i) = \frac{|\text{M}_{\text{Lm}1}(i) - \text{M}_{\text{Lm}3}(i)|}{\sqrt{\text{SD}_{\text{Lm}1}(i)^2 + \text{SD}_{\text{Lm}3}(i)^2 - 2 \text{r}_{\text{Lm}3}(i) \text{SD}_{\text{Lm}1}(i) \text{SD}_{\text{Lm}3}(i)});
\]

\[
\text{figure}(1);
\]

\[
% \text{title('Effect size Cohen''s d_{z}','FontSize',12);
\]

\[
\text{rectangle('Position',[4,-1,1,10],'EdgeColor',[0.6 0.6 0.6],'FaceColor',[0.9 0.9 0.9]);
\]

\[
\text{hold on};
\]

\[
\text{rectangle('Position',[6.5,-1,1,10],'EdgeColor',[0.6 0.6 0.6],'FaceColor',[0.9 0.9 0.9]);
\]

\[
\text{hold on};
\]

\[
\text{h1 = text(4.4, 1.40, 'Multiplicand','FontSize',12); set(h1, 'rotation', 90)
\]

\[
\text{h2 = text(6.9, 1.40, 'Multiplier' ,'FontSize',12); set(h2, 'rotation', 90)
\]

\[
\text{for } k = 1:3
\]

\[
\text{plot(linspace(0,length(Cohensdz_{Lm}k)/120,length(Cohensdz_{Lm}k))),(Cohensdz_{Lm}k),'Color',Color{k},'LineWidth',1); hold on;
\]

\[
\text{xlim([0 17.5])
\]

\[
\text{ylim([0 2])
\]

\[
\text{end
\]

\[
\text{xlabel('Time (s)');
\]

\[
\text{ylabel('Cohen''s d_{z}');
\]

\[
\text{legend('Level 1-2','Level 2-3','Level 1-3','Location','NorthWest');
\]

\[
\text{set(gca,'XTick',0:1:18)
\]

\[
\text{figure(1);
\]

\[
% \text{title('Mean Pupil Diameter Change - Level 3 - Correct VS Incorrect Responses','FontSize',12);
\]

\[
\text{rectangle('Position',[4,-1,1,10],'EdgeColor',[0.6 0.6 0.6],'FaceColor',[0.9 0.9 0.9]);
\]

\[
\text{hold on};
\]

\[
\text{rectangle('Position',[6.5,-1,1,10],'EdgeColor',[0.6 0.6 0.6],'FaceColor',[0.9 0.9 0.9]);
\]

\[
\text{hold on};
\]

\[
\text{h1 = text(4.4, 0.20, 'Multiplicand','FontSize',12); set(h1, 'rotation', 90)
\]

\[
\text{h2 = text(6.9, 0.20, 'Multiplier' ,'FontSize',12); set(h2, 'rotation', 90)
\]

\[
\text{for } k = 3
\]

\[
\text{plot(linspace(0,length(Ln{k})/120,length(Ln{k}))),(nanmean(Ln{k}).*1000,'Color',Color{k},'LineWidth',1); hold on;
\]

\[
\text{plot(linspace(0,length(LCorrectn{k})/120,length(LCorrectn{k}))),(nanmean(LCorrectn{k}).*1000,'g','LineWidth',1); hold on;
\]

\[
\text{plot(linspace(0,length(LIncorrectn{k})/120,length(LIncorrectn{k}))),(nanmean(LIncorrectn{k}).*1000,'r','LineWidth',1); hold on;
\]

\[
\text{xlabel('Time (s)');
\]

\[
\text{ylabel('MPDC (mm)');
\]

\[
\text{legend('Level 3 All','Level 3 Correct (}'
\]

\[
\text{sprintf('%.1f',size(LCorrect{1,k}(any(~isnan(LCorrect{1,k}),2)),:),1).*size(L{1,k}(any(~isnan(L{1,k}),2)),:),1).*100) '}%')','Level 3 Incorrect ('
\]

\[
\text{sprintf('%.1f',size(LIncorrect{1,k}(any(~isnan(LIncorrect{1,k}),2)),:),1).*size(L{1,k}(any(~isnan(L{1,k}),2)),:),1).*100) '}%')
\]

\[
\text{),'Location','NorthEast');
\]
%% Pupil Diameter Change With and Without Filter
figure(3);
% title('Pupil Diameter Change With and Without Filter','FontSize',12);
rectangle('Position',[4,-1,1,10],'EdgeColor',[0.6 0.6 0.6],'FaceColor',[0.9 0.9 0.9]);
hold on;
rectangle('Position',[6.5,-1,1,10],'EdgeColor',[0.6 0.6 0.6],'FaceColor',[0.9 0.9 0.9]);
h1 = text(4.4, 0.25, 'Multiplicand','FontSize',12); set(h1, 'rotation', 90)
h2 = text(6.9, 0.25, 'Multiplier','FontSize',12); set(h2, 'rotation', 90)
plot(linspace(0,length(L{1}(22,:))/120,length(L{1}(22,:))),L{1}(22,:).*1000,'k','LineWidth',1);
hold on;
plot(linspace(0,length(LSmooth{1}(22,:))/120,length(LSmooth{1}(22,:))),LSmooth{1}(22,:).*1000,'r','LineWidth',1)
xlim([0 17.5])
ylim([-0.9 0.7])
xlabel('Time (s)');
ylabel('MPDC (mm)');
legend('Unfiltered Trial','Filtered Trial');
set(gca,'XTick',0:1:18)
%
figure(4);
% title('Mean Pupil Diameter Change per Level','FontSize',12);
rectangle('Position',[4,-1,1,10],'EdgeColor',[0.6 0.6 0.6],'FaceColor',[0.9 0.9 0.9]);
hold on;
rectangle('Position',[6.5,-1,1,10],'EdgeColor',[0.6 0.6 0.6],'FaceColor',[0.9 0.9 0.9]);
h1 = text(4.4, 0.20, 'Multiplicand','FontSize',12); set(h1, 'rotation', 90)
h2 = text(6.9, 0.20, 'Multiplier','FontSize',12); set(h2, 'rotation', 90)
for k = 1:3
plot(linspace(0,length(Ln{k})/120,length(Ln{k})),nanmean(Ln{k}).*1000,'Color',Color{k},'LineWidth',1);
hold on;
xlim([0 17.5])
ylim([-0.2 0.4])
xlabel('Time (s)');
ylabel('MPDC (mm)');
legend('Level 1','Level 2','Level 3','Location','NorthEast')
set(gca,'XTick',0:1:18)
end
L4 = nan(size(Ln{1})); L5 = nan(size(Ln{2})); L6 = nan(size(Ln{3}));
LL = {L4 L5 L6}; clear L4 L5 L6
for k = 1:3
TEMP1 = (Ln{k}-nanmean(Ln{1}(:,6.5*120))).*1000;
LL{k} = nanmean(TEMP1(:,1:17.5*120));
clear TEMP1
end
%
figure(5);
% title('Mean Pupil Diameter Velocity per Level','FontSize',12);
rectangle('Position',[4,-1,1,10],'EdgeColor',[0.6 0.6 0.6],'FaceColor',[0.9 0.9 0.9]);
hold on;
rectangle('Position',[6.5,-1,1,10],'EdgeColor',[0.6 0.6 0.6],'FaceColor',[0.9 0.9 0.9]);
hold on;
for i = 0:n-1
    for k = 1:3
        LDiffn1{k}(i+1,:) = nanmean(LDiff1{k}(i*15+1:(i+1)*15,:),1);
    end
end

87
h1 = text(4.4, 0.14, 'Multiplicand', 'FontSize', 12); set(h1, 'rotation', 90); % 0.28 ipv
h2 = text(6.9, 0.14, 'Multiplier', 'FontSize', 12); set(h2, 'rotation', 90); % 0.28 ipv
for k = 1:3
    plot(linspace(0,length(LDiffn1{k})/120,length(LDiffn1{k})),nanmean(LDiffn1{k}).*1000, 'Color', Color(k), 'LineWidth', 1); hold on;
    xlim([0 17.5]); ylim([-0.15 0.25]); %ylim([0.15 0.35]);
end
xlabel('Time (s)');
ylabel('MPDV (mm/s)'); %ylabel('Abs MPDV (mm/s)');
legend('Level 1', 'Level 2', 'Level 3', 'Location', 'NorthEast')
set(gca, 'XTick', 0:1:18)

%% Mean Pupil Diameter Change per Level, 2
figure(6);
% title('Mean Pupil Diameter Change per Level', 'FontSize', 12);
rectangle('Position', [4, -1, 1, 2], 'EdgeColor', [0.6 0.6 0.6], 'FaceColor', [0.9 0.9 0.9]); hold on;
rectangle('Position', [6.5, -1, 1, 2], 'EdgeColor', [0.6 0.6 0.6], 'FaceColor', [0.9 0.9 0.9]); hold on;
rectangle('Position', [7.5, -1, 1.5, 2], 'EdgeColor', [0.8 0.8 0.8], 'FaceColor', [0.95 0.95 0.95]); hold on;
rectangle('Position', [9.0, -1, 1.5, 2], 'EdgeColor', [0.8 0.8 0.8], 'FaceColor', [0.95 0.95 0.95]); hold on;
rectangle('Position', [10.5, -1, 1.5, 2], 'EdgeColor', [0.8 0.8 0.8], 'FaceColor', [0.95 0.95 0.95]); hold on;
rectangle('Position', [12.0, -1, 1.5, 2], 'EdgeColor', [0.8 0.8 0.8], 'FaceColor', [0.95 0.95 0.95]); hold on;
rectangle('Position', [13.5, -1, 1.5, 2], 'EdgeColor', [0.8 0.8 0.8], 'FaceColor', [0.95 0.95 0.95]); hold on;
rectangle('Position', [15.0, -1, 1.5, 2], 'EdgeColor', [0.8 0.8 0.8], 'FaceColor', [0.95 0.95 0.95]); hold on;
h1 = text(4.4, 0.20, 'Multiplicand', 'FontSize', 12); set(h1, 'rotation', 90)
h2 = text(6.9, 0.20, 'Multiplier', 'FontSize', 12); set(h2, 'rotation', 90)
P1 = text(6.90, -0.08, '(1)', 'FontSize', 12);
P2 = text(8.15, -0.08, '(2)', 'FontSize', 12);
P3 = text(9.65, -0.08, '(3)', 'FontSize', 12);
P4 = text(11.15, -0.08, '(4)', 'FontSize', 12);
P5 = text(12.65, -0.08, '(5)', 'FontSize', 12);
P6 = text(14.15, -0.08, '(6)', 'FontSize', 12);
P7 = text(15.65, -0.08, '(7)', 'FontSize', 12);
for k = 1:3
    plot(linspace(0,length(LL{k})/120,length(LL{k})),LL{k},'Color', Color(k), 'LineWidth', 1); hold on;
    xlim([6.5 17.5])
    ylim([-0.10 0.35])
end
xlabel('Time (s)');
ylabel('MPDC (mm)');
set(gca, 'XTick', 0:0.5:17.5)
plot(0,0,'xk','LineWidth',1.5);
plot(0,0,'ok','LineWidth',1.5);
plot(0,0,'^k','LineWidth',1.5);
plot(0,0,'*k','LineWidth',1.5);
plot(0,0,'>k','LineWidth',1.5);
plot(0,0,'+k','LineWidth',1.5);
plot(0,0,'<k','LineWidth',1.5);
plot(0,0,'sk','LineWidth',1.5);
for k = 1:3
    plot(XPoints(1,k)/120,LL{k}(XPoints(1,k)+1),'x','Color', Color(k), 'LineWidth', 2);
    plot(XPoints(2,k)/120,LL{k}(XPoints(2,k)),'o','Color', Color(k), 'LineWidth', 2);
plot(XPoints(3,k)./120,LL{k}(XPoints(3,k)), '^', 'Color', Color{k}, 'LineWidth', 2);
plot(XPoints(4,k)./120,LL{k}(XPoints(4,k)), 'o', 'Color', Color{k}, 'LineWidth', 2);
plot(XPoints(5,k)./120,LL{k}(XPoints(5,k)), '>', 'Color', Color{k}, 'LineWidth', 2);
plot(XPoints(6,k)./120,LL{k}(XPoints(6,k)), '+', 'Color', Color{k}, 'LineWidth', 2);
plot(XPoints(7,k)./120,LL{k}(XPoints(7,k)), '<', 'Color', Color{k}, 'LineWidth', 2);
plot(XPoints(8,k)./120,LL{k}(XPoints(8,k)), 's', 'Color', Color{k}, 'LineWidth', 2);
end

legend('Level 1', 'Level 2', 'Level 3', 'P1', 'P2', 'P3', 'P4', 'P5', 'P6', 'P7', 'P8', 'Location', 'NorthEastoutside')

%% Mean Pupil Diameter Change - Bar Graph
S1 = size(Ln{1},1); S2 = size(Ln{2},1); S3 = size(Ln{3},1);
PD_Points = nan(max([S1 S2 S3]),3,NPoints+1);
for k = 1:3
    for j = 1:NPoints
        for i = 1:size(Ln{k},1);
            PD_Points(i,k,j) = (Ln{k}(i,XPoints(j,k)) - nanmean(Ln{1}(:,6.5*120))).*1000;
        end
    end
end

PD_Point1 = PD_Points(:,:,1); PD_Point1 = PD_Point1(~any(isnan(PD_Point1),2),:);
PD_Point2 = PD_Points(:,:,2); PD_Point2 = PD_Point2(~any(isnan(PD_Point2),2),:);
PD_Point3 = PD_Points(:,:,3); PD_Point3 = PD_Point3(~any(isnan(PD_Point3),2),:);
PD_Point4 = PD_Points(:,:,4); PD_Point4 = PD_Point4(~any(isnan(PD_Point4),2),:);
PD_Point5 = PD_Points(:,:,5); PD_Point5 = PD_Point5(~any(isnan(PD_Point5),2),:);
PD_Point6 = PD_Points(:,:,6); PD_Point6 = PD_Point6(~any(isnan(PD_Point6),2),:);
PD_Point7 = PD_Points(:,:,7); PD_Point7 = PD_Point7(~any(isnan(PD_Point7),2),:);
PD_Point8 = PD_Points(:,:,8); PD_Point8 = PD_Point8(~any(isnan(PD_Point8),2),:);

Name_PD = {'P1', 'P2', 'P3', 'P4', 'P5', 'P6', 'P7', 'P8'};
M_PD = nan(3,NPoints);
SE_PD = nan(3,NPoints);
SD_PD = nanstd(PD_Points,0,1);

% One-Way Repeated Measure ANOVA
Pos1 = [0.78 1 1.22; 1.78 2 2.22; 2.78 3 3.22; 3.78 4 4.22; 4.78 5 5.22; 5.78 6 6.22; 6.78 7 7.22; 7.78 8 8.22];
PD_All = {PD_Point1 PD_Point2 PD_Point3 PD_Point4 PD_Point5 PD_Point6 PD_Point7 PD_Point8};
c_PD = nan(3,6,NPoints); m_PD = nan(3,2,NPoints);
figure(7); hold on; % title('Mean Pupil Diameter Change');
for i = 1:NPoints
    [p_PD{i}, table_PD{i}, stats_PD{i}] = anova2(PD_All{i},1,'off');
etap2_PD{i} = table_PD{i}(1,2)/table_PD{i}(1,2)+table_PD{i}(1,4); etag2_PD{i} = table_PD{i}(1,2)/table_PD{i}(1,2)+table_PD{i}(1,4); % [c_PD(:,i),m_PD(:,i)] = multcompare(stats_PD{i},'Display','off'); % post-hoc multiple comparisons
    M_PD(:,i)=m_PD(:,1,i);
    SE_PD(:,i)=m_PD(:,2,i);
end
Bar_PD = bar(squeeze(M_PD(:,1:NPoints))'); hold on;
set(gca, 'XTickLabel', Name_PD(:,1:NPoints), 'XTick', 1:1+numel(Name_PD(:,1:NPoints)), 'FontSize', 12);
xlim([0.4 8.6]);
ylim([-0.15 0.45]);
legend('Level 1', 'Level 2', 'Level 3', 'Location', 'NorthEastoutside')
ylabel('MPDC (mm)');
set(gca, 'XTick', -0.15:0.05:0.45)
Bar_PD(1).FaceColor = Color{1};
Bar_PD(2).FaceColor = Color{2};
Bar_PD(3).FaceColor = Color{3};
CI_PD = 1.96.*SE_PD;
errorbar(Pos1(1:NPoints,:),squeeze(M_PD(:,1:NPoints))',squeeze(CI_PD(:,1:NPoints))','.k
hold on;
for i = 1:NPoints
    if c_PD(1,end,i) < 0.05
        line([i-0.22; i],[0.36; 0.36],'Color',[0 0 0],'LineWidth',1.2)
        line([i-0.22 i; i-0.22 i],[0.35 0.35; 0.36 0.36],'Color',[0 0 0])
    end
if c_PD(3,end,i) < 0.05
    line([i;i+0.22],[0.36; 0.36],'Color',[0 0 0],'LineWidth',1.2)
    line([i i+0.22; i i+0.22],[0.35 0.35; 0.36 0.36],'Color',[0 0 0])
end
if c_PD(2,end,i) < 0.05
    line([i-0.22; i+0.22],[0.39; 0.39],'Color',[0 0 0],'LineWidth',1.2)
    line([i-0.22 i+0.22; i-0.22 i+0.22],[0.38 0.38; 0.39 0.39],'Color',[0 0 0])
end
end
%% Mean Pupil Diameter Change Rate - Bar Graph
PDCR = nan(size(PD_Point1,1),3,NPoints-1);
for k = 1:3
    for j = 1:(NPoints-1)
        for i = 1:size(PD_All{k},1);
            PDCR(i,k,j) = (PD_All{j+1}(i,k)-PD_All{j}(i,k))./(XPoints(j+1,k)-XPoints(j,k));
        end
    end
end
Name_MPDCR = {'(1)' '(2)' '(3)' '(4)' '(5)' '(6)' '(7)'};
M_MPDCR = nan(3,NPoints-1);
SE_MPDCR = nan(3,NPoints-1);
SD_MPDCR = nanstd(PDCR,0,1);
% One-Way Repeated Measure ANOVA
Pos1 = [0.78 1 1.22; 1.78 2 2.22; 2.78 3 3.22; 3.78 4 4.22; 4.78 5 5.22; 5.78 6 6.22; 6.78 7 7.22];
c_MPDCR = nan(3,6,NPoints-1); m_MPDCR = nan(3,2,NPoints-1);
figure(8); hold on; % title('Mean Pupil Diameter Change Rate');
for i = 1:(NPoints-1)
    [p_MPDCR{i}, table_MPDCR{i}, stats_MPDCR{i}] = anova2(PDCR(:,i),1,'off');
    etap2_MPDCR{i} = table_MPDCR{1,i}{2,2}/(table_MPDCR{1,i}{2,2}+table_MPDCR{1,i}{4,2});
    etag2_MPDCR{i} = table_MPDCR{1,i}{2,2}/(table_MPDCR{1,i}{2,2}+table_MPDCR{1,i}{3,2}+table_MPDCR{1,i}{4,2});
    [c_MPDCR(:,:,i),m_MPDCR(:,:,i)] = multcompare(stats_MPDCR{i},'Display','off'); % post-hoc multiple comparisons
    M_MPDCR(:,:,i)=m_MPDCR(:,:,i);
    SE_MPDCR(:,:,i)=m_MPDCR(:,:,i);
end
Bar_MPDCR = bar(squeeze(M_MPDCR')); hold on;
set(gca,'XTickLabel',Name_MPDCR,'XTick',1:numel(Name_MPDCR),'FontSize',12)
xlim([0.4 7.6]); ylim([-8e-4 1.4e-3]);
legend('Level 1','Level 2','Level 3','Location','NorthEastoutside')
ylabel('MPDCR (mm/sample)');
set(gca,'YTick',-8e-4:2e-4:1.4e-3)
Bar_MPDCR(1).FaceColor = Color{1};
Bar_MPDCR(2).FaceColor = Color{2};
Bar_MPDCR(3).FaceColor = Color{3};
CI_MPDCR = 1.96.*SE_MPDCR;
errorbar(Pos1,squeeze(M_MPDCR)',squeeze(CI_MPDCR)','.k'; hold on;
for i = 1:(NPoints-1)
  if c_MPDCR(1,end,i) < 0.05
    line([i-0.22; i], [1e-3; 1e-3], 'Color', [0 0 0], 'LineWidth', 1.2)
    text(i-0.12,1.05e-3, '*', 'FontSize', 15)
  end
  if c_MPDCR(2,end,i) < 0.05
    line([i;i+0.22], [1e-3; 1e-3], 'Color', [0 0 0], 'LineWidth', 1.2)
    line([i i+0.22; i i+0.22], [1.15e-3 1.15e-3; 1.15e-3 1.15e-3], 'Color', [0 0 0], 'LineWidth', 1.2)
    text(i,1.20e-3, '*', 'FontSize', 15)
  end
  if c_MPDCR(3,end,i) < 0.05
    line([i-0.22; i+0.22], [0.95e-3 0.95e-3; 1e-3 1e-3], 'Color', [0 0 0], 'LineWidth', 1.2)
    line([i-0.22 i+0.22; i-0.22 i+0.22], [1e-3 1e-3; 1.15e-3 1.15e-3], 'Color', [0 0 0], 'LineWidth', 1.2)
    text(i+0.08,1.05e-3, '*', 'FontSize', 15)
  end
end
%% NASA-TLX - Bar Graph
NASA_All = nan(30,3,7);
for i = 1:30
  for k = 1:3
    NASA_All(i,Diff(i,k),1) = mean(Nasa(i).N(:,k));
    NASA_All(i,Diff(i,k),2:7) = Nasa(i).N(:,k);
  end
end
Name_NASA = {'Total'; 'Mental'; 'Physical'; 'Temporal'; 'Performance'; 'Effort'; 'Frustration'};
% Names of all
M_NASA = nan(3,7);
SE_NASA = nan(3,7);
SD_NASA = std(NASA_All,0,1);
% One-Way Repeated Measure ANOVA
Pos2 = [0.78 1 1.22; 1.78 2 2.22; 2.78 3 3.22; 3.78 4 4.22; 4.78 5 5.22; 5.78 6 6.22; 6.78 7 7.22];
c_NASA = nan(3,6,7); m_NASA = nan(3,2,7);
figure(9); hold on;
title('NASA-TLX Results');
for i = 1:7
  [p_NASA{i}, table_NASA{i}, stats_NASA{i}] = anova2(NASA_All(:,:,i),1,'off');
  etap2_NASA{i} = table_NASA{1,i}{2,2}/(table_NASA{1,i}{2,2}+table_NASA{1,i}{4,2});
  etag2_NASA{i} = table_NASA{1,i}{2,2}/(table_NASA{1,i}{2,2}+table_NASA{1,i}{3,2}+table_NASA{1,i}{4,2});
  [c_NASA(:,;i),m_NASA(:,;i)] = multcompare(stats_NASA{i}, 'Display', 'off'); % post-hoc multiple comparisons
  M_NASA(:,i)=m_NASA(:,;i);
  SE_NASA(:,i)=m_NASA(:,;2,i);
end
Bar_NASA = bar(squeeze(M_NASA)');
set(gca, 'XTickLabel', Name_NASA, 'XTick', 1:6, 'FontSize', 12);
xlim([0.4 7.6]); ylim([0 100]); ylabel('TLX Score (%); legend('Level 1', 'Level 2', 'Level 3', 'Location', 'northeastoutside');
Bar_NASA(1).FaceColor = Color{1};
Bar_NASA(2).FaceColor = Color{2};
Bar_NASA(3).FaceColor = Color{3};
CI_NASA = 1.96.*SE_NASA;
errorbar(Pos2,squeeze(M_NASA)',squeeze(CI_NASA)','.k'); hold on;

for i = 1:7
    if c_NASA(1,end,i) < 0.05
        line([i-0.22; i],[82; 82], 'Color', [0 0 0], 'LineWidth', 1.2)
        line([i-0.22 i; i-0.22 i],[80 82; 82 82], 'Color', [0 0 0], 'LineWidth', 1.2)
        text(i-0.11-0.02,84,'*','FontSize',15)
    end
    if c_NASA(3,end,i) < 0.05
        line([i;i+0.22],[82; 82], 'Color', [0 0 0], 'LineWidth', 1.2)
        line([i i+0.22; i i+0.22],[80 82; 82 82], 'Color', [0 0 0], 'LineWidth', 1.2)
        text(i+0.11-0.02,84,'*','FontSize',15)
    end
    if c_NASA(2,end,i) < 0.05
        line([i-0.22; i+0.22],[88; 88], 'Color', [0 0 0], 'LineWidth', 1.2)
        line([i-0.22 i+0.22; i-0.22 i+0.22],[86 88; 88 88], 'Color', [0 0 0], 'LineWidth', 1.2)
        text(i-0.02,90,'*','FontSize',15)
    end
end

%% Scatter Plot & Correlation Analys
figure(11); hold on; % title('Scatter plot MPDC Level 1 VS 3 - Point 5');
for i = 1:3
    subplot(1,3,i);
    if i == 1
        scatter(PD_Point1(:,1),PD_Point1(:,3),25, 'MarkerEdgeColor', Color{1}, 'MarkerFaceColor', Color{1});
        title('P1')
    end
    if i == 2
        scatter(PD_Point5(:,1),PD_Point5(:,3),25, 'MarkerEdgeColor', Color{1}, 'MarkerFaceColor', Color{1});
        title('P5')
    end
    if i == 3
        scatter(PD_Point8(:,1),PD_Point8(:,3),25, 'MarkerEdgeColor', Color{1}, 'MarkerFaceColor', Color{1});
        title('P8')
    end
    line([-1 1],[-1 1], 'LineWidth', 2, 'Color', 'k');
    xlim([-0.4 1]); ylim([-0.4 1]);
    xlabel('MPDC Level 1'); ylabel('MPDC Level 3');
axis square
end

% Correlation MPDC - TLX
Corr_All2 = nan(NPoints,3);
PCorr_All2 = nan(NPoints,3);
corr2 = nan(NPoints,1); pvalue2 = nan(NPoints,1);
for i=1:NPoints
    [c1,p1] = corr(PD_All{1,i},NASA_All([1:8 10:30],:,1));
    Corr_All2(i,:) = diag(c1);
    PCorr_All2(i,:) = diag(p1);
    [corr2(i), pvalue2(i)] = corr(mean(PD_All{1,i},2),mean(NASA_All([1:8 10:end],:,1),2));
end
disp('MPDC vs. Overall NASA-TLX')
disp(Corr_All2)
disp(PCorr_All2)
disp(corr2)
disp(pvalue2)
% Correlation MPDC - Responses

TT = nan(30, 15, 3);
for i = 1:30
    for j = 1:15
        for k = 1:3
            if Task(i).T(((k-1)*15+j), 3) == Task(i).T(((k-1)*15+j), 4)
                TT(i, j, Diff(i, k)) = 0;
            else
                TT(i, j, Diff(i, k)) = 1;
            end
        end
    end
end

TTT = squeeze(mean(TT, 2));

Corr_All3 = nan(NPoints, 3); PCorr_All3 = nan(NPoints, 3);
corr3 = nan(NPoints, 1); pvalue3 = nan(NPoints, 1);
for i = 1:NPoints
    [c, p] = (corr(PD_All{1, i}, TTT([1:8 10:30], :), 1));
    Corr_All3(i, :) = diag(c);
    PCorr_All3(i, :) = diag(p);
    [corr3(i), pvalue3(i)] = corr(mean(PD_All{1, i}, 2), mean(TTT([1:8 10:end], :), 2));
end

disp('MPDC vs. % Incorrect responses')
disp((Corr_All3*100)/100)
disp((PCorr_All3*100)/100)
disp(corr3)
disp(pvalue3)

[c10, p10] = corr(NASA_All(:,:,1), TTT(::_,:));
[c11, p11] = corr(mean(NASA_All(:,:,1), 2), mean(TTT(::_,:), 2));
disp('Overall NASA-TLX vs. % Incorrect responses')
disp(diag(c10))
disp(diag(c11))
disp(diag(p10))
disp(diag(p11))
%% ANALYSIS MENTAL MULTIPLICATION EXPERIMENT
% Gerhard Marquart
% #1538446

clear all; close all; clc;

%% LOAD DATA
n = 30; % Number of participants
for i = 1:n
    load(['RESULTS\DATA\DATA' num2str(i) '.mat']); Data(i).D = DATA; clear DATA
    load(['RESULTS\INFO\INFO' num2str(i) '.mat']); Info(i).I = INFO; clear INFO
    load(['RESULTS\NASA\NASA' num2str(i) '.mat']); Nasa(i).N = NASA; clear NASA
    load(['RESULTS\TASK\TASK' num2str(i) '.mat']); Task(i).T = TASK; clear TASK
end
load('Diff.mat');
Color = {[0.0 0.45 0.74] [0.85 0.33 0.10] [0.93 0.69 0.13]};

% X coordinates of # points
XPoints = 120.*[6.5 6.5 6.5; 7.5 7.5 7.5; 9.0 9.0 9.0; 10.5 10.5 10.5; 12.0 12.0 12.0; 13.5 13.5 13.5; 15.0 15.0 15.0; 16.5 16.5 16.5];
NPoints = size(XPoints,1);

%% DATA PROCESSING
L1 = nan(15*n,2200); L2 = nan(15*n,2200); L3 = nan(15*n,2200); L = {L1 L2 L3}; LSmooth = {L1 L2 L3}; LDiff = {L1 L2 L3}; LCorrect = {L1 L2 L3}; LIncorrect = {L1 L2 L3}; LEO = {L1 L2 L3};
LS1 = {L1 L2 L3}; LS2 = {L1 L2 L3}; LS3 = {L1 L2 L3}; LS4 = {L1 L2 L3}; LS5 = {L1 L2 L3}; LS6 = {L1 L2 L3};
l1 = 1; l2 = 1; l3 = 1;
l = {l1 l2 l3}; clear l1 l2 l3
Freq = 120*2;
Time = 17.5;
coeff = ones(1, Freq/4)/(Freq/4);
Delay = (length(coeff)-1)/2;
for i = 1:30
    for j = 1:15
        for k = 1:3
            clear PD PD1 EO PDQ
            EO = Data(i).D(j,k).EO;
            EO(EO==0)=NaN;
            PD = Data(i).D(j,k).PD;
            LS1{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
            PD(PD==0)=NaN;
            LS2{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
            PDQ = Data(i).D(j,k).PDQ;
            PDQ(PDQ==0)=NaN;
            if sum(~isnan(PD)) == 0.95*17.5*120
            end
        end
    end
end
EO(:, isnan(EO)) = interp1(find(~isnan(EO)), EO(:, ~isnan(EO)), find(isnan(EO)), 'linear');
PD(:, isnan(PD)) = interp1(find(~isnan(PD)), PD(:, ~isnan(PD)), find(isnan(PD)), 'linear');
PDQ(:, isnan(PDQ)) = interp1(find(~isnan(PDQ)), PDQ(:, ~isnan(PDQ)), find(isnan(PDQ)), 'linear');
end

LS3{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
eolim = 0.75*nanmean(EO);
PD = blinkremoval(PD, EO, PDQ, 0.75, eolim);
LS4{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
if sum(~isnan(PD)) >= 0.70*17.5*120
PD(:, isnan(PD)) = interp1(find(~isnan(PD)), PD(:, ~isnan(PD)), find(isnan(PD)), 'linear');
LS5{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
end

L{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
LEO{Diff(i,k)}(l{Diff(i,k)},1:length(EO)) = EO;
TEMP_Smooth = filter(coeff,1,PD);
LSsmooth{Diff(i,k)}(l{Diff(i,k)},1:(length(PD)-(Freq/8)+1)) =
TEMP_Smooth(Freq/8:end);
LDiff{Diff(i,k)}(l{Diff(i,k)},1:(length(PD)-(Freq/8))) = abs(diff(TEMP_Smooth(Freq/8:end))./(1/120));
if Task(i).T(((k-1)*15+j),3) == Task(i).T(((k-1)*15+j),4)
LCorrect{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
else
LIncorrect{Diff(i,k)}(l{Diff(i,k)},1:length(PD)) = PD;
end
l{Diff(i,k)} = l{Diff(i,k)}+1;
end
end

%% Mean Pupil Diameter Change of all trials per level of difficulty
for i = 0:n-1
for k = 1:3
Ln{k}(i+1,:) = nanmean(L{k}(i*15+1:(i+1)*15,:),1);
LCorrectn{k}(i+1,:) = nanmean(LCorrect{k}(i*15+1:(i+1)*15,:),1);
LIncorrectn{k}(i+1,:) = nanmean(LIncorrect{k}(i*15+1:(i+1)*15,:),1);
end
end
figure(1);
% title('Mean Pupil Diameter - Level 3 - Correct VS Incorrect','FontSize',12);
rectangle('Position',[4,-1,1,10], 'EdgeColor',[0.6 0.6 0.6], 'FaceColor',[0.9 0.9 0.9]);
hold on;
rectangle('Position',[6.5,-1,1,10], 'EdgeColor',[0.6 0.6 0.6], 'FaceColor',[0.9 0.9 0.9]);
hold on;
h1 = text(4.4, 4.10, 'Multiplicand', 'FontSize',12); set(h1, 'rotation', 90)
h2 = text(6.9, 4.10, 'Multiplier', 'FontSize',12); set(h2, 'rotation', 90)
for k = 3
plot(linspace(0,length(Ln{k})/120,length(Ln{k})),nanmean(Ln{k}).*1000,'Color',Color{k}, 'LineWidth',1); hold on;
plot(linspace(0,length(LCorrectn{k})/120,length(LCorrectn{k})),nanmean(LCorrectn{k}).*1000,'g','LineWidth',1); hold on;
plot(linspace(0,length(LIncorrectn{k})/120,length(LIncorrectn{k})),nanmean(LIncorrectn{k}).*1000,'r','LineWidth',1); hold on;
xlim([0 17.5])
ylim([3.6 4.3])
xlabel('Time (s)');
ylabel('MPD (mm)');
legend('Level 3 All',...
    ['Level 3 Correct (' sprintf('%.1f',size(LCorrect{1,k})(any(~isnan(LCorrect{1,k}),2),:),1).*100) ' %')',...
    'Level 3 Incorrect (' sprintf('%.1f',size(LIncorrect{1,k})(any(~isnan(LIncorrect{1,k}),2),:),1).*100) ' %'),'Location','SouthEast');
set(gca,'XTick',0:1:18)

Sn = size(Ln{3},1); Sc = size(LCorrectn{3},1); Si = size(LIncorrectn{3},1);
MPD_Points_c = nan(max([Sn Sc Si]),3,NPoints);
for k = 3
    for j = 1:NPoints
        for i = 1:size(Ln{k},1);
            MPD_Points_c(i,1,j) = (Ln{k}(i,XPoints(j,k))).*1000;
        end
        for i = 1:size(LCorrectn{k},1);
            MPD_Points_c(i,2,j) = (LCorrectn{k}(i,XPoints(j,k))).*1000;
        end
        for i = 1:size(LIncorrectn{k},1);
            MPD_Points_c(i,3,j) = (LIncorrectn{k}(i,XPoints(j,k))).*1000;
        end
    end
end

MPD_Point1_c = MPD_Points_c(:,:,1); MPD_Point1_c = MPD_Point1_c(~any(isnan(MPD_Point1_c),2),:);
MPD_Point2_c = MPD_Points_c(:,:,2); MPD_Point2_c = MPD_Point2_c(~any(isnan(MPD_Point2_c),2),:);
MPD_Point3_c = MPD_Points_c(:,:,3); MPD_Point3_c = MPD_Point3_c(~any(isnan(MPD_Point3_c),2),:);
MPD_Point4_c = MPD_Points_c(:,:,4); MPD_Point4_c = MPD_Point4_c(~any(isnan(MPD_Point4_c),2),:);
MPD_Point5_c = MPD_Points_c(:,:,5); MPD_Point5_c = MPD_Point5_c(~any(isnan(MPD_Point5_c),2),:);
MPD_Point6_c = MPD_Points_c(:,:,6); MPD_Point6_c = MPD_Point6_c(~any(isnan(MPD_Point6_c),2),:);
MPD_Point7_c = MPD_Points_c(:,:,7); MPD_Point7_c = MPD_Point7_c(~any(isnan(MPD_Point7_c),2),:);
MPD_Point8_c = MPD_Points_c(:,:,8); MPD_Point8_c = MPD_Point8_c(~any(isnan(MPD_Point8_c),2),:);

Name_MPD_c = {'P1' 'P2' 'P3' 'P4' 'P5' 'P6' 'P7' 'P8'};
M_MPD_c = nan(3,NPoints);
SE_MPD_c = nan(3,NPoints);
SD_MPD_c = nanstd(MPD_Points_c,0,1);

% One-Way Repeated Measure ANOVA
Pos1_c = [0.78 1 1.22; 1.78 2 2.22; 2.78 3 3.22; 3.78 4 4.22; 4.78 5 5.22; 5.78 6 6.22; 6.78 7 7.22; 7.78 8 8.22];
MPD_All_c = {MPD_Point1_c MPD_Point2_c MPD_Point3_c MPD_Point4_c MPD_Point5_c MPD_Point6_c MPD_Point7_c MPD_Point8_c};
c_MPD_c = nan(3,6,NPoints); m_MPD_c = nan(3,2,NPoints);
figure(10); hold on; %title('Mean Pupil Diameter');
for i = 1:NPoints
    [p_MPD_c{i}, table_MPD_c{i}, stats_MPD_c{i}] = anova2(MPD_All_c{i},1,'off');
etap2_MPD_c{i} = table_MPD_c{1,i}{2,2}/(table_MPD_c{1,i}{2,2}+table_MPD_c{1,i}{4,2});
etag2_MPD_c{i} = table_MPD_c{1,i}{2,2}/(table_MPD_c{1,i}{2,2}+table_MPD_c{1,i}{3,2}+table_MPD_c{1,i}{4,2});
c_MPD_c(:,:,i), m_MPD_c(:,:,i) = multcompare(stats_MPD_c{i}, 'Display', 'off');

post-hoc multiple comparisons
M_MPD_c(:,:,i)=m_MPD_c(:,:,i);
SE_MPD_c(:,:,i)=m_MPD_c(:,:,i);

Bar_MPD_c = bar(squeeze(M_MPD_c)'); hold on;
set(gca, 'XTickLabel', Name_MPD_c, 'XTick', 1:numel(Name_MPD_c), 'FontSize', 12);
xlim([0.4 8.6]);
ylim([3.80 4.40]);
legend('Level 3 All', 'Level 3 Correct', 'Level 3 Incorrect', 'Location', 'NorthEastoutside')
ylabel('MPD (mm)');
Bar_MPD_c(1).FaceColor = Color{3};
Bar_MPD_c(2).FaceColor = 'g';
Bar_MPD_c(3).FaceColor = 'r';
CI_MPD_c = 1.96.*SE_MPD_c;
errorbar(Pos1_c, squeeze(M_MPD_c)', squeeze(CI_MPD_c)', '.k'); hold on;
for i = 1:NPoints
    if c_MPD_c(1,end,i) < 0.05
        line([i-0.22; i], [4.32; 4.32], 'Color', [0 0 0], 'LineWidth', 1.2)
        line([i-0.22 i; i-0.22 i], [4.30 4.30; 4.32 4.32], 'Color', [0 0 0], 'LineWidth', 1.2)
        text(i-0.11, 4.33, '*', 'FontSize', 15)
    end
    if c_MPD_c(3,end,i) < 0.05
        line([i+0.22; i], [4.32; 4.32], 'Color', [0 0 0], 'LineWidth', 1.2)
        line([i i+0.22; i i+0.22], [4.30 4.30; 4.32 4.32], 'Color', [0 0 0], 'LineWidth', 1.2)
        text(i+0.09, 4.33, '*', 'FontSize', 15)
    end
    if c_MPD_c(2,end,i) < 0.05
        line([i-0.22; i+0.22], [4.38; 4.38], 'Color', [0 0 0], 'LineWidth', 1.2)
        line([i-0.22 i+0.22; i-0.22 i+0.22], [4.36 4.36; 4.38 4.38], 'Color', [0 0 0], 'LineWidth', 1.2)
        text(i-0.01, 4.41, '*', 'FontSize', 15)
    end
end

figure(2);
% title('Mean Pupil Diameter per Level', 'FontSize', 12);
rectangle('Position', [4, -1, 1, 10], 'EdgeColor', [0.6 0.6 0.6], 'FaceColor', [0.9 0.9 0.9]);
hold on;
rectangle('Position', [6.5, -1, 1, 10], 'EdgeColor', [0.6 0.6 0.6], 'FaceColor', [0.9 0.9 0.9]);
hold on;
h1 = text(4.4, 4.10, 'Multiplicand', 'FontSize', 12); set(h1, 'rotation', 90)
h2 = text(6.9, 4.10, 'Multiplier', 'FontSize', 12); set(h2, 'rotation', 90)
for k = 1:3
    plot(linspace(0, length(Ln{k})/120, length(Ln{k})), nanmean(Ln{k}).*1000, 'Color', Color{k},
        'LineWidth', 1); hold on;
    xlim([0 17.5])
    ylim([3.6 4.3])
end
xlabel('Time (s)');
ylabel('MPD (mm)');
legend('Level 1', 'Level 2', 'Level 3', 'Location', 'NorthEast')
set(gca, 'XTick', 0:1:18)

figure(3);
% title('Mean Pupil Diameter per Level', 'FontSize', 12);
rectangle('Position', [4, -1, 1, 10], 'EdgeColor', [0.6 0.6 0.6], 'FaceColor', [0.9 0.9 0.9]);
hold on;
rectangle('Position', [6.5, -1, 1, 10], 'EdgeColor', [0.6 0.6 0.6], 'FaceColor', [0.9 0.9 0.9]);
hold on;
rectangle('Position',[7.5,-1,1.5,10],'EdgeColor',[0.8 0.8 0.8],'FaceColor',[0.95 0.95 0.95]); hold on;
rectangle('Position',[9.0,-1,1.5,10],'EdgeColor',[0.8 0.8 0.8],'FaceColor',[0.95 0.95 0.95]); hold on;
rectangle('Position',[10.5,-1,1.5,10],'EdgeColor',[0.8 0.8 0.8],'FaceColor',[0.95 0.95 0.95]); hold on;
rectangle('Position',[12.0,-1,1.5,10],'EdgeColor',[0.8 0.8 0.8],'FaceColor',[0.95 0.95 0.95]); hold on;
rectangle('Position',[13.5,-1,1.5,10],'EdgeColor',[0.8 0.8 0.8],'FaceColor',[0.95 0.95 0.95]); hold on;
rectangle('Position',[15.0,-1,1.5,10],'EdgeColor',[0.8 0.8 0.8],'FaceColor',[0.95 0.95 0.95]); hold on;

h1 = text(4.4, 4.10,'Multiplicand','FontSize',12); set(h1,'rotation',90);
h2 = text(6.9, 4.10,'Multiplier','FontSize',12); set(h2,'rotation',90);
P1 = text(6.90, 3.64,'(1)','FontSize',12);
P2 = text(8.15, 3.64,'(2)','FontSize',12);
P3 = text(9.65, 3.64,'(3)','FontSize',12);
P4 = text(11.15, 3.64,'(4)','FontSize',12);
P5 = text(12.65, 3.64,'(5)','FontSize',12);
P6 = text(14.15, 3.64,'(6)','FontSize',12);
P7 = text(15.65, 3.64,'(7)','FontSize',12);

for k = 1:3
plot(linspace(0,length(Ln{k})/120,length(Ln{k})),nanmean(Ln{k}).*1000,'Color',Color{k},'LineWidth',1); hold on;
xlim([6.5 17.5])
ylim([3.6 4.3])
Ln2{k} = nanmean(Ln{k}).*1000;
end
xlabel('Time (s)');
ylabel('MPD (mm)');
set(gca,'XTick',6.5:0.5:17.5)

plot(0,0,'xk','LineWidth',1.5);
plot(0,0,'ok','LineWidth',1.5);
plot(0,0,'^k','LineWidth',1.5);
plot(0,0,'*k','LineWidth',1.5);
plot(0,0,'>k','LineWidth',1.5);
plot(0,0,'<k','LineWidth',1.5);
plot(0,0,'sk','LineWidth',1.5);

for k = 1:3
plot(XPoints(1,k)./120,Ln2{k}(XPoints(1,k)+1),'x','Color',Color{k},'LineWidth',2);
plot(XPoints(2,k)./120,Ln2{k}(XPoints(2,k)),'o','Color',Color{k},'LineWidth',2);
plot(XPoints(3,k)./120,Ln2{k}(XPoints(3,k)),'^','Color',Color{k},'LineWidth',2);
plot(XPoints(4,k)./120,Ln2{k}(XPoints(4,k)),'*','Color',Color{k},'LineWidth',2);
plot(XPoints(5,k)./120,Ln2{k}(XPoints(5,k)),'>','Color',Color{k},'LineWidth',2);
plot(XPoints(6,k)./120,Ln2{k}(XPoints(6,k)),'<','Color',Color{k},'LineWidth',2);
plot(XPoints(7,k)./120,Ln2{k}(XPoints(7,k)),'s','Color',Color{k},'LineWidth',2);
end
legend('Level 1','Level 2','Level 3','P1','P2','P3','P4','P5','P6','P7','P8','Location','NorthEastoutside')

S1 = size(Ln{1},1); S2 = size(Ln{2},1); S3 = size(Ln{3},1);
MPD_Points = nan(max([S1 S2 S3]),3,6);
for k = 1:3
for j = 1:NPoints
for i = 1:size(Ln{k},1);
MPD_Points(i,k,j) = (Ln{k}(i,XPoints(j,k))).*1000;
end
end
end
MPD_Point1 = MPD_Points(:,:,1); MPD_Point1 = MPD_Point1(~any(isnan(MPD_Point1),2),:);
MPD_Point2 = MPD_Points(:,:,2); MPD_Point2 = MPD_Point2(~any(isnan(MPD_Point2),2),:);
MPD_Point3 = MPD_Points(:,:,3); MPD_Point3 = MPD_Point3(~any(isnan(MPD_Point3),2),:);
MPD_Point4 = MPD_Points(:,:,4); MPD_Point4 = MPD_Point4(~any(isnan(MPD_Point4),2),:);
MPD_Point5 = MPD_Points(:,:,5); MPD_Point5 = MPD_Point5(~any(isnan(MPD_Point5),2),:);
MPD_Point6 = MPD_Points(:,:,6); MPD_Point6 = MPD_Point6(~any(isnan(MPD_Point6),2),:);
MPD_Point7 = MPD_Points(:,:,7); MPD_Point7 = MPD_Point7(~any(isnan(MPD_Point7),2),:);
MPD_Point8 = MPD_Points(:,:,8); MPD_Point8 = MPD_Point8(~any(isnan(MPD_Point8),2),:);

Name_MPD = {'P1' 'P2' 'P3' 'P4' 'P5' 'P6' 'P7' 'P8'};
M_MPD = nan(3,NPoints);
SE_MPD = nan(3,NPoints);
SD_MPD = nanstd(MPD_Points,0,1);

% One-Way Repeated Measure ANOVA
Pos1 = [0.78 1 1.22; 1.78 2 2.22; 2.78 3 3.22; 3.78 4 4.22; 4.78 5 5.22; 5.78 6 6.22; 6.78 7 7.22; 7.78 8 8.22];
MPD_All = {MPD_Point1 MPD_Point2 MPD_Point3 MPD_Point4 MPD_Point5 MPD_Point6 MPD_Point7 MPD_Point8};
c_MPD = nan(3,6,NPoints); m_MPD = nan(3,2,NPoints);
figure(4); hold on;
%title('Mean Pupil Diameter');
for i = 1:NPoints
    [p_MPD{i}, table_MPD{i}, stats_MPD{i}] = anova2(MPD_All{i},1,'off');
etap2_MPD{i} = table_MPD{1,i}{2,2}/(table_MPD{1,i}{2,2}+table_MPD{1,i}{4,2});
etag2_MPD{i} = table_MPD{1,i}{2,2}/(table_MPD{1,i}{2,2}+table_MPD{1,i}{3,2}+table_MPD{1,i}{4,2});
[c_MPD(:,i),m_MPD(:,:,i)] = multcompare(stats_MPD{i},'Display','off'); % post-hoc multiple comparisons
M_MPD(:,i)=m_MPD(:,1,i);
SE_MPD(:,i)=m_MPD(:,2,i);
end
Bar_MPD = bar(squeeze(M_MPD)'); hold on;
set(gca,'XTickLabel',Name_MPD,'XTick',1:numel(Name_MPD),'FontSize',12)
xlim([0.4 8.6]); ylim([3.6 4.40]);
legend('Level 1','Level 2','Level 3','Location','NorthEastoutside')
ylabel('MPD (mm)');
Bar_MPD(1).FaceColor = Color{1};
Bar_MPD(2).FaceColor = Color{2};
Bar_MPD(3).FaceColor = Color{3};
CI_MPD = 1.96.*SE_MPD;
errorbar(Pos1,squeeze(M_MPD)',squeeze(CI_MPD)','.k'); hold on;
for i = 1:NPoints
    if c_MPD(1,end,i) < 0.05
        line([i-0.22; i],[4.26; 4.26],'Color',[0 0 0],'LineWidth',1.2)
        line([i-0.22 i; i-0.22 i],[4.24 4.24; 4.26 4.26],'Color',[0 0 0],'LineWidth',1.2)
        text(i-0.11,4.28,'*','FontSize',15)
    end
    if c_MPD(3,end,i) < 0.05
        line([i;i+0.22],[4.26; 4.26],'Color',[0 0 0],'LineWidth',1.2)
        line([i i+0.22; i+0.22 i+0.22],[4.24 4.24; 4.26 4.26],'Color',[0 0 0],'LineWidth',1.2)
        text(i+0.11,4.28,'*','FontSize',15)
    end
    if c_MPD(2,end,i) < 0.05
        line([i-0.22; i+0.22],[4.32; 4.32],'Color',[0 0 0],'LineWidth',1.2)
        line([i-0.22 i+0.22; i-0.22 i+0.22],[4.30 4.30; 4.32 4.32],'Color',[0 0 0],'LineWidth',1.2)
        text(i,4.34,'*','FontSize',15)
    end
end

%% Scatter Plot & Correlations
figure(11); hold on; % title('Scatter plot MPDC Level 1 VS 3 - Point 5');
for i = 1:3
    subplot(1,3,i);
    if i == 1
        scatter(MPD_Point1(:,1),MPD_Point1(:,3),25,'MarkerEdgeColor',Color{1},'MarkerFaceColor',Color{1});
        title('P1')
    end
    if i == 2
        scatter(MPD_Point5(:,1),MPD_Point5(:,3),25,'MarkerEdgeColor',Color{1},'MarkerFaceColor',Color{1});
        title('P5')
    end
    if i == 3
        scatter(MPD_Point8(:,1),MPD_Point8(:,3),25,'MarkerEdgeColor',Color{1},'MarkerFaceColor',Color{1});
        title('P8')
    end
    line([-1 5.5],[-1 5.5], 'LineWidth',2, 'Color','k');
    xlim([2.5 5.5]); ylim([2.5 5.5]);
    xlabel('MPD Level 1'); ylabel('MPD Level 3');
    axis square
end
clear all; close all; clc;

%% LOAD DATA
n = 30; % Number of participants
for i = 1:n
    load(['RESULTS\DATA\DATA' num2str(i) '.mat']); Data(i).D = DATA; clear DATA
    load(['RESULTS\INFO\INFO' num2str(i) '.mat']); Info(i).I = INFO; clear INFO
    load(['RESULTS\NASA\NASA' num2str(i) '.mat']); Nasa(i).N = NASA; clear NASA
    load(['RESULTS\TASK\TASK' num2str(i) '.mat']); Task(i).T = TASK; clear TASK
end
load('Diff.mat');
Color = {{[0.0 0.45 0.74] [0.85 0.33 0.10] [0.93 0.69 0.13]};

%% DATA PROCESSING
L1 = nan(15*n,2200); L2 = nan(15*n,2200); L3 = nan(15*n,2200);
LEO = {L1 L2 L3}; clear L1 L2 L3
L1 = nan(15*n,2); L2 = nan(15*n,2); L3 = nan(15*n,2);
Count = {L1 L2 L3}; clear L1 L2 L3
l1 = 1; l2 = 1; l3 = 1;
l = {l1 l2 l3}; clear l1 l2 l3
Freq = 120;
Time = 17.5;
coeff = ones(1, Freq/4)/(Freq/4);
Delay = (length(coeff)-1)/2;
Times = [2 6.5 13].*120;
for i = 1:30
    for j = 1:15
        for k = 1:3
            clear EO CC
            EO = Data(i).D(j,k).EO;
            EO(EO==0)=NaN;
            if sum(~isnan(EO)) >= 1
                EO(:,isnan(EO)) = interp1(find(~isnan(EO)), EO(:,~isnan(EO)),
                find(isnan(EO)), 'linear');
                Count{Diff(i,k)}{l{Diff(i,k)},1} = 0;
                Count{Diff(i,k)}{l{Diff(i,k)},2} = 0;
            end
            LEO{Diff(i,k)}{l{Diff(i,k)},1:length(EO)} = EO;
eolim = 0.75*nanmean(EO);
EO = EO-eolim; % 0.010
    end
    for m=Times(1):Times(2)
        if EO(m)<0 && EO(m-1)>0
            Count{Diff(i,k)}{l{Diff(i,k)},1} = Count{Diff(i,k)}{l{Diff(i,k)},1}+1;
        end
    end
    for m=Times(2):Times(3)
        if EO(m)<0 && EO(m-1)>0
            Count{Diff(i,k)}{l{Diff(i,k)},2} = Count{Diff(i,k)}{l{Diff(i,k)},2}+1;
        end
    end
end
l{Diff(i,k)} = l{Diff(i,k)}+1;
end
end

L1 = nan(n,2); L2 = nan(n,2); L3 = nan(n,2);
Countn = {L1 L2 L3}; clear L1 L2 L3;
for i = 0:n-1
  for k = 1:3
    Countn{k}(i+1,1) = nanmean(Count{k}(i*15+1:(i+1)*15,1),1);
    Countn{k}(i+1,2) = nanmean(Count{k}(i*15+1:(i+1)*15,2),1);
  end
end
ML1 = mean(Count{1});
ML2 = mean(Count{2});
ML3 = mean(Count{3});

% One-Way Repeated Measure ANOVA
Pos1 = [0.78 1 1.22; 1.78 2 2.22];
MBR_All = {{[Countn{1}(:,1) Countn{2}(:,1) Countn{3}(:,1)]./((Times(2)-Times(1))/120)
  [Countn{1}(:,2) Countn{2}(:,2) Countn{3}(:,2)]./((Times(3)-Times(2))/120)
}
  c_MBR = nan(3,6,2); m_MBR = nan(3,2,2);
Name_MBR = {'Low Demands (2-6.5 s)' 'High Demands (6.5-13 s)'};
SD_MBR1 = nanstd(MBR_All{1},0,1);
SD_MBR2 = nanstd(MBR_All{2},0,1);
SD_MBR = [SD_MBR1; SD_MBR2];
figure(4); hold on; %title('Mean Blink Rate');
for i = 1:2
  [p_MBR{i}, table_MBR{i}, stats_MBR{i}] = anova2(MBR_All{i},1,'off');
etap2_MBR{i} = table_MBR{1,i}{2,2}/(table_MBR{1,i}{2,2}+table_MBR{1,i}{4,2});
etag2_MBR{i} = table_MBR{1,i}{2,2}/(table_MBR{1,i}{2,2}+table_MBR{1,i}{3,2}+table_MBR{1,i}{4,2});
  [c_MBR(:,i),m_MBR(:,i)] = multcompare(stats_MBR{i},'Display','off'); % post-hoc multiple comparisons
  M_MBR(:,i)=m_MBR(:,1,i);
  SE_MBR(:,i)=m_MBR(:,2,i);
end
Bar_MBR = bar(squeeze(M_MBR)); hold on;
set(gca,'XTickLabel',Name_MBR,'XTick',1:numel(Name_MBR),'FontSize',12)
ylim([0 0.7]);
legend('Level 1','Level 2','Level 3','Location','NorthEastoutside')
ylabel('MBR (blinks/s)');
Bar_MBR(1).FaceColor = Color(1);
Bar_MBR(2).FaceColor = Color(2);
Bar_MBR(3).FaceColor = Color(3);
CI_MBR = 1.96.*SE_MBR;
errorbar(Pos1,squeeze(M_MBR)',squeeze(CI_MBR)','.k'; hold on;
for i = 1:2
  if c_MBR(1,end,i) < 0.05
    line([i-0.22 i],[0.54 0.54],'Color',[0 0 0],'LineWidth',1.2)
    text(i-0.11,0.56,'**','FontSize',15)
  end
  if c_MBR(3,end,i) < 0.05
    line([i i+0.22],[0.54 0.54],'Color',[0 0 0],'LineWidth',1.2)
    text(i+0.11,0.56,'**','FontSize',15)
  end
  if c_MBR(2,end,i) < 0.05
    line([i-0.22 i+0.22],[0.60 0.60],'Color',[0 0 0],'LineWidth',1.2)
    text(i+0.11,0.56,'**','FontSize',15)
  end
end
function PD = blinkremoval(PD,EO,PDQ,PDQlim,eolim)
% Blink identification and removal
PD(EO>0.020)=NaN; % unrealistically large eyelid opening

Freq = 120; % Recorded frequency
del1 = Freq*0.1; % del before blink
del2 = Freq*0.3; % del after blink

for i1=1:length(PD)
    if EO(i1)< eolim % 0.012
        if i1<=del1
            PD(i1+(0:del2))=NaN;
        elseif i1>(length(PD)-del2)
            PD(i1-(1:del1))=NaN;
            PD(i1:end)=NaN;
        else
            PD(i1-(1:del1))=NaN;
            PD(i1+(0:del2))=NaN;
        end
    end
end

for i1=1:length(PD)
    if PDQ(i1)<PDQlim
        PD(i1)=NaN;
    end
end