Pupil Response to Precision in Surgical Task Execution

Xianta Jiang\textsuperscript{a,b}, Bin Zheng\textsuperscript{c}, Geoffrey Tien\textsuperscript{b}, and M. Stella Atkins\textsuperscript{b,1}

\textsuperscript{a}Affiliation College of Computer Science and Technology, Zhejiang University
\textsuperscript{b}School of Computing Science, Simon Fraser University
\textsuperscript{c}Department of Surgery, University of Alberta

Abstract. Task-evoked pupil response (TEPR) has been extensively studied and well proven to be sensitive to mental workload changes. We aimed to explore how TEPR reflects mental workload changes in a surgical environment. We conducted a simulated surgical task that has 3 different subtasks with different levels of motor precision and different mental workload requirements. We found a significant effect among these different subtask groups by measuring pupil diameter change rate. This finding may improve patient safety in a real operating room by non-intrusively monitoring the surgeon’s mental workload while performing a surgery using an eye-tracking system.

Keywords. Pupil response, eye-tracking, surgical simulation, mental workload

Introduction

The pupil has been well proven to react not only to the ambient light and psychological changes, but also to the intensity of mental effort demanded by the task performance [1]. Hess and Polt [2] found that the pupil gradually dilated when preparing the answer to a multiplication problem and reached a peak immediately before the answer was orally reported; then, it rapidly constricted back to the original size. They also found that the mean pupil dilation was a function of the level of difficulty of the problem.

The pupil size changes not only in response to the task difficulty overall, but also with respect to critical events during an information processing task, called the task-evoked pupillary response (TEPR) [3]. TEPR has been extensively studied and proven to be an efficient index of mental workload while participants are performing tasks. This has been tested in tasks of driving vehicles [4], interacting with computer interfaces [5], and performing surgery [6].

Objectively measuring mental workload in performing a laparoscopic operation has been proposed by several researchers [7-9]. Berguer et al. [9] found that performing laparoscopic surgery causes higher mental stress than open surgery by measuring physiological signals, i.e., skin conductance level (SCL) and blinks, collected from the participants when performing simulated tasks in both laparoscopic and open operating situations. Results have been confirmed using psychological assessment [7]. Zheng et al. [10] found that blink rate and mental workload of the participants were correlated during a simulated surgical task. Richstone et al. [6] used eye movement behaviors

\textsuperscript{1}E-mail: stella@sfu.ca
including pupil movements and blinks to predict surgical skill which distinguished expert and novice surgeons at 81.0% and 90.7% accuracy respectively in the live operating room setting. To examine the relationship between TEPR and surgical workload, we conducted an experiment in a surgical simulation lab where the participants were required to perform a simulated laparoscopic task with their pupillary movements recorded. The task includes 3 groups of subtasks, demanding different levels of intensity of mental workload. We hypothesized that the change rate of the pupil diameter would reflect the precision requirement of the subtask, i.e. the pupil diameter would increase faster during performing higher precision subtasks than lower precision subtasks. Similarly, the pupil would constrict rapidly while performing subtasks requiring low precision.

We hypothesize that TEPR indicates the level of precision of subtask in a surgical task execution. Thus pupil measurements can be used for monitoring the mental workload of a surgeon when performing a surgery, leading to improved patient safety.

1. Methods & Materials

1.1. Experiment setting & Apparatus

This study was conducted in a surgical simulation room in Vancouver General Hospital. 12 subjects were recruited including surgeons and office staff. Each participant signed a consent form.

![Figure 1a. The experimental setting.](image)

![Figure 1b. Illustration of the task.](image)

As shown in Figure 1a, the manual task was performed inside a laparoscopic training box manufactured by 3-D Technical Services equipped with a single grasper in one of the four entrance ports, detailed in [11]. The remote eye-tracker (Tobii 1750) has a built-in 17” LCD display, and records eye gaze points on the display at 50 Hz. The web camera recorded facial expressions of the operator for validation purpose.

1.2. Task & Procedure

The task was to transport the rubber peg between three dishes using the grasper. Each participant was given a brief oral description of the task and practiced a few minutes before starting to perform the task. Each participant performed five trials, with a pause...
between each trial. At the beginning and end of each trial, a camera flash was given for synchronization purpose. The ambient light was constant and controlled for all trials.

A trial has 9 subtasks as shown in Figure 1b, which can be grouped into three basic movements with different precision requirements: reaching and grasping the object (RG), transporting and releasing the object (TR), and bringing the instrument to the home position in the white central square (H). Taking the first three subtasks as an example, starting from the home position, the grasper was moved to the top dish (6 mm) and picked up the peg (2 mm) (RG). The peg was then transported from the top to the left bottom dish using the grasper, which was opened to release the peg into the left bottom dish (TR). After releasing the peg, the empty grasper was moved to the home position (H). Each basic movement was repeated three times until the peg was brought back to the original place and the grasper back to the home position.

When the participants were performing reaching and grasping (RG), greater mental effort was demanded since speed and motor need to be well controlled; the operator had to control the grasper by slowing down and opening the grasper when it approached the target after a relatively fast move from the home position. Then the operator had to decide and locate a proper position for the tool tip of the grasper to stop for the peg, and the grasping action had to be performed very carefully to avoid dropping the peg. Transporting the object to a cup (TR) might be also demanding as it required the subject to carefully place the peg into the 6 mm dish. In contrast, bringing the empty grasper back to the home position (H) was less demanding.

1.3. Data Analysis

Surgical videos were captured from the display on the Tobii monitor while participants were performing the task, and were manually annotated by recording the start time of each subtask in milliseconds. The criteria for judging the start of a subtask are as follows. For the start of the RG movement: the first frame that the tool starts to move towards the target dish after being in the home position. For the start of the TR movement: the first frame that the tool moves after successfully grasping the peg and lifting it up above the dish. For the start of the H movement: the first frame that the empty grasper moves towards the home position after successfully releasing the peg into the dish.

We need to adjust the pupil diameter (ranging from 3.18 mm to 5.60 mm among the 12 subjects) by an appropriate baseline diameter. The baseline for each trial is calculated based on the average pupil diameter of the samples in the 400 ms period around the start of RG1 (200 ms before and 200 ms after the start of RG1).

2. Results

From the possible 540 subtasks produced by the 12 participants each performing 5 trials, 6 subtasks were excluded because of peg dropping. For each of the 534 valid subtasks, the average adjusted pupil diameter and the average rate of pupil diameter change (the slope of the adjusted pupil diameter over time) were calculated.

We performed a one-way ANOVA analysis on the data to find pupillary response differences among the three groups of subtasks, based on adjusted pupil diameter and the slope of rate of change of pupil diameter. The output of the analysis is shown in Table 1.
Table 1. The output of the single ANOVA analysis on the average adjusted pupil diameter and average rate of change of diameter (slope) over three types of subtasks.

<table>
<thead>
<tr>
<th></th>
<th>RG</th>
<th>TR</th>
<th>H</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted pupil size (mm)</td>
<td>0.427±0.202</td>
<td>0.603±0.288</td>
<td>0.473±0.228</td>
<td>0.227</td>
</tr>
<tr>
<td>Slope (mm/sec)</td>
<td>0.163±0.108</td>
<td>0.035±0.033</td>
<td>-0.299±0.179</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Adjusted pupil size did not show significant differences among the three types of movement (Row 1 in Table 1, F(2,33) = 1.553, p = 0.227). The largest pupil size was recorded in performing the TR movement, rather than in RG movement. Using the slope of rate of change of pupil diameter for one-way ANOVA analysis, we found a significant effect among RG, TR, and H groups on value (F(2,33) = 41.837, p < 0.001), as shown in row 2 in Table 1. Post-hoc analysis using Tukey HSD test shows that the average slope of H subtask was less than that of TR subtask (p < 0.001) and RG subtask (p < 0.001), and the average slope of the TR subtask is less than that of RG subtask (p = 0.05). Performing RG subtasks demands the highest mental workload causing rapid pupil dilation (mean slope = 0.163 mm/sec); the H subtask is easy which causes rapid pupil relaxing (mean slope = -0.299 mm/sec); the TR subtask is in-between, with slightly enlarged pupil size (mean slope = 0.035 mm/sec). The changes in pupil diameter over time for a single trial are shown in Fig. 2.

3. Discussion

Our hypothesis was partially supported, although the largest pupil size was not recorded during the most challenging movement (RG), but in the TR movement, likely caused by the order of the types of subtasks. The pupil diameter in RG starts to increase from a rest state with a small pupil size to an active movement with a larger pupil size. On average pupil diameter increased from the RG baseline (0.427 mm) and continued to enlarge slightly in performing the TR task (0.603 mm). The pupil diameter in H subtasks drops down rapidly towards the baseline, resulting in a size close to that in RG. Pupil size failed to show significant difference among the three types of movement, suggesting that average pupil diameter does not correlate directly with the task difficulty.
However, pupil size changes over time (calculated by the slope) supports our hypothesis well. Pupil size increased significantly during RG movements and increased slightly during TR movements; while in H movements, the pupil started to shrink rapidly. The pupil enlarged less during the TR movement than in RG because the mental workload demanded by the TR is lower than the RG subtask since releasing an object is relatively easier than grasping an object (the cup size is larger than the size of peg). Also the TR subtask follows the RG subtask where the pupil diameter nearly reached its maximum and has to decrease. The pupil diameter in most of the TR subtasks first underwent a decrease and then increased again in a V-shape so the average pupil diameter and the slope during the TR subtask did not reflect the real difficulty of the task. Performing the H subtask was easier since there was no peg to carry or to pick up, and the target (the white central square) is relatively big.

4. Conclusions

Although TEPR has been extensively studied and well proven to be sensitive to mental workload changes, there are few works examining how TEPR correlates with mental workload intensity during surgical task execution. This study shows that the rate of change of pupil diameter matches well to the change of precision requirement of a surgical task and can serve as a better behavioral indicator for assessing mental workload of a surgeon than the pupil diameter. This finding may allow non-intrusive monitoring of a surgeon’s mental workload during a real surgery using a remote eye-tracking system. Future work will explore individual differences among surgeons.

References