Contents lists available at ScienceDirect

Int. J. Human-Computer Studies

Pupil responses to continuous aiming movements ☆

Xianta Jiang [a,*], Bin Zheng [b], Roman Bednarik [c], M. Stella Atkins [a]

[a] School of Computing Science, Simon Fraser University, Burnaby, BC, Canada V5A 1S6
[b] Department of Surgery, University of Alberta, Edmonton, AB, Canada T6G 2E1
[c] School of Computing, University of Eastern Finland, Joensuu FIN-80101, Finland

ABSTRACT

Pupillary response is associated with perceptual and cognitive loads in visual and cognitive tasks, but no quantitative link between pupil response and the task workload in visual – motor tasks has been confirmed. The objective of this study is to investigate how the changes of task requirement of a visual – motor task are reflected by the changes of pupil size. In the present study, a simple continuous aiming task is performed and the task requirement is manipulated and measured by Fitts' Index of Difficulty (ID), calculated for different combinations of the target size and movement distance. Pupil response is recorded using a remote eye-tracker. The results show that event-triggered pupil dilations in continuous aiming movements respect Fitts' Law, such that higher task difficulty evokes higher peak pupil dilation and longer peak duration. These findings suggest that pupil diameter can be employed as a physiological indicator to task workload evoked by the task requirement in visual – motor tasks.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Evaluating the task workload of visual – motor tasks, and specifically the tasks' impact on the mental load of the user, is of great importance in monitoring and managing the workload in various tasks and systems, such as designing human – machine interfaces (Bailey and Iqbal, 2008; Goldberg and Kotval, 1999; Iqbal et al., 2005; Pomplun and Sunkara, 2003) and evaluating the mental workload of operators in high skill demanding work environments such as surgeons (Carswell et al., 2005; Zheng et al., 2010) and vehicle and aviation operators (Kun et al., 2012; Palinko and Kun, 2012; Veltman and Gaillard, 1998). Pupillary response has been extensively investigated and found to be a reliable indicator for the changes of cognitive loads in various cognitive tasks such as mental arithmetic and memory recall (Ahern and Jackson, 1979; Hess and Polt, 1964; Kahneman and Jackson, 1966); however, the relationship between pupil response and the changes of mental workload induced by physical demands in visual – motor tasks such as goal-directed movements have not been thoroughly explored. The confirmation of such relationship would expand the ability of using pupil diameter to indicate mental workload in visual – motor tasks.

The pioneering work, conducted by Richer and his colleagues in the 1980s (Richer and Beatty, 1985; Richer et al., 1983), examined the pupil response to a simple finger flexion movement during key-pressing. They found a connection between pupillary response and the task complexity; when the subjects were required to press buttons with increasing number of fingers, the amplitude of pupil response increased. Typically, the pupil started to dilate around 1.5 s before the finger movement and the pupil reached its peak size 0.5–1.0 s after the movement. However, Richer's task, only involving various numbers of fingers in the movement, was not a testing of real-world eye – hand coordination; the participants' eyes fixed at the center of the screen throughout the task for the purpose of pupil size recording and isolating perceptual load out of movement. In ordinary everyday motor tasks, eyes are usually involved and the movements are continuous and complex, such as walking, driving, and playing ball games (Land, 2006). Furthermore, the task difficulty of Richer's study did not have a quantitative definition of the task requirement, i.e. the task difficulty was represented by the number of the fingers involved in the flexion instead of being scaled such as using Fitts' Index of Difficulty.

To investigate the quantitative association between pupillary responses and the task requirement of motor tasks, we adopted the concept from Fitts' study of the information processing model between environmental stimulation and human response. In aiming tasks, Fitts defined task requirement using the Index of Difficulty (ID), where increasing ID is predicted by increases of tool travel distance and decreases in target size, and the performance

---

(movement time) is correlated with the ID, which is named Fitts' law (Fitts, 1954; Fitts and Peterson, 1964). Fitts' law is a fundamental method of quantitating task difficulty evaluation in HCI research and design due to its strong predictive power (Kopper et al., 2010; MacKenzie, 1992).

Before examining the pupil responses to the task requirement in continuous pointing movements, we examined the pupil responses to the task requirement in a discrete Fitts' pointing task (Jiang et al., 2014a), where the subjects were required to move a tool to touch a circle with varying target sizes and distances and every movement was preceded by a 10 s wait. We found a small but significant dilation starting about 1.5 s before the tool started to move, followed by a slight constriction, the "valley" in the pupil size profile. Before the tool touched the target (2 s after the tool starts to leave), the pupil reached its peak size. Both the pupil dilation and the duration from Valley-to-Peak size positively correlate with the increase of IDs. This evidence indicates that the change of pupil diameter is regulated by task requirement. This finding was confirmed by a second study to determine whether the target size or target distance has an independent influence on the pupil response (Jiang et al., 2014b).

The above two studies documented the connection between pupillary response and task difficulty in discrete pointing tasks. The subjects were instructed to wait 10 s before taking the next aiming movement to ensure that the recorded pupil response would not be affected by the previous movement, and the pupil had time to return its baseline size. Examples of discrete visual−motor tasks in daily life include inserting a key into a lock, shooting a basketball, and mouse-clicking at a specific location in an editor. However, in reality of everyday interactive tasks, continuous tasks are more common, such as steering a vehicle, playing ping-pong, and selecting an item in a multiple-level cascade menu. In many cases, the continuous movement frequency is higher than the pupil response frequency which is typically lower than 0.5 Hz (Jiang et al., 2014a; Richer and Beatty, 1985), pupil response is inevitably affected by multiple movements. It is a high time to carefully examine pupil response and develop a method to distinguish if pupil response is a reaction to an upcoming movement or is just a residual effect from a previous movement. We therefore explore the pupil responses to the change of task requirements in a continuous movement such as continuous aiming tasks, with the following research questions in mind. First, is there a difference between the patterns of pupil size responses between discrete and continuous visual−motor tasks? Second, is the change of pupil size still able to distinguish task difficulty in continuous visual−motor tasks?

We conducted the present study using a similar experimental setting as that in the discrete movement study (Jiang et al., 2014a) but here the participants performed a continuous pointing task without an extra waiting time between movements. We hypothesized that the pupil dilation will respect Fitts' Law in continuous movements, such that higher task difficulty evokes higher peak pupil dilation and longer peak duration. If the hypothesis holds, it may be possible to employ pupil diameter as a reliable physiological indicator to quantitatively measure task workload in continuous visually-guided motor tasks. Such measurements can be used for continuously adjusting proactive responses of user interfaces, for example in medical educational simulations involving visual−motor tasks.

2. Related works

2.1. Mental workload, task difficulty, and measurement methods

Mental workload is a finite mental resource that one uses to perform a task under specific environmental and operational conditions (Cain, 2004; Cassenti and Kelley, 2006). The amount of mental resource is limited for each individual. To achieve higher performance, the mental resource of a human operator must be managed effectively. For example, knowing how users' mental load fluctuates during interaction is critical in optimizing the human-centered interface design (Iqbal et al., 2005). Pupil diameter has been employed as an objective indicator of mental workload in various interactive tasks (Bailey and Iqbal, 2008; Iqbal et al., 2005; Wang et al., 2013). For example, Iqbal et al. (2005) explored pupil diameter changes in route planning and document editing tasks, purporting to find proper moments for low cost interruption; they found that the pupil diameter is relatively low during task boundaries, which are suitable for interruption with lower mental workload.

According to Wickens' 4-D multiple resource model (Wickens, 2002, 2008), mental workload is generally induced by task difficulty, generated from various sources such as perceptual load, cognitive load, and physical load. These sources correspond to different stages of perception, cognition, and manual response respectively. Perceptual load is the requirement to perceive more items during a visual searching task, cognitive load is related to the task demands on working memory in cognitive tasks such as mental arithmetic tasks, and physical load arises from physical demands typically in motor tasks (Backs et al., 1994; Chen and Epps, 2014). A task may involve multiple sources of loads. For example, a mental arithmetic task may involve perceptual and cognitive loads: the subject has to take in the question from either visual or acoustic channel (perceptual load), and then calculate from the items in the working memory (cognitive load). Chen et al. (2014) designed an experiment to separate perceptual load and cognitive load as two distinct sources of task difficulty, by manipulating five levels of difficulty of an arithmetic task performed in low and high perceptual load situations respectively.

In a visual−motor task such as target-pointing task, both perceptual load (visual) and physical load (manual−motor response) contribute to the task difficulty, but the latter usually dominates (Backs et al., 1994). Backs et al. (1994) separated perceptual and physical loads in a manual tracking task. In the case of a target-pointing task, visual perception is involved at the beginning of the task to perceive the global visual field of the task setting before the hand/tool moves towards the target to intake the specific target position information (Abrams et al., 1990; Elliott et al., 2001). In this paper, the difficulty of the designed target-pointing task refers to the difficulty arising from the physical demands shaped by the target size and target distance.

In order to evoke different levels of mental workload of a user in a study, the difficulty of the task has to be carefully manipulated. Most past studies manipulated the task difficulty by changing related task factors such as the complexity of the task. For example, Richer and Beatty (1985) defined four levels of task difficulty by varying the complexity of finger movement: one-finger flexion, two-finger flexion of one hand, one finger flexion of both hands, and three-finger flexion in one hand. This was not an analytic description of the task difficulty was not quantitative. The difficulty of two-finger flexion was not necessarily twice as hard as that of single-finger flexion, and even two-finger flexion in one hand might not be easier than one-finger flexion in both hands. In goal-directed movement tasks, the task difficulty is mostly from physical demands shaped by the target size and distance, and is governed by the law of speed-accuracy trade-off.

Fitts' law is a traditional model of human movement by analogy to the transmission of information (Fitts, 1954), and serves as a quantitative definition of difficulty in a variety of research areas, including kinematics, human factors, and human-computer interaction (HCI) (Guiard and Beaudouin-Lafon, 2004; Kopper et al., 2010; MacKenzie, 1992; Soukoreff and MacKenzie, 2004), and even

recently in the surgical environment (Prytz et al., 2012). Kourtis et al. (2012) explored the correlations between the electroence-phalography (EEG) signals and the Fitts' Index of Difficulty of an aiming movement, and found that the index of difficulty of the planned movement correlated linearly with the amplitudes of the specific components of EEG signals (N2 and P3b). In an eye-based interaction task where the participants were instructed to move their gaze from home position to target locations of various distance, Fitts' law has been found be valid to saccadic parameters, that is the peak saccadic velocity and the variability of the endpoint of saccades positively correlate to the gaze travel distance (Abrams et al., 1989; Al-Aidroos et al., 2008).

Once a concrete and operationalized and analytic description of the task difficulty such as Fitts' Index of Difficulty is established, an analytic model such as linear regression between the difficulty and other physiological indicators such as EEG and pupil responses could be built (Backs et al., 1994; Jiang et al., 2014b). The established model that maps the pupil responses to task difficulty could be further applied to situations where the task difficulty cannot be measured by Fitts' law.

However, Fitts' law has been rarely employed in experimental research on mental workload studies. Mental workload measurement techniques are usually categorized into three types: subjective rating scales (self-assessment), performance measures (including primary and secondary task measures), and psychophysiological measures (Gawron, 2008). The subjective rating methods are easy to perform, as they only require the operator to answer survey questions about the mental stress associated with a task. However, the answers to the survey questions are biased by the task difficulty and the measurement is not continuous as the survey is usually conducted at the end of the task (Zheng et al., 2010).

The subjective shortcomings may be avoided by using a secondary task to the primary task where decreased secondary task performance means increased mental workload caused by the primary task. However, there may be extra workload caused by the secondary task that affects the performance of the primary task.

Although psychophysiological measures are objective and generally do not interfere with task performance, most of them present an intrusion, as they need sensors attached to the user. Modern eye-tracking systems overcome these disadvantages as they can monitor the eye movements remotely and output pupil size data as a by-product.

2.2. Pupil dilation as an indicator of mental workload in cognitive tasks

The pioneer work exploring pupil size changes relating to cognitive workload was done by Hess and Polt (1964) in a mental arithmetic experiment. The authors found that the pupil size of the subjects gradually dilated along with the time elapse of presentation of a multiplication problem and reached a peak immediately before the production was orally reported; then constricted rapidly back to the original size. The mean pupil dilation was also found to be a function of the level of difficulty of the problem. Following this work, extensive studies have shown that the changes of pupil size reflect mental workload and the level of difficulty in different tasks, e.g., mental arithmetic tasks (Bradshaw, 1967), recall or memory tasks (Goldinger and Papesh, 2012; Otero et al., 2011), and visual search tasks (Privitera et al., 2010).

Pupil size also responds to the critical event during information processing, which is called Task-Evoked Pupil Response (TEPR), appearing at the event onset with a short latency (averaging between 100 ms and 200 ms), and terminating rapidly following the completion of the event (Beatty, 1982). The TEPR has been employed as an approach to capture and evaluate the mental workload changes during a variety of tasks (Beatty, 1982).

2.3. Pupillary responses to mental workload in visual motor tasks

In the field of eye – hand coordination in health care training, Jiang et al. (2013) investigated the pupil changes in a simulated laparoscopic task, and found that the pupil dilates faster during the execution of higher task difficulty subtasks, such as dropping a peg into a tiny cup, than during the execution of tasks with lower difficulty, such as touching a centrally located home position with the operating tool. However, the difficulty for each of the tasks was not well defined in their study.

Marshall (2000) reported the Index of Cognitive Activity (ICA) that is capable of capturing subtle cognitive changes from pupil metrics, and was used to predict the expertise of surgeons, together with other eye metrics (Richstone et al., 2010). However, details of the pupil response to motor tasks were not reported. It would be interesting to pinpoint the pupil dilation pattern during a very elementary motor task, which would provide a solid foundation for developing methods of objectively measuring mental workload of motor tasks using pupil diameter.

The pioneering work of exploring how the pupil responds to a single movement was done by Richer and Beatty (1985). The participants in this study were required to perform a self-paced finger flexion (key-pressing) task while looking at the center of the screen with their pupil diameter recorded. The authors found the typical pupil dilation pattern to the preparation and execution of a simple movement is that pupil dilates at around 1.5 s before the movement and peaks afterwards at around 0.5 s.

Before examining the pupil responses to the task requirement in continuous pointing movements, we examined the pupil responses to the task requirement in a discrete Fitts' pointing task (Jiang et al., 2014a), where the subjects were required to move a tool to touch pairs of circles with varying target sizes and distances and every movement was preceded by a 10 s wait. We found a small but significant dilation starting about 1.5 s before the tool started to move, followed by a slight constriction, the "valley" in the pupil size profile. Before the tool touched the target (2 s after the tool starts to leave), the pupil reached its peak size. Both the pupil dilation and the duration from Valley-to-Peak size positively correlate with the increase of IDs. This evidence indicates that the change of pupil diameter is regulated by task requirement. This finding was confirmed by a second study to determine whether the target size or target distance has an independent influence on the pupil response (Jiang et al., 2014b).

The above two studies documented the connection between pupillary response and task difficulty in discrete pointing tasks. The subjects were instructed to wait 10 s before taking the next aiming movement to ensure that the recorded pupil response would not be affected by the previous movement, and the pupil had time to return its baseline size. However, in goal-directed interactive movements, such as aiming at a target and selecting a menu item, the tool or hand move continuously with the coordination of eye movements, which evoke a different pupil response pattern. Therefore, the present study investigates pupil responses to continuous visual – motor tasks, and confirms whether Fitts' law applies to the pupil responses to task workload in continuous visual – motor tasks. The discovery of this knowledge would have important implications. For example, interactive environments could continuously adjust the presented information to accommodate the workload of the user in real-time. Design guidelines for fundamental movements of a tool could be objectively evaluated, and training and simulation systems could be developed with specific tasks toward the improvement of motor and aiming skills.

3. Methods

3.1. Participants

Fourteen participants (three females) were recruited to the study—including three graduate students, seven undergraduate students, and four staff members from the University of Alberta. All were right-handed and had normal or corrected-to-normal vision. None were previously trained in surgical procedures. The research was approved by the Health Research Ethics Board of University of Alberta. Each subject signed consent form before starting to perform the trial.

3.2. Apparatus

The apparatus is shown in Fig. 1, where a one-handed grasper is used to perform motor actions inside an illuminated training box, where the scene inside the box is recorded by a built-in camera and displayed on a monitor. This kind of box is commonly used for training surgeons in endoscopic techniques (Atkins et al., 2012). The participants had to move the grasper horizontally inside the training box to point at the horizontal circles printed on a paper laid flat on the bottom of the training box, while viewing the scene inside the box that was displayed on a monitor. The parameters of the target (the circles) setting, i.e., the target size and the distance between the targets, shown in Fig. 2, are the same as those in the discrete study (Jiang et al., 2014a). We designed three difficulty levels in this experiment according to the Fitts' Index of Difficulty (ID) (Fitts, 1954), i.e. Easy (ID1=2.7 bits/response, W1=0.9 cm, A1=3 cm), Medium (ID2=4.3 bits/response, W2=0.6 cm, A2=6 cm), and Hard (ID3=5.9 bits/response, W3=0.3 cm, A3=9 cm). The equation for calculating ID1, ID2, and ID3 is in (1).
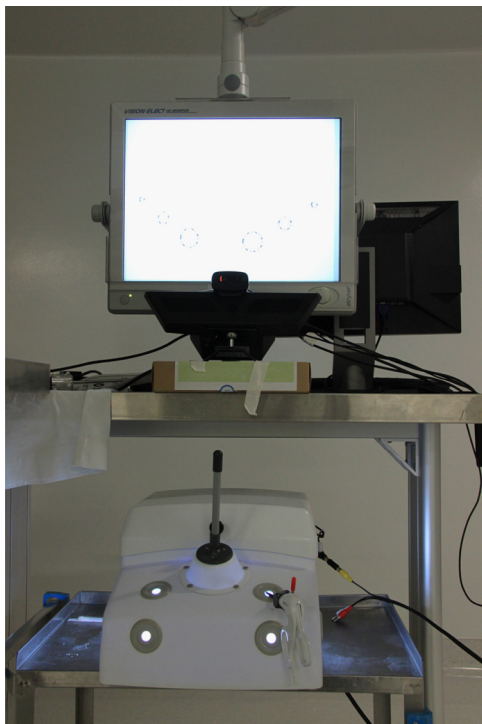
$$ID = \log_2 \frac{2A}{W} \qquad (1)$$

Fig. 1. Experimental setting including Tobii X50, training box, and web camera.
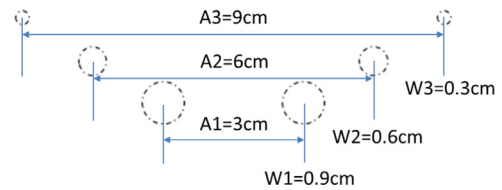

Fig. 2. The target setting of the task with combined variation of target size (W1 – W3) and distance (A1 – A3). The task was to move a tool to point between the circles in a Laparoscopic training box.

The three difficulty levels of targets were organized in two target settings by their execution sequence. Target setting 1 is shown in Fig. 2, executed from Easy task (bottom circles) to Hard task (top circles), and target setting 2 is similar to target setting 1, except the Hard target was on the bottom, so the task was executed from Hard to Easy.

The experimental setting includes a remote eye-tracker (Tobii X50, Tobii Technology AB), a laparoscopic training box (Laparoscopic Trainer, 3-D Technical Services, Franklin, OH), and a computer system. The participants performed the task using the training box, holding a surgical grasper at a standing pose about 60 cm from the eye-tracker. The work area inside the training box was illuminated and captured at 30 Hz by a built-in camera, and was projected on a 17 in. display screen. The eye-tracker was fixed on a height-adjustable desk and the display was hung over the eye-tracker using a surgical bracket (Stryker, USA); both were adjusted to fit the height of the participants. The eye movements of the participants on the screen during the task execution were recorded using the eye-tracker. A web camera was attached on the top center of the eye-tracker to record the facial expressions of the participant for the purpose of identifying eye blinks and head motions that result in lost data. The eye-tracker and cameras were integrated and automatically synchronized using Tobii software, Clearview 2.7.0 in the computer system. The setting was located in a well-isolated surgical simulation room where the ambient lighting was kept relatively stable. The brightness and contrast of the display were constant and set to a moderate level to make the pupil to work at the center of the length-tension curve of the iris muscle, for the best pupil response to task requirements (Privitera et al., 2010).

3.3. Tasks and procedure

The task was to continuously move the surgical tool to point to the circles printed on a piece of paper placed horizontally at the bottom inside the training box. A trial consisted of 6 steps executing from bottom to top, requiring touching pairs of targets (phase 1, steps 1–3) and then from top to bottom (phase 2, steps 4–6), as shown in Fig. 3. The trial started by placing the tooltip on the right bottom circle for 20 s for the baseline recording for the whole trial. The participant was then aurally instructed to move the tool to the left bottom circle and back to the right circle repetitively for ten times without any waiting time in-between movements (step (1)). After step 1 the tool was moved to the middle right circle after staying for 10 s on the bottom right circle, then similarly steps 2 and 3 were performed (Fig. 3, Phase 1). Right after step 3, on the same target setting, step 4 started with a 10 s pause on the top right circle, then steps 5 and 6 were completed with a 10 s pause between each step (Fig. 3, phase 2). The trial was ended by stopping the tooltip on the right bottom circle for 10 s after step (6). Only the 120 horizontal movements were used for analysis.

The participants were instructed to move the tool and touch the target as accurately and quickly as possible. Each trial took about 7 min.
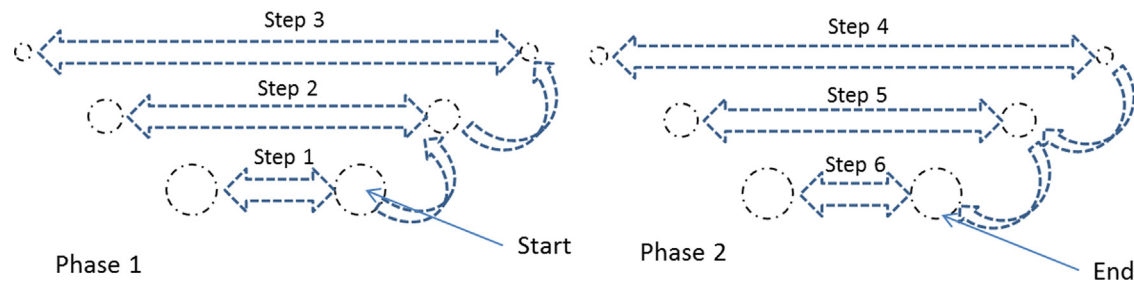
Fig. 3. The execution sequence of target setting 1, including phase 1 (Step 1 − Step 3) and phase 2 (Step 4 − Step 6). Each step includes moving the tool-tip from right to left and then left to right 10 times. The horizontal arrows represent 20 tool movements from one circle to another within a step, and the vertical arrows represent Transporting to the next step. Before moving to the next step, the tool stops on the right side circle for 10 s for baseline recording.

Each participant read and signed the consent form before entering the study, and then read the instructions. The ethical consent was approved by the Research Ethics office of University of Alberta. The participants practiced the task for a few minutes, until they felt ready to begin. Each participant performed one trial either on task setting 1 (from easy to hard and then from hard to easy as shown in Fig. 3) or task setting 2 (by flipping the paper, the same execution sequence as in Fig. 3 but starting with the hard targets on the bottom and easy targets on the top). For counter-balance, half of the participants started the target setting 1 first and then performed the task setting 2, and the other half of the participants did the reverse. We did not consider other ordering options like executing from medium difficulty to hardest to easiest, because the ANOVA results showed there was no significant difference in movement time between the two groups, and no interaction effect of ID and group order.

We decided to use this typical aiming task which clearly relates to Fitts' law and has been employed in many motor control studies (Adam et al., 2000; Adam and Paas, 1996; Binsted et al., 2001; Elliott et al., 2001; Lavrysen et al., 2007; Smits-Engelsman et al., 2002). The task was designed to be self-paced, without any explicit stimulus to indicate the start of a movement during a trial, because we tried to measure the task workload in as natural a setting as possible. The task was performed under a laparoscopic environment using a long shaft surgical tool because we intended to increase the difficulty of the motor part of task which would induce higher physical load compared to the perceptual load (Backs et al., 1994; Prytz et al., 2012). Also, the eye movements were easily recorded under this setting by attaching a remote eye-tracker to the screen, because the eye gaze was mostly on the screen. We decided to use the target setting with fixed numbers and sizes of target circles on the paper, to minimize the perceptual load.

3.4. Data analyses

3.4.1. Tooltip locations

The tooltip positions were automatically extracted from the task videos using a customized video processing algorithm. The details of the algorithm can be found in the discrete movement study (Jiang et al., 2014a). The algorithm involves three major steps, as illustrated in Fig. 4. First, the RGB video was read in frame by frame and transferred to the gray-scale image format, and then was transformed into the black-and-white image format via binary-threshold, where mostly only the tool remained in the image. Second, the biggest connected object was searched for and identified as the tool, as shown in the white dashed rectangle in Fig. 4. Third, the coordinates of the left-top corner of the tool rectangle were used as tooltip position, as in the study the tool was always orientated towards the northwest. The determined tooltip positions (x and y coordinates in pixels) along with time-stamps were stored in a text file for further analysis. The tooltip

data were smoothed with a running-average-filtered using equally weighted four samples window.

3.4.2. Tool movement phases

A movement time for this continuous aiming movement is calculated from when the tool starts to leave (Tool-leave), to when the tool touches the target and leaves for the next movement cycle (Next Tool-leave) in Fig. 5. This definition of movement time follows that in Fitts's (1954) study, which includes the dwelling time (Adam and Paas, 1996) which is from when the tooltip touches the target to the start of the next movement.

As shown in Fig. 5, a movement is further divided into Transport and Landing phases (Elliott et al., 2001) to better distinguish the impact on the pupil response from a continuous movement. The Transport phase starts when the tool leaves (Tool-leave) for the target circle and stops when the tooltip arrives the vicinity above the target circle (Tool-reach). In the Landing phase the tool is slowly adjusted to touch the target circle before the immediate start of the next movement (next Tool-leave). The Transport and Landing phases can be detected automatically from the tooltip positions in the video from the detected moments of Tool-leave and Tool-reach. An algorithm written in Matlab script is used to detect the moments of Tool-leave and Tool-reach (Jiang et al., 2014a). The algorithm first finds the moment of the absolute tooltip movement peak velocity during a movement and then searches backward and forward respectively for the moments of Tool-leave and Tool-reach, by checking whether the absolute velocity of the tool is lower than a set threshold. The velocity thresholds for detecting the Tool-leave and Tool-reach moments are empirically determined, since there are abrupt changes of tool velocity at the moments of Tool-leave and Tool-reach. The Tool-leave and Tool-reach thresholds were both set to 30 pixels/s.

3.4.3. Pupil data processing

The pupil diameter data was processed and synchronized with the trajectory of the tooltip for the analysis of the pupil responses during tool movements, as shown in Fig. 5. The pupil diameter data was exported from Tobii Clearview to the text file (Combined Data file, CMD) and the tooltip positions were derived from the task video recorded by the built-in camera. The pupil data were in 50 Hz and the task videos were in 30 Hz with a resolution of 352×288 pixels.

Then a Butterworth low-pass filter with a cut-off frequency of 4 Hz was applied to the pupil diameter data, since frequency above 2 Hz of the pupil is considered as noise (Privitera et al., 2010).

3.4.4. Pupil responses to movement detection

Since the magnitude of workload-related pupil dilation is usually small (less than 0.5 mm) and mixed with other simultaneously ongoing changes caused by light reflex and other brain activities, it is a challenge for us to detect the pupil dilation caused
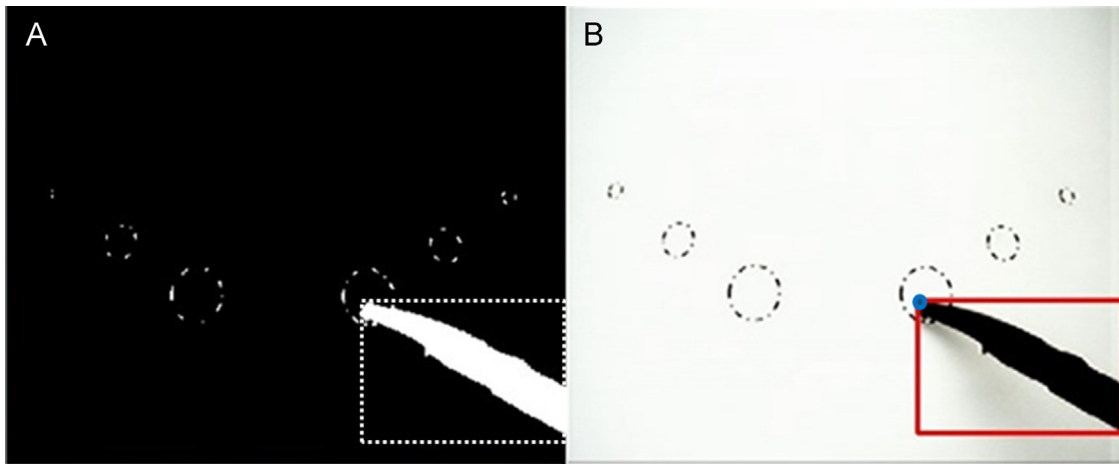
Fig. 4. An illustration of the process of tooltip detection from the task videos. Panel A shows the binary-thresholded image with mostly the tool left (in the dashed rectangle). Panel B shows the recognized tool (in red rectangle) and tooltip (the blue dot). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
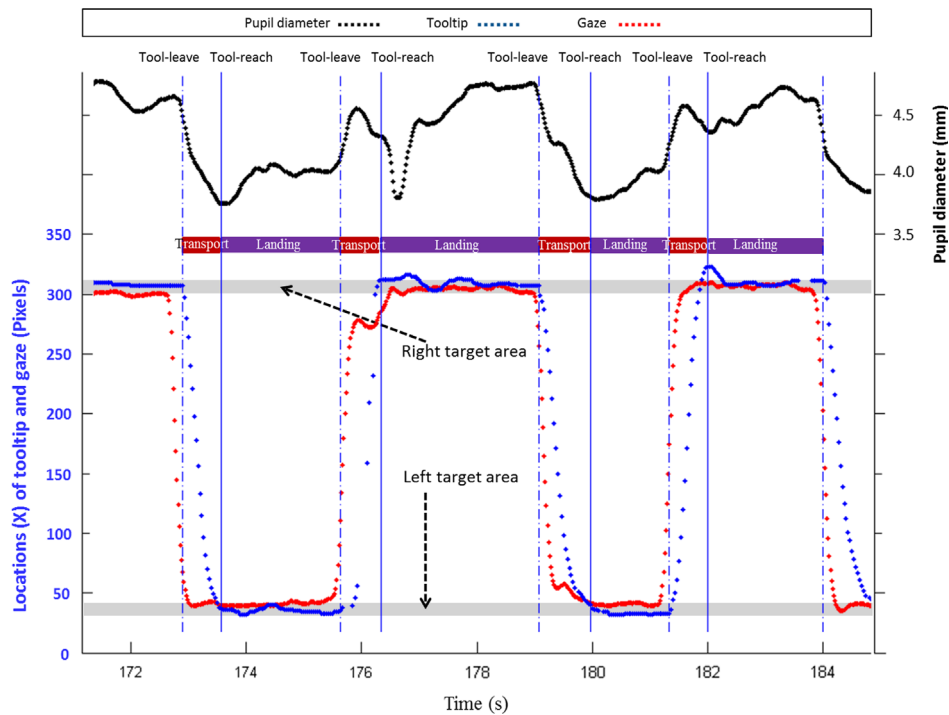

Fig. 5. An example showing four movements of hard ID execution (ID3, during 173–182 s in the trial from subject 1). The black curve on the top is pupil diameter, the blue curve is the tooltip position (*X* coordinate) in pixels along the horizontal line between the circles, and the red curve is the gaze locations (*X* coordinate). The vertical dash-dotted and solid blue lines represent the moments of start of a movement and when the tool reaches the vicinity above the target. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

by the workload directly among other influences to pupil dilations. One solution is to average tasks performed under the same condition to level off non-task related influence on the pupil response. The assumption behind this approach is, once the repetitions of the tasks were sufficient, the noise and factors affect pupil dilation would be averaged to zero.

The key step of successfully applying this signal averaging technique is to align repetitive tasks at a specific time stamp, such as at the moment of movement start. In our previous study on discrete tasks, the repetitive 7 s-windows of pupil responses were aligned at the moment of Tool-leave (the tart of Transport phase) to preserve the pupil dilation 1–2 s before the movement which is considered to be a reaction to the preparation of a movement (Jiang et al., 2014a; Jiang et al., 2014b). However, in the present continuous movement study, the preparation period for a

movement is overlaid with the Landing phase. Therefore, we decided to align the task repetitions at the moment of Tool-reach (the end of Transport phase) to preserve the pupil diameter changes in both Transport and Landing phases. Also, a shorter time window i.e. 4 s-window was employed to extract and show pupil diameter data in this continuous movement study, since the longest movement time of the hardest task was 3.2 s.

Another problem in pupil response analysis is that the pupil diameter baseline may drift during a task, and usually baseline techniques are used to normalize the pupil changes and eliminate the effect of drift of mean pupil size. The way that the baseline was chosen, e.g. the position of the baseline, may greatly affect the results of the data analysis (Bednarik et al., 2012). The length of baseline is usually from 100 ms to 1000 ms (Chen and Epps, 2014; Moresi et al., 2008; Privitera et al., 2010). The ideal baseline for

event-related pupil diameter change analysis should be derived from the period in the vicinity of the event where the pupil is as free from the effects of the events as possible. In the present continuous aiming task, since there was no waiting time between movements, we considered the 400 ms period at the beginning of the Landing phase, when the eyes start to focus on the new target and the pupil starts to dilate. We did not consider a shorter baseline e.g. 200 ms or 100 ms because ANOVA showed no significant difference of mean baseline pupil diameter between 400 ms, 200 ms, and 100 ms baseline length.

Adjusted pupil diameter changes in the 4 s-window were derived by subtracting the baseline pupil diameter from each sample. The baseline was the mean of the pupil size during the pre-determined period. All the data in the windows were aligned at the moments of Tool-reach (2 s into the 4 s window), and the mean pupil diameter change was calculated for each time point in the window across all horizontal tooltip movements from all trials for all subjects. Similarly, the mean pupil diameter changes were calculated across all movements from all trials for each ID. The mean pupil diameter changes in the 4-s window were plotted in a graph for visual analysis.

To examine which parts of the pupil size changes in the 4-second window have significant differences between the three IDs, we applied a paired *t*-test to the same time point samples across the windows and examined all the *p*-values along the time axis to determine which segments of the curves are significantly different. Due to the temporal autocorrelation of pupil waveform, we considered a series of more than 4 consecutive samples (80 ms) with *p*-values < 0.05 as significantly different (Privitera et al., 2010; Siegle et al., 2004).

3.5. Experiment design

The experiment had an independent variable task difficulty with three ID levels of Easy, Medium, and Hard. The dependent variables were movement time, Valley-to-Peak pupil dilation, Valley-to-Peak pupil dilation duration, Peak-to-Valley pupil constriction, and Peak-to-Valley pupil constriction duration.

The total number of movements was 1680 (14 participants, each performing 1 trial, each trial has 6 ID executions (each ID was executed twice), with each ID execution having 20 moves—10 times from right to left and from left to right, for a total of 120 movements for each subject).

4. Results

From the 1680 movements recorded by the 14 participants, we excluded the first 2 moves and last 2 moves from each ID execution (96 moves remained for each subject), and also discarded a total of 19 movements due to mis-operation, for example when the subject moved the pointing tool to a wrong target. Therefore we obtained 1325 valid movement recordings.

4.1. Accuracy

Among the 1325 movements, a total of 73 movements' endpoints (the tooltip positions when touching the target) were outside the target circle (63 for Hard ID, 10 for Medium ID, and none for Easy ID). Most movements' end points are within 2.5 mm on the target paper (corresponding to 10 mm displayed on the 17 in. screen) to their corresponding target edges (46 movements). The 27 movements with their endpoint distance greater than 2.5 mm to their target circle edges were discarded. The remaining 1298 valid movements were further analyzed.

4.2. Missing data

During execution of the 1298 valid movement trials, missing data accounted for 1.6% of the time. Segments of missed pupil data shorter than 100 ms mainly due to blinks (0.07%) were linearly interpolated. Longer missing data segments were ignored; the pupil size data during these periods was not included in the statistics. All subjects provided enough good quality data that no trials were discarded.

4.3. Performance

The movement time (MT) for a complete movement started from when the tool left the current target, and ended when the tool left that target for the next movement (the next Tool-leave), such that the MT was the interval between two consecutive Tool-leave moments. The mean MT for all movements was 2.0 ± 0.8 s, and the MT increased when the difficulty level increased (Fig. 6). There was a significant main effect of ID in terms of mean MT ($F_{2,1295} = 412.258$, $p < 0.0001$). Post Hoc test (Tukey HSD) showed significant differences ($p < 0.0001$) between pairs of the three IDs, with mean MT of Easy, Medium, and Hard IDs being 1.3 ± 0.5 s, 1.9 ± 0.6 s, and 2.6 ± 0.8 s respectively.

To distinguish the impact on the pupil response from a continuous movement, we divided the entire movement into Transport and Landing phases (Elliott et al., 2001). The phase separation was defined by the kinematics of the tool. The Transport phase started from the moment of Tool-leave to Tool-reach (when the tool fast moved to the vicinity above the target, see Methods), and the Landing phase occurred between the Tool-reach and the next Tool-leave (when the tool descended and touched the target, see Methods). The mean MT of both Transport and Landing phases were significantly different ($p < 0.0001$) between the three IDs.

4.4. Pupil responses to tool movements

Fig. 7 shows the mean changes of pupil diameter averaged for all IDs for all subjects. Data were aligned at the Tool-reach moment which was the end of Transport phase. The mean pupil diameter change was shown from the baseline over time. The baseline pupil diameter was the mean pupil diameter over a 400 ms period at the Tool-reach moment in each window, as shown in the dashed red rectangle. The vertical red line indicates the moment of Tool-reach where all the data were aligned. The blue and yellow bars represent Transport and Landing phases respectively. The pupil started to constrict about 150 ms into the Transport phase and dilated in the Landing phase with an approximate 100 ms delay; this is the typical pattern of pupil size changes found in a single movement in this study.

Fig. 8 shows the mean changes of pupil diameter in a 4 s-window aligned at the Tool-reach moment for each of the three IDs. The curves of the pupil response to the three IDs are clearly separated in both Transport and Landing phases; the hard ID constricts from a higher peak pupil size to the baseline in the Transport phase and dilates from the baseline to a higher peak pupil size in the Landing phase.

Table 1 shows the extent and duration of the pupil diameter changes between the three IDs in Transport and Landing phases. ANOVA results showed that there were significant main effects in terms of Peak-to-Valley pupil constriction ($F_{2,1295} = 157.463$, $p < 0.0001$) and Peak-to-Valley duration ($F_{2,1295} = 144.223$, $p < 0.0001$) in the Transport phase and Valley-to-Peak pupil dilation ($F_{2,1295} = 157.052$, $p < 0.0001$) and Valley-to-Peak duration ($F_{2,1295} = 408.863$, $p < 0.0001$) in the Landing phase between the three IDs. The values of the above four pupil diameter parameters
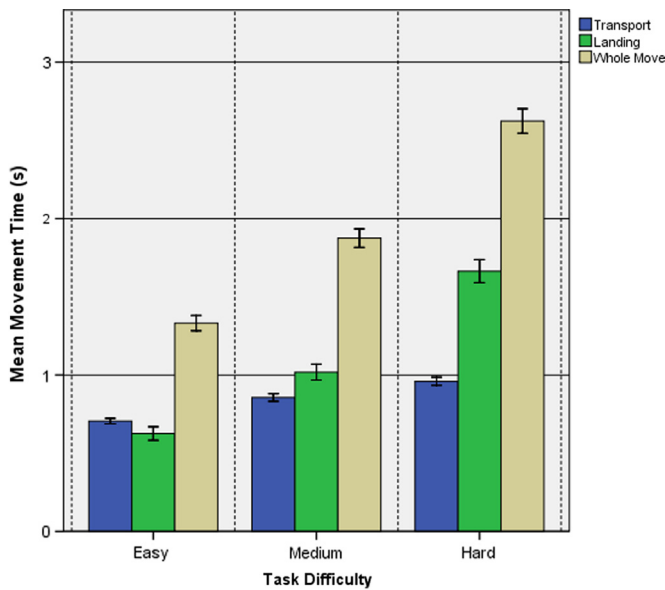
Fig. 6. Mean movement time for movement phases and a complete movement for different IDs for all subjects. The blue bar is the movement time for the Transport phase, the green bar is the movement time for the Landing phase, and the mustard bar is the movement time for the complete movement. The error bars are for 95% confidence intervals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
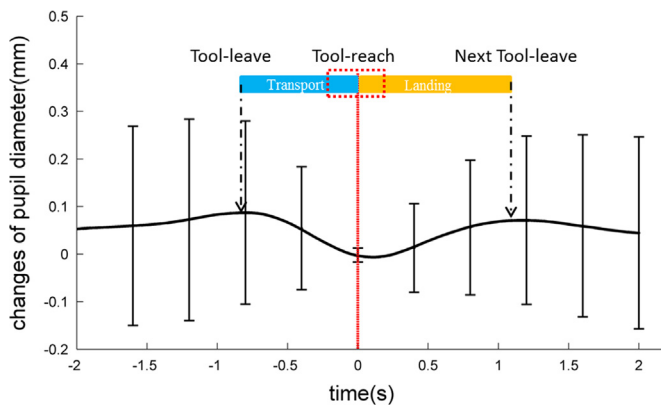

Fig. 7. Mean pupil diameter changes for 1298 valid movements from all IDs from all 14 subjects. Data were aligned over a 4-second window starting 2 s before the Tool-reach, shown by the vertical red line. The solid black curve is the mean pupil diameter change from the baseline over time. The pupil baseline diameter is the mean pupil diameter over a 400 ms period at the Tool-reach moment, as shown in the dashed red rectangle. The error bars for 1 std. dev. are drawn every 400 ms. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Peak-to-Valley pupil constriction, Peak-to-Valley duration, Valley-to-Peak pupil dilation, and Valley-to-Peak duration) positively correlated with IDs ($p < 0.0001$).

A paired t-test was performed for each time point in each pair of IDs containing samples from the 14 subjects, with the results as shown in the bottom colored bars in Fig. 8. The colored areas of the bars at the bottom of Fig. 8 indicate significant differences ($p < 0.05$), with the black comparing Easy and Hard, the blue comparing Medium and Easy, and the purple comparing Hard and Medium. Almost all moments along the 4 s window are significantly different between pairs of three IDs except in the area of the Tool-reach moment (baseline).

Fig. 9 shows mean pupil diameter changes for 3 different IDs for each subject. Most subjects have distinguishable pupil dilation patterns between three IDs except subject 11 which exhibited a relatively small extent of Peak-to-Valley constriction and Valley-
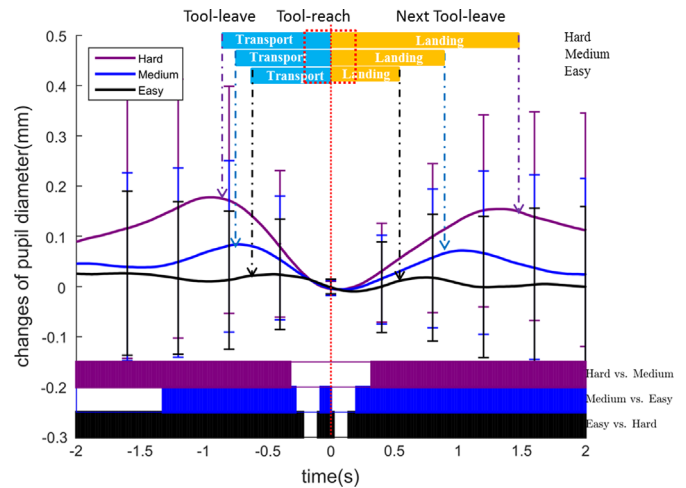

Fig. 8. Mean pupil diameter changes (subtracted by baseline) for 3 different IDs from all subjects; data are aligned over a 4-second window 2 s before Tool-reach, shown by the vertical red line. The three colored curves — black, blue, and purple — are the mean pupil diameter changes for Easy, Medium, and Hard ID respectively. The pupil baseline diameter is the mean pupil diameter over a 400 ms period at the 0 s (Tool-reach moment), as shown in the dashed red rectangle. The three rows of bars at the top represent Transport (light blue) and Landing (light orange) phases for the three IDs Hard, Medium, and Easy, respectively. The three colors of bars at the bottom indicate statistically significant differences in pupil dilation between Easy, Medium and Hard ID with black representing Easy vs. Hard, blue representing Medium vs. Easy, and purple representing Hard vs. Medium. The error bars for 1 std. dev. are drawn every 400 ms. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
The mean pupil diameter changes and duration over all subjects for transport and landing phases, for different IDs.

Transport phase	Easy	Medium	Hard
Peak-to-Valley constriction (mm)	0.09 ± 0.10	0.16 ± 0.15	0.28 ± 0.21
Peak-to-Valley duration (s)	0.52 ± 0.33	0.68 ± 0.35	0.97 ± 0.48
Landing phase			
Valley-to-Peak dilation (mm)	0.08 ± 0.09	0.15 ± 0.14	0.26 ± 0.20
Valley-to-Peak duration (s)	0.53 ± 0.35	0.91 ± 0.46	1.44 ± 0.58

to-Peak pupil dilation.[1] ANOVA was performed for each subject, and all showed significant main effects of the task difficulty on the four pupil dilation parameters (Peak-to-Valley constriction and duration and Valley-to-Peak pupil dilation and duration).

5. Discussion

The results support our hypothesis, that the pupil responses to a movement in this continuous aiming task respect Fitts' Law. Our data revealed a pattern of pupil response to a single goal-directed movement in the continuous aiming task, where the pupil constricts during the Transport phase and dilates during the Landing phase (Fig. 7); this holds for different difficulty levels of movements, where harder task IDs elicit a higher magnitude of pupil response and longer pupil dilation/constriction duration (Fig. 8). Both ANOVA on the pupil diameter features (Peak-to-Valley pupil constriction and duration in Transport phase and Valley-to-Peak

[1] Although subject11 does not show clearly separation between three curves of pupil dilation in Landing phase in Figure 10, ANOVA shows significant effects of task difficulty on Valley-to-Peak pupil dilation and duration with ($F_{2,93}=7.491$, $p < .005$) and ($F_{2,93}=11.344$, $p < .0001$), respectively, since the Hard ID has significant higher Valley-to-Peak pupil dilation and longer Valley-to-Peak duration than those of Easy and Medium IDS.
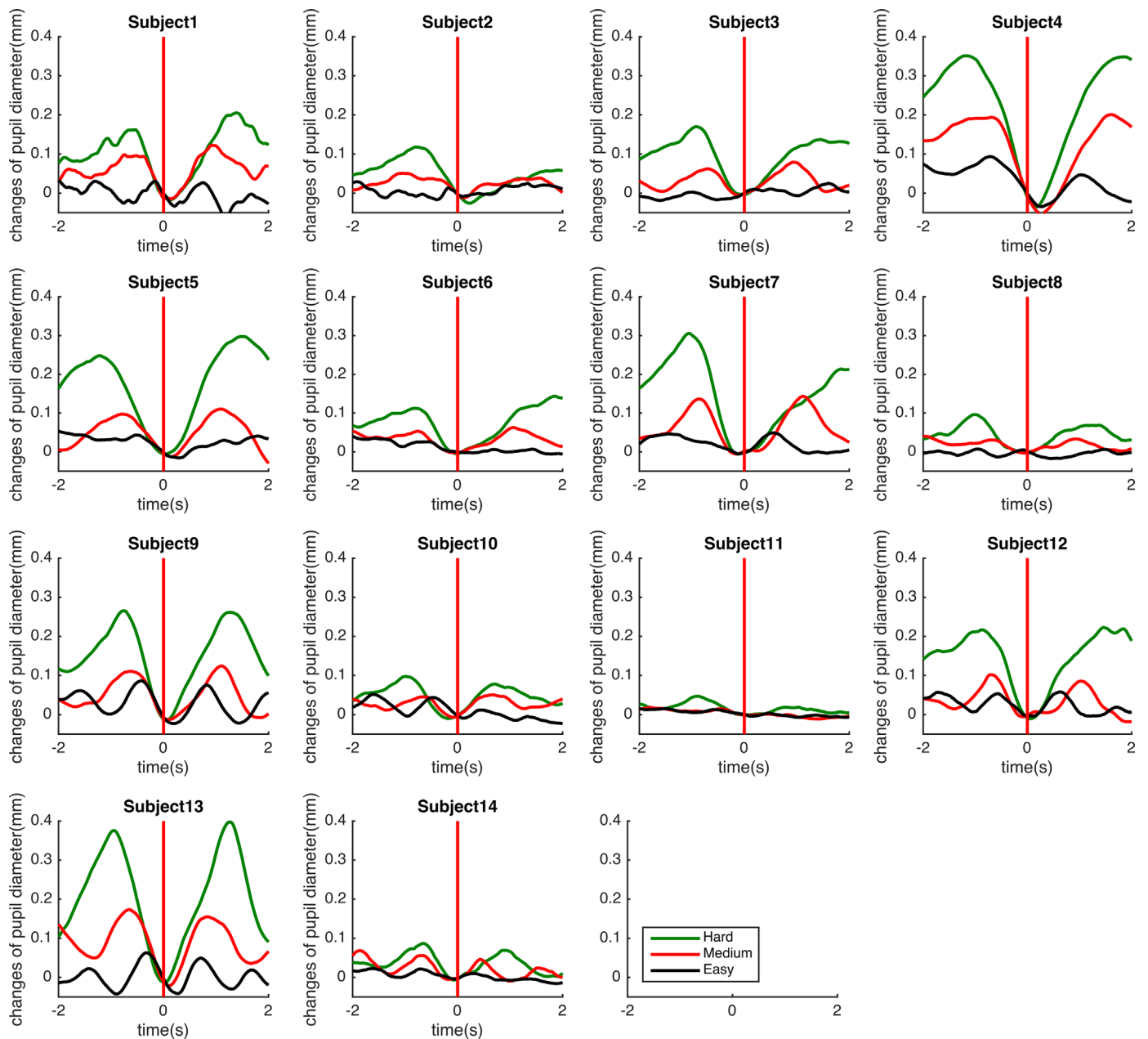
Fig. 9. Mean pupil diameter changes (subtracted by baseline) for 3 different IDs per each subject; data are aligned over a 4-s window 2 s before Tool-reach, shown by the vertical red line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

pupil dilation and duration in Landing phase), and moment-to-moment based t-test analysis on the changes of pupil size during a movement (the horizontal color bars on the bottom of Fig. 8) show significant differences between three IDs. The pupil size of Easy ID did not show a significant peak (Fig. 8). We believe the reason to be that in the continuous aiming task, the frequency of tool movement for Easy ID is higher (0.7 Hz) than that of pupil's response (typically lower than 0.5 Hz) (Jiang et al., 2014a; Richer and Beatty, 1985).

This research differs from previous pupillary response studies in several facets. First, it employed Fitts' Index of Difficulty to define task difficulty. Given the pupil-to-ID mappings derived from the present study such as linear regression between pupil size and ID value, we would be able to apply this model to measure task workload in motor tasks by only using pupil diameter when target size and distance are not measurable. For example in many surgery tasks such as grasping an object and knotting and cutting

a thread, the task difficulty cannot be measured by Fitts' law; whereas it is possible to measure these task workloads using the pupil diameter. Second, this study focused on measuring mental workload in visual–motor tasks, specifically exploring the pupil response to physical demands. The findings would expand the measurement of mental workload of cognitive tasks to motor tasks by using pupil diameter in human-centered HCI. For example, the existing interaction systems could add a model to automatically and proactively adapt the users' mental workload indicated by the pupil. Third, this study measured motor task workload in a natural setting employing a continuous and self-paced target-pointing task. There were challenges in data analysis, but the results offered pupil diameter a better opportunity to serve as an indicator of task workload in natural motor tasks.

As we intended to measure movement-evoked pupil responses in the present study, the cognitive and perceptive loads had to be excluded or minimized in the task design. In this simple target-

pointing task, there was no cognitive load involved; specifically, the subjects did not have to work with working memory, such as memorizing, calculating, or recalling items. The target layout was relatively simple with six fixed thin and dashed-line target circles on a light grey background. There was no need of searching for new items during movements. Perceptual load might only be slightly involved at the beginning of the task to perceive the global scene, but this had been diluted by adding the practicing session before the task – the participants had already been familiar with the target setting and the setting was unchanged throughout the whole experiment.

New challenges emerged during the analysis of pupil data in continuous, self-paced movement tasks due to the interference between two consecutive movements and no explicit stimuli onset. The challenges included how to identify the moment of stimuli onset and how to separate the effect of the overlap of pupil dilation. We employed video processing techniques to locate the moment of the tool starting to move as the stimuli onset of a movement. To separate the effect of the overlap of pupil responses to continuous tool movements, we divided a movement into Transport and Landing phases according to the kinematics of the tool movement, which was consistent with the definition in previous motor control studies (Abrams et al., 1989; Binsted et al., 2001). More importantly, the end of Transport phase was found to be a good timestamp for aligning windows, where the pupil starts to dilate in response to the Landing phase, after a constriction because of the tool' fast travel (the Transport phase). By dividing a movement into two phases and aligning tasks at the end of tool transportation, a clear pupil response to the continuous aiming movement was discovered. The pupil constricted during the Transport phase and dilated during the Landing phase; both the amplitude and duration of the pupil constriction and dilation positively correlated to Fitts' ID.

By doing this, we also found new results which were not identified in the prior discrete aiming movement studies (Jiang et al., 2014a; Jiang et al., 2014b), where the data were aligned at the moment of Tool-leave, with the major peak near the tool-reach moment. In this work, we saw peak pupil dilations occurred at around both the start of Transport and at the end of Landing phases (Figs. 7 and 8) with a constriction during the transport phase. The first peak at the beginning of the Transport phase in Fig. 8 is a result of the mixed effect of both the pupil response to the preparation of the tool movement and the residual of the pupil dilation from the previous movement. This is because, in the continuous movement, the $1-2$ s period before Tool-leave is also a Landing phase for the previous target, and pupil dilation is a combined response to the preparation for the Tool-leave and the execution of touching the target. By aligning the data at the moment of Tool-reach, we can see that pupil dilation is a combined response to these two events.

Second, pupil response at the moment of the Tool-leave differs most significantly between the three IDs (Fig. 8), where the end of the movement is followed immediately by a new movement. The effect of the pupil dilation after the end of movement is overlapped by the new movement which yields a significant pupil response. In contrast, in the discrete task (Jiang et al., 2014a), the most significantly distinguishable phase is around the end of each movement, since the pupil diameter develops to its peak followed by the 10 s waiting time.

This pupil response detection technique is very practical; in real-world applications, once the end of tool transportation is detected, the pupil diameter features including Peak-to-Valley constriction and duration and Valley-to-Peak pupil dilation and duration can be employed to distinguish the task difficulty of the movements. Furthermore, in cases of analyzing pupil responses to mental workload in multiple-tasks involving cognitive and motor components, which is

ubiquitous in real world situations such as listening to the radio while driving a car, we might be able to separate pupil responses to different types of workloads such as perception and physical loads, by using the present technique by aligning at different time points. Specifically, we could derive a pattern of pupil responses to the motor task by aligning the pupil signal windows at the moment of tool/hand reaching, and could obtain another pattern of pupil responses to the cognitive task by aligning the pupil signal windows at the onset of the cognitive stimulus. The individual differences shown in Fig. 9 should be considered when developing a system to access task workload in real world applications using pupil diameter. For example, a calibration procedure could be employed before a user's first use of the system.

The lighting condition and screen illumination of this experiment were well controlled. The ambient lighting in the case study room (a windowless room) was kept constant throughout the entire experiment data collection. The luminance of monitor that the subjects looked at during performing the task was kept constant as well. The background of the screen was uniformly grey. Furthermore, the baseline subtraction performed for each window and the averaging of the windows (signal averaging) would cancel most of the possible unrelated pupil responses.

Besides lighting conditions, one may be concerned that the pupil dilation or constriction are affected by the viewing angle to the eye-tracker camera due to the changes of gaze location (Brisson et al., 2013; Gagl et al., 2011) during target-pointing. In our case, the tool quickly moves from a target circle to another and the eyes mostly fixate on the targets, as shown in the red curve in Fig. 5. To clarify the gaze angle problem, we only need to check whether the baseline pupil diameter of movements in two different directions (moving from right to left vs. from left to right) are significantly different. In our two discrete studies (Jiang et al., 2014a; Jiang et al., 2014b), there is no significant difference between the baselines of left and right movements. Furthermore, we calculated a per-sample correlation between horizontal gaze position and pupil size using data from the center 10 s period of the 20 s baseline recorded at the beginning of each trial (see Tasks and Procedure) when the participants were free viewing a blank screen, and found no significant correlation between pupil size and horizontal gaze position ($Y = -0.000538 \times X + 3.7106$ for the left eye and $Y = -0.000357 \times X + 3.6825$ for the right eye, where Y: pupil diameter, X: horizontal gaze location in pixels from the left top corner of the screen). Therefore, we are sure that gaze angle does not affect the measured pupil size in this study.

Similarly, concerns may also be raised that the pupil response could be affected by the variation of the illumination for different sizes of the dashed black target circles (e.g. Easy ID has the biggest circles) (Kun et al., 2012). However, there was no significant effect of the target size (task difficulty) on the pupil baseline ($F_{2,1295} = 2.476$, $p > 0.05$).

The movement-related pupil response is a result of cortical activities which could be recorded more obtrusively by using scalp potential recording or neuroimaging technologies such as electro-encephalography (EEG) and magnetoencephalography (MEG) (Kilner et al., 2004; Kourtis et al., 2012; Nagamine et al., 1996). These techniques require attaching sensors to the human body, and are expensive. However, modern eye-tracking systems can monitor the eye movements remotely and unobtrusively, and pupil size data are available from most eye-tracking systems.

6. Conclusions

We examined the validity of pupil diameter as a physiological indicator to quantitatively, non-invasively, and continuously measure workload of motor tasks in continuous aiming movements. We showed a novel method of measuring the task requirements

during visually-guided physical tasks using Fitts' ID-to-pupil mapping. This allows measuring task workload in situations where the target size and/or distance are not readily discernable, such as during laparoscopic surgery.

Previous discrete movement studies showed that the pupil dilation positively correlates to task requirement modeled by Fitts' Index of Difficulty. The present study shows that pupil responses still respect Fitts' law in continuous aiming tasks, which are more complex and common than discrete aiming movements. This implies that pupil size changes can provide indicators for measuring task workload in goal-directed movements.

Acknowledgments

References

Abrams, R.A., Meyer, D.E., Kornblum, S., 1989. Speed and accuracy of saccadic eye movements: characteristics of impulse variability in the oculomotor system. J. Exp. Psychol.: Human Percept. Perform. 15, 529–543.

Abrams, R.A., Meyer, D.E., Kornblum, S., 1990. Eye – hand coordination: oculomotor control in rapid aimed limb movements. J. Exp. Psychol.: Human Percept. Perform. 16, 248–267.

Adam, J.J., Nieuwenstein, J.H., Huys, R., Paas, F.G.W.C., Kingma, H., Willems, P., Werry, M., 2000. Control of rapid aimed hand movements: the one-target advantage. J. Exp. Psychol.: Human Percept. Perform. 26, 295–312.

Adam, J.J., Paas, F.G.W.C., 1996. Dwell time in reciprocal aiming tasks. Human Mov. Sci. 15, 1–24.

Ahern, S., Jackson, B., 1979. Pupillary responses during information processing vary with scholastic aptitude test scores. Science 205, 1289–1292.

Al-Aidroos, N., Fischer, M.H., Adam, J.J., Pratt, J., 2008. Structured perceptual arrays and the modulation of Fitts's law: examining saccadic eye movements. J. Motor Behav. 40, 155–164.

Atkins, M.S., Jiang, X., Tien, G., Zheng, B., 2012. Saccadic delays on targets while watching videos ETRA 2012. Proceedings of the Symposium on Eye Tracking Research and Applications, pp. 405–409.

Backs, R.W., Ryan, A.M., Wilson, G.F., 1994. Psychophysiological measures of workload during continuous manual performance. Human Factors: J. Human Factors Ergon. Soc. 36, 514–531.

Bailey, B.P., Iqbal, S.T., 2008. Understanding changes in mental workload during execution of goal-directed tasks and its application for interruption management. ACM Trans. Comput. – Human Interact. (TOCHI) 14, 1–28.

Beatty, J., 1982. Task-evoked pupillary responses, processing load, and the structure of processing resources. Psychol. Bull. 91, 276–292.

Bednarik, R., Vrzakova, H., Hradis, M., 2012. What do you want to do next: a novel approach for intent prediction in gaze-based interaction, Proceedings of the Symposium on Eye Tracking Research and Applications, pp. 83–90.

Binsted, G., Chua, R., Helsen, W., Elliott, D., 2001. Eye – hand coordination in goal-directed aiming. Human Mov. Sci. 20, 563–585.

Bradshaw, J., 1967. Pupil size as a measure of arousal during information processing. Nature 216, 515–516.

Brisson, J., Mainville, M., Mailloux, D., Beaulieu, C., Serres, J., Sirois, S., 2013. Pupil diameter measurement errors as a function of gaze direction in corneal reflection eyetrackers. Behav. Res. Methods 45, 1322–1331.

Cain, B., 2004. A review of the mental workload literature. English 0, 1–34.

Carswell, M., Clarke, D., Seales, B., 2005. Assessing mental workload during laparoscopic surgery. Surg. Innovat. 12, 80–90.

Cassenti, D.N., Kelley, T.D., 2006. Towards the shape of mental workload. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50, pp. 1147–1151.

Chen, S., Epps, J., 2014. Using task-induced pupil diameter and blink rate to infer cognitive load. Human–Comput. Interact. 29, 390–413.

Elliott, D., Helsen, W.F., Chua, R., 2001. A century later: Woodworth's (1899) two-component model of goal-based aiming. Psychol. Bull. 127, 342–357.

Fitts, P.M., 1954. The information capacity of the human motor system in controlling the amplitude of movement. J. Exp. Psychol. 47, 381–391.

Fitts, P.M., Peterson, J.R., 1964. Information capacity of discrete motor responses. J. Exp. Psychol. 67, 103–112.

Gagl, B., Hawelka, S., Hutzler, F., 2011. Systematic influence of gaze position on pupil size measurement: analysis and correction. Behav. Res. Methods 43, 1171–1181.

Gawron, V.J., 2008. Human workload, Human Performance, Workload, and Situational Awareness Measures Handbook, second edition CRC Press, pp. 87–230.

Goldberg, J.H., Kotval, X.P., 1999. Computer interface evaluation using eye movements: methods and constructs. Int. J. Ind. Ergon. 24, 631–645.

Goldinger, S.D., Papesh, M.H., 2012. Pupil dilation reflects the creation and retrieval of memories. Curr. Direct. Psychol. Sci. 21, 90–95.

Guiard, Y., Beaudouin-Lafon, M., 2004. Fitts' law 50 years later: applications and contributions from human–computer interaction. Int. J. Human – Comput. Stud. 61, 747–750.

Hess, E.H., Polt, J.M., 1964. Pupil size in relation to mental activity during simple problem-solving. Science 143, 1190–1192.

Iqbal, S.T., Adamczyk, P.D., Zheng, X.S., Bailey, B.P., 2005. Towards an index of opportunity: Understanding changes in mental workload during task execution, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 311–320.

Jiang, X., Atkins, M.S., Tien, G., Bednarik, R., Zheng, B., 2014a. Pupil responses during discrete goal-directed movements, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2075–2084.

Jiang, X., Atkins, M.S., Tien, G., Zheng, B., Bednarik, R., 2014b. Pupil dilations during target-pointing respect fitts' law ETRA 2014. Proceedings of the Symposium on Eye Tracking Research and Applications, pp. 175–182.

Jiang, X., Zheng, B., Tien, G., Atkins, M.S., 2013. Pupil response to precision in surgical task execution. Stud. Health Technol. Inform. 184, 210–214.

Kahneman, D., Jackson, B., 1966. Pupil diameter and load on memory. Science 154, 1583–1585.

Kilner, J.M., Vargas, C., Duval, S., Blakemore, S.-J., Sirigu, A., 2004. Motor activation prior to observation of a predicted movement. Nature Neurosci. 7, 1299–1301.

Kopper, R., Bowman, D.A., Silva, M.G., Mcmahan, R.P., 2010. A human motor behavior model for distal pointing tasks. Int. J. Human – Comput. Stud. 68, 603–615.

Kourtis, D., Sebanz, N., Knoblich, G., 2012. Eeg correlates of Fitts's law during preparation for action. Psychol. Res. 76, 514–524.

Kun, A.L., Palinko, O., Razumenić, 2012. Exploring the effects of size and luminance of visual targets on the pupillary light reflex, Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 183–186.

Land, M.F., 2006. Eye movements and the control of actions in everyday life. Prog. Retinal Eye Res. 25, 296–324.

Lavrysen, A., Elliott, D., Buekers, M.J., Feys, P., Helsen, W.F., 2007. Eye – hand coordination asymmetries in manual aiming. J. Motor Behav. 39, 9–18.

Mackenzie, I.S., 1992. Fitts' law as a research and design tool in human – computer interaction. Human Comput. Interact. 7, 91–139.

Marshall, S.P., 2000. United States Patent No. 6090051.

Moresi, S., Adam, J.J., Rijcken, J., Van Gerven, P.W.M., Kuipers, H., Jolles, J., 2008. Pupil dilation in response preparation. Int. J. Psychophysiol. 67, 124–130.

Nagamine, T., Kajola, M., Salmelin, R., Shibasaki, H., Hari, R., 1996. Movement-related slow cortical magnetic fields and changes of spontaneous meg- and eeg-brain rhythms. Electroencephalogr. Clin. Neurophysiol. 99, 274–286.

Otero, S.C., Weekes, B.S., Hutton, S.B., 2011. Pupil size changes during recognition memory. Psychophysiology 48, 1346–1353.

Palinko, O., Kun, A.L., 2012. Exploring the effects of visual cognitive load and illumination on pupil diameter in driving simulators, Proceedings of the Symposium on Eye Tracking Research and Applications, pp. 413–416.

Pomplun, M., Sunkara, S., 2003. Pupil dilation as an indicator of cognitive workload in human-computer interaction, Proceedings of the 10th International Conference on Human – Computer Interaction (HCII), pp. 542–546.

Privitera, C.M., Renninger, L.W., Carney, T., Klein, S., Aguilar, M., 2010. Pupil dilation during visual target detection. J. Vis., 10, pp. 1–14.

Prytz, E., Montano, M., Scerbo, M.W., 2012. Using Fitts' law for a 3D pointing task on a 2D display: effects of depth and vantage point. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56, pp. 1391–1395.

Richer, F., Beatty, J., 1985. Pupillary dilations in movement preparation and execution. Psychophysiology 22, 204–207.

Richer, F., Silverman, C., Beatty, J., 1983. Response selection and initiation in speeded reactions: a pupillometric analysis. J. Exp. Psychol.: Human Percept. Perform. 9, 360–370.

Richstone, L., Schwartz, M.J., Seideman, C., Cadeddu, J., Marshall, S., Kavoussi, L.R., 2010. Eye metrics as an objective assessment of surgical skill. Ann. Surg. 252, 177–182.

Siegle, G.J., Steinhauer, S.R., Thase, M.E., 2004. Pupillary assessment and computational modeling of the stroop task in depression. Int. J. Psychophysiol. 52, 63–76.

Smits-Engelsman, B.C.M., Galen, G.P., Duysens, J., 2002. The breakdown of Fitts' law in rapid, reciprocal aiming movements. Exp. Brain Res. 145, 222–230.

Soukoreff, R.W., Mackenzie, I.S., 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. Int. J. Human – Comput. Stud. 61, 751–789.

Veltman, J.A., Gaillard, A.W., 1998. Physiological workload reactions to increasing levels of task difficulty. Ergonomics 41, 656–669.

Wang, W., Li, Z., Wang, Y., Chen, F., 2013. Indexing cognitive workload based on pupillary response under luminance and emotional changes, Proceedings of the 2013 International Conference on Intelligent User Interfaces, pp. 247–256.

Wickens, C.D., 2002. Multiple resources and performance prediction. Theor. Issues Ergon. Sci. 3, 159–177.

Wickens, C.D., 2008. Multiple resources and mental workload. Human Factors: J. Human Factors Ergon. Soc. 50, 449–455.

Zheng, B., Cassera, M.A., Martinec, D.V., Spaun, G.O., Swanstrom, L.L., 2010. Measuring mental workload during the performance of advanced laparoscopic tasks. Surg. Endosc. 24, 45–50.