# A potential application of pupillometry in web-usability research

*Eszter* Józsa

## Abstract

*Despite the increasing popularity of eye tracking technique, measurement of changes in pupil diameter has received considerably less attention.*

*However, for over two millennia it is unambiguous, by studies of acclaimed researchers (Eckhard Hess, Michel Pierre Janisse or Lowenstein and Loewenfeld) that cognitive and emotional brain activities are in connection with pupil size changes. The area, evolved by their profound and comprehensive research is referred as pupillometry.*

*Current review describes a potential application of pupillometry in web-usability research by applying ASL Mobile Eye device.*

## Keywords

*pupillometry · web-usability · eye tracking*

**Eszter Józsa**

Department of Ergonomics and Psychology, BME, 1111 Budapest, Egry J. u. 1., Hungary

e-mail: eszter.jozsa@gmail.com

## 1 Introduction

The pupils of the human eyes have been referred to as 'windows to the soul' by different writers of the past (Joshua Sylvester, 1591, and Gauillaume de Salluste) as 'observation points for a person's innermost thoughts' [1, p. 218]. Although examination of pupils is much less romantic than implied by these early writers, scientists have discovered some interesting links between pupillary motion and behavior.

Our pupil is solely an opening at the center of the iris, surrounded by the iris muscle through which light passes. The iris is capable of stretching and reducing the size of the pupil opening. The pupil of the eye can constrict to 1.5 mm in diameter and can dilate to about 8 to 9 mm [7].

Early devices in the measurement of pupil size used infrared photography, which technique permitted the assessment of pupil diameter, but it was limited in accuracy. At the beginning of the second half of the $20^{th}$ century, electronically based systems for observing and monitoring of pupil, were interfaced with newer and faster computer technologies to provide more accurate measurements. Today's commonly used systems can resolve better than 0.025 mm in diameter on individual measurements, at rates of 60 Hz, or even up to 240 Hz. Although the technology is continuously improving it is still important in pupillometric studies, which use visual stimuli, to carefully control psychophysiological aspects of the stimuli, such as light or hue [1, 2].

There are a number of environmental factors that can effect pupil size changes. Table 1 above has been adapted from W W. Tryon[12].

The general term pupillometry 'refers to the measurement of variations in the diameter of the pupillary aperture of the eye', as an indicator of attention, interest or emotion [1, p. 218].

### 1.1 Pupillometric measures in web-usability

Despite the increasing popularity of eye tracking technique, measurement of changes in pupil diameter, which can be recorded by most of the eye tracker devices, has received considerably less attention. Andreassi states that 'although many observations about changes in pupil size under different emotional or performance conditions have been reported, it is only

| Factor | Effect on pupil size |
|---|---|
| Darkness reflex | Momentary dilation due to interruption of a constant adapting light |
| Consensual reflex | Stimulation of one eye affects both eyes equally |
| Near reflex | Constriction due to decreasing the point of focus |
| Lid-closure reflex | Momentary contraction followed by redilation |
| Psycho sensory reflex | Restoration of diminished reflexes due to external stimulation |
| Age | Decreased diameter and increased variability with age |
| Habituation | Pupil diameter decreases, speed of contraction increases, and magnitude of reflex decreases with continued stimulus presentation |
| Binocular summation | Constriction greater when both eyes are stimulated simultaneously |

**Tab. 1.** Factors that may influence pupil size changes

within the last 30 years that the development of practical and reliable instrumentation has enabled the precise measurement of the pupillary aperture under different experimental condition and psychological states' ([1, p. 219]). Sirevaag and Stern assert that 'It is somewhat surprising that relatively little attention has been paid to this measure over the years' [10]. One reasonable answer could be the difficulty in quantifying the pupil diameter; however it has largely been eliminated with improved image processing technologies.

Pupil size varies in response to different levels of cognitive engagement or effort [5], and the interest value of stimuli [3]. Oliveira, Aula & Russel suggest in their recent publication that 'measure of pupil size have the potential to serve as non-intrusive (in case of a non-intrusive eye-tracker device) and objective physiological correlates of user interest'. This, along with the increasing utilization of eye trackers in usability laboratories, makes pupillometry an 'intriguing and promising technique for HCI research'[8]. In their book, Tullis & Albert [13] have committed themselves to the idea that measuring pupil diameter in everyday usability tests may be useful, particularly in situations where the focus is on the amount of mental concentration or emotional arousal.

The aim of our *pilot* study was to trigger, measure and analyze the known phenomenon that mental effort such as a cognitively demanding task causes increased pupil dilation, and when the solution is found a pupillary constriction occurs [4]), [6]). In other words, 'the amount of pupil response during mental activity has been shown to be a function of how hard an individual has to work'. [1, p. 224]) Besides measuring changes in pupillary dilation, eye gaze movements and gaze directions have also been recorded.

In current pupillometry based case study we have formulated three Hypotheses, which will be evaluated in the Discussion part.

### Hypothesis 1
The effect of increasing mental effort can be observed through the changes in pupil diameter under the described laboratory conditions with ASL Mobile Eye device.

### Hypothesis 2
During the failure task (Task 4) increasing mental effort can be observed and measured.

### Hypothesis 3
A pupillary constriction occurs when the participant gives the right solution. Our assumption is that this constriction can be measured in web-usability experiments as well, in case of successfully accomplished tasks.

**During pupillometric studies lightning conditions must be carefully controlled.**

### 2 Method
#### 2.1 Participants, Device, Tracking procedure, Stimuli
Following the guidelines of Nielsen Norman Group a total of 5 female and 3 male Hungarian participants ranging in age from 19 to 29 were invited to the study. In a recent publication the Nielsen Norman Group recommends a minimum number of 6 test users in qualitative eye tracking studies, which use gaze replays for the purpose of evaluation [9].

The eye tracking session lasted approximately 30 minutes. Users participated individually and were not paid. They all have clear vision; do not wear contact lenses or prescription glasses. Upon entering the hall, an explanation of the study's objectives was given, and they read a handout with the instructions. Afterwards, participants were guided to the experimental room where they were seated in front of the projector screen on which the stimulus webpage was projected from above. The centre of the screen was located at the subject's eye height and measured 1700 × 1300 mm. The distance between the eye and the screen was 2200 mm.

For the purpose of our experiment an ASL Mobile Eye device (see also Fig. 1) was applied. The eye tracker was calibrated on the participants' right eye. The Mobile Eye records data on 60 Hz by interleaving images taken from two cameras. One camera records the eye movements while the scene camera records the environment being observed by the user. Both image streams are documented on the same digital videotape by alternating the frames. Therefore, the actual functional sampling of data is 30 Hz.

The stimuli were a series of 5 tasks. Participants were asked to accomplish them at a certain webpage. They had to solve the tasks within the webpage but they were allowed to use search engines if they felt it necessary. The tasks were placed on the table in front of the test user, printed one task per sheet, upside down, in the correct order. Additional 'fake' pages were attached to the

**Fig. 1.** ASL mobile eye

handout. Participants were asked to follow the given order, do not return to a previous task, and turn a page if they are ready or want to quit the task. A blank sheet of paper and a pen was also placed on the table and the users were requested to write down their answers if the task requires it.

### 2.2 Tasks

**Task 1:** Navigate to the webpage (URL printed on the first sheet) from Google's main page!

**Task 2:** Please find the number of persons who were interested in and asked about Scholl products during 'Scholl'-display efficiency research!

**Task 3:** Please find and name an FMCG display design!

**Task 4:** Please find the company's work for 'Lindt' chocolates and write down the year of the display design! *(There is no display designed for 'Lindt' chocolates among the company's work.)*

**Task 5:** How many female employees work with the company?

The composition consists of three tasks that can be solved with increasing mental effort. The fourth is a failure task to trigger even harder effort, finally frustration and quitting. The last one is an easy, straightforward finalizing task.

### 2.3 Analysis

For the purpose of data processing and analysis the following method was implemented.

ASL Mobile Eye's default software is Eye Vision. The data output contains an .xls and an .avi file. The first file includes the time stamps, the coordinates of fixation location and pupil's position together with the pupil radius size.

Firstly, using gaze replays we have unfolded and reorganized the .xls files, in a way that we have put each task, from the beginning to the end, in a separate table, and visualized them on a separate graph. Successes and failures are represented in the

table below (Table 2) and also on each graph. 0 means failure/quitting while 1 is successful task accomplishment.

Due to insufficient calibration in case of Participant 3 and 5, respectively due to incomplete eye recording in case of Participant 8, starred items had to be dropped from the database. Items shaded with dark grey refer to participants, who have found out that there is no existing display design for 'Lindt' chocolates at the company's webpage, so thus they have successfully 'accomplished' the failure task. Light grey shading has been applied for participants who have previous marketing-related experience or knowledge.

For the purpose of demonstration, data of Participants 2, 4, 6 and 7 have been chosen. On the graphs below Axis X shows the elapsed time in mm:ss format while Axis Y represents the pupil radius data in number of pixels. Because of the significant differences in pupil size between individuals, direct comparing is not possible; pupil's construction and dilation can be analyzed for each participant individually. During the evaluation process we have compared the video replays with the data files, paying special attention to the very moments when participants found the solution or gave up the task. Through the following 'Results' section a detailed analysis will be given about our findings.

### 3 Results
#### 3.1 TASK 1

Since all the users hold Internet proficiency, they can easily handle a browser (Mozilla Firefox in our case), we have applied Task 1 as a warm-up task, participants could get used to the unusual circumstances (laboratory environment, eye tracker glasses) and their pupil returned to resting baseline.

#### 3.2 TASK 2

*Please find the number of individuals who were interested in and asked about 'Scholl' products during the company's 'Scholl'-display efficiency research!*

Our aim with this task was to induce participants to explore the page's content and its structure. The design worked well, users spent 4 minutes on average to solve the task (minimum of 1:14 mm:ss in case of Participant 2 and maximum of 8:33 mm:ss in case of Participant 6).

The solution can be found at the 'News' menu point, which is located in the main menu bar. To find the menu point is simple but the article is extensive, therefore to hit on this piece of information is time-consuming and requires very accurate search. Additionally, there is no built-in search module on the page and the article is in picture format, so thus search engines cannot produce any results. 5 participants out of 8 successfully accomplished the task and 3 of them gave it up.

Fig. 2a below represents the pupil radius data of Participant 4 during solving the task. The graph's minimum point at 02:35 refers to the point when the user finds the answer.

On Fig. 2b pupil data of Participant 2 is displayed. Using the video replays we have revealed that at 03:00 (the minimum point

| Task | Participant 1 | Participant 2 | Participant 3 | Participant 4 | Participant 5 | Participant 6 | Participant 7 | Participant 8 |
|------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Task1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Task2 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 |
| Task3 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| Task4 | 0 | 1 | 1* | 0 | 0 | 1 | 0 | 1* |
| Task5 | 1 | 1 | 1* | 1 | 1* | 1 | 1 | 1* |

of the graph) the participant turned a page but he did not start to read the next task. Apparently he worked on solving the task for an additional 40 seconds and started to read Task 3 at 03:40.

### 3.3 TASK 3

*Please find and name an FMCG hostess display design!*

Task 3 has been applied to increase the amount of mental effort needed to solve the task. There are two possibly right solutions (displays for Gilan Trading Kft. and for Procter & Gamble).

The sought products are located at the main page (landing page), however participants had to scroll down to the bottom of the page and use the two slides to choose the type of the display (hostess) and the field of usage (FMCG).

Participants spent 2:27 mm:ss on average to solve the task, minimum 1:14 mm:ss in case of Participant 2, maximum 4:32 mm:ss in case of Participant 4. 5 users have successfully accomplished Task 3, while 3 of them quitted.

Using the same methodology as previously, we have found the following results:

Fig. 3a below demonstrates the pupil radius changes of Participant 2. At 04:38 he has found one of the hostess displays and at this moment his pupil size reached the minimum.

Participant 6 has found both solutions. The moment when she found the first display is demonstrated on Fig. 3b. at 11:00, while the second solution was found at 11:49.

### 3.4 TASK 4

*Please find the company's work for 'Lindt' chocolates and write down the year of production!*

There is no existing display-design for 'Lindt' chocolates on the analyzed webpage. The task has been created to enhance mental effort, evoke frustration and finally abdication.

The most obvious way to find display designs is to search among the references; located at the 'References' menu point, in the main menu bar. All users started to solve the task in this way. Participants 2, 3, 6 and 8, after they did not find the 'Lindt' display, put the 'Lindt' or 'Lindt display' collocation into Google's search engine. Participants were allowed to use search engine if they felt it necessary. When they realized that Google has no relevant results for their search, respondents wrote down to their paper that there is no solution for this task. 4 participants out of the total 8 unveiled the trick. They spent 6 minutes on average to find the display, minimum of 3:06 mm:ss in case of Participant

2 and maximum of 13:53 mm:ss in case of Participant 4. Although user 4 spent extremely long time on finding the display finally she gave up the task and turned to the next page.

Fig. 4a represents the data of Participant 2. At 07:48, when the user's pupil radius reached its minimum point he started to type the collocation 'Lindt display' into the Google search bar, located on the top-right corner of the browser page. On the grounds of the video replays in case of Participant 6 (see Fig. 4)b the graph's minimum point refers to the moment when the user started to type the searching keywords into Google.

To summarize our findings so far they show the followings: participants' pupil size reaches its minimum when the user finds the right solution furthermore the same phenomenon occurs when a new idea crosses through the participant's mind.

### 3.5 TASK 5

*How many female employees work with the company?*

After all participants discovered the page and browsed it through more than once, when they looked at Task 5, they immediately knew the answer (2), or at least where to find it. There is a separated menu point 'Employees' in the main menu bar. All of our respondents have clicked directly to this menu point after reading the task. It took them 00:18 mm:ss on average to find the number of female employees, write down the number and check it once more, with the minimum of 00:09 mm:ss for Participant 2, respectively with the maximum of 00:26 mm:ss for Participant 4.

Even though they did not know that this one is the last task, because of the additional fake pages attached to the handout; participants' pupil size has started to decrease precipitously. At the moment when the users wrote down the correct number their pupil radius reached the minimum. In case of Participant 2 at 08:44 (see Fig. 5a), Participant 6 at 20:11 (see Fig. 5b), Participant 4 at 23:58 (see Fig. 5c), Participant 7 at 11:14 (see Fig. 5d.).

## 4 Discussion

### Hypothesis 1

The effect of increasing mental effort can be observed through the changes in pupil diameter under the described laboratory conditions with ASL Mobile Eye device.

Current pilot study revealed that ASL Mobile Eye device can be a successful tool in web-usability studies if the stimuli are presented on a projected screen. A normal-sized monitor screen (14"-24") provides too small space for this type of tracker de-
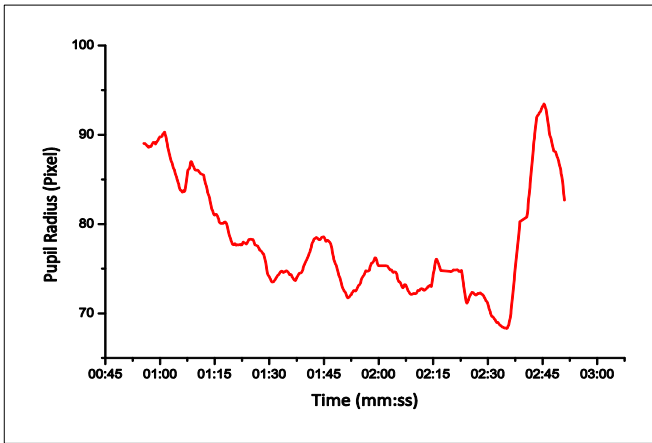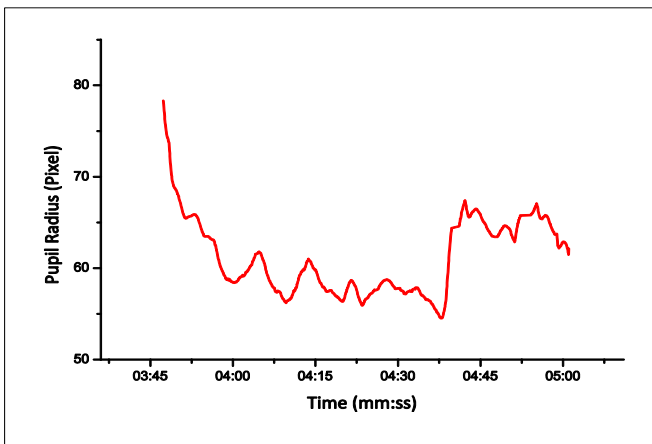
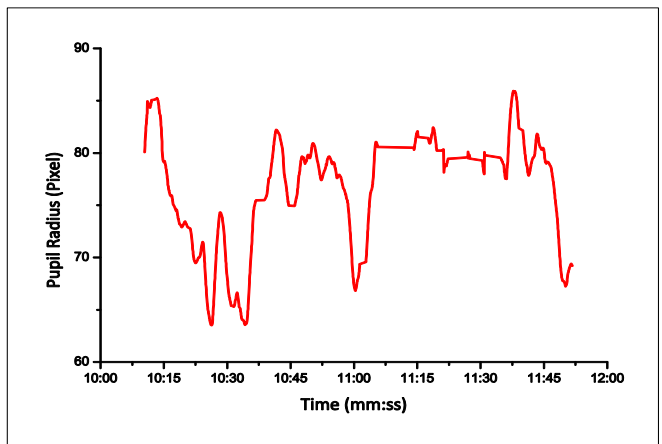(a) Participant 4 Task 2 Success – Solution 02:35
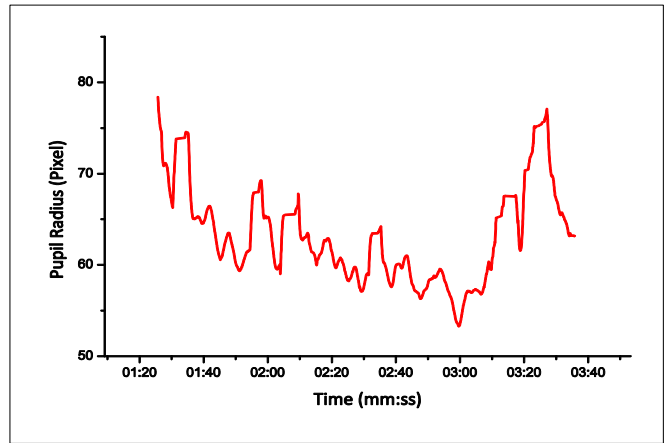


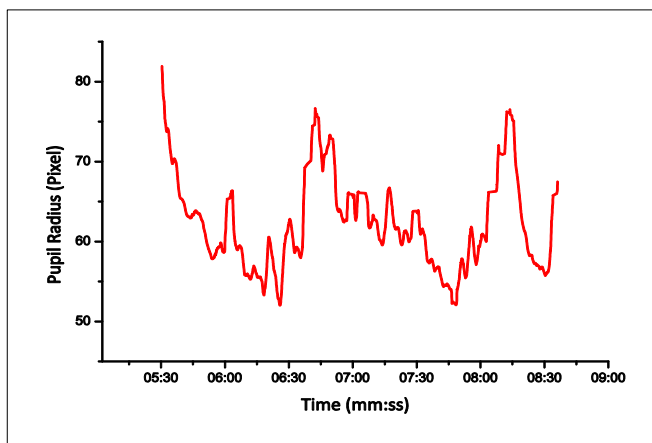(b) Participant 2 Task 2 Failure – Abdication 03:00

**Fig. 2.**



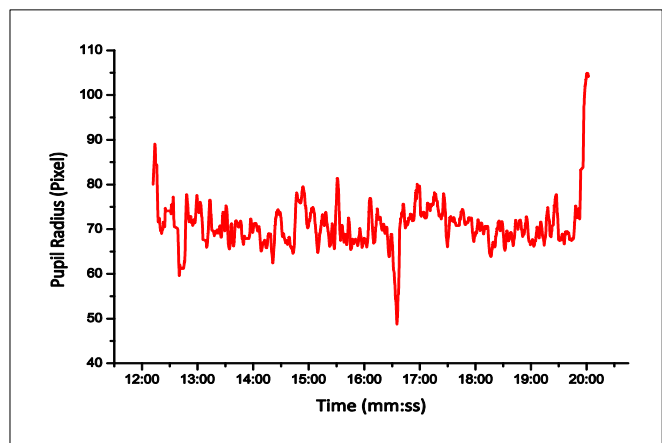(a) Participant 2 Task 3 Success – Solution 04:38



(b) Participant 6 Task 3 Success – Solution I. 11:00, Solution II. 11:49

**Fig. 3.**



(a) Participant 2 Task 4 Success – Google Search 07:48



(b) Participant 6 Task 4 Success – Google Search 16:35

**Fig. 4.**

vice, causing poor eye tracking data. Although the tracker device and the screen size was appropriate for the purpose of our experiment, the dilation occurs in pupil size along with increasing mental effort has been masked in our case. It could be caused by the effect of novelty, excitement and unusual, laboratory circumstances.
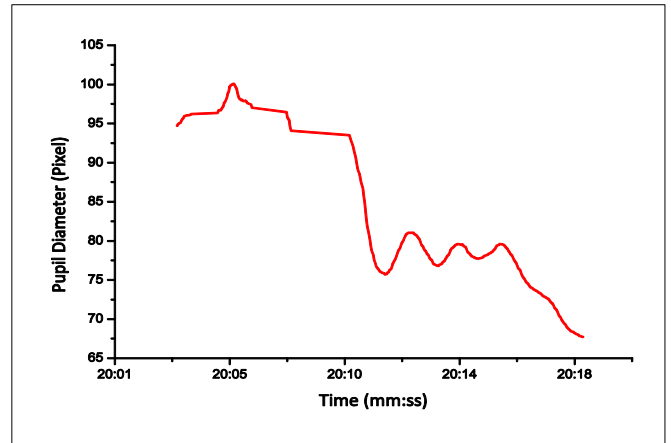
## Hypothesis 2

During the failure task (Task 4) increasing mental effort can be observed and measured.

Concerning the second Hypothesis we can affirm our previous observation (Hypothesis 1). Environmental effects or the lack of smoothly increasing workload, masked the increased pupillary
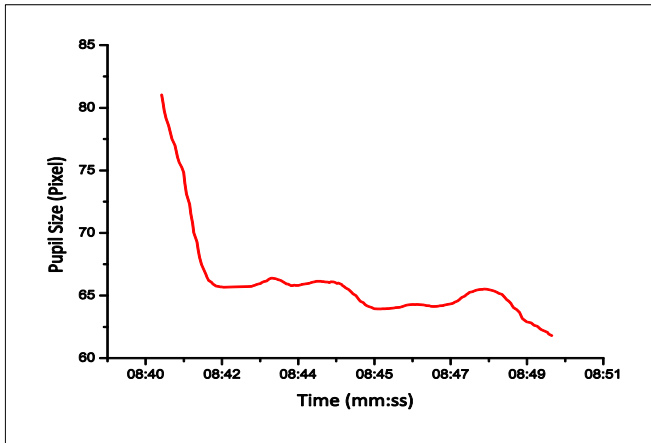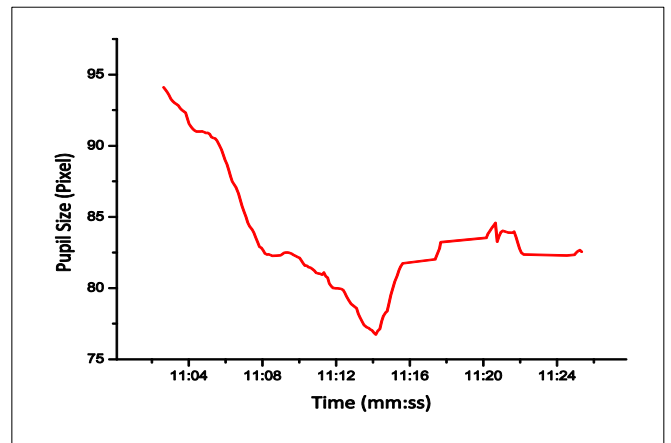
(a) Task 5 Participant 2 Success – Solution 08:44



(b) Task 5 Participant 6 Success – Solution 20:13
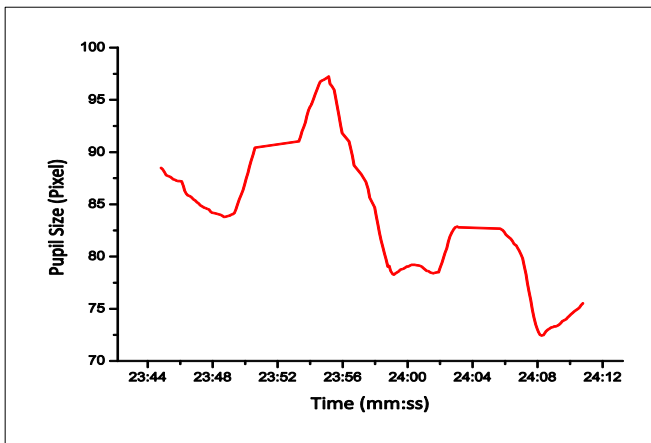


(c) Task 5 Participant 4 Success – Solution



(d) Task 5 Participant 7 Success – 11:14

**Fig. 5.**

dilation.

### Hypothesis 3

In pupil diameter a constriction occurs when the participant gives the right solution [4]. Our assumption was that this constriction can be measured in web-usability experiments as well, in case of successfully accomplished tasks.

Using the video replays, made by the eye tracker, we can assert that the pupillary constriction occurs at the moment when the participant gives the right solution, can be measured with ASL Mobile Eye device under the described laboratory conditions. In each successfully accomplished case, the graphs' minimum point refers to the moment, when the participant finishes the task, as the sign of relief.

Besides the 3 Hypotheses, we have made two complementary observations: not only the correct answer induces the decline in pupil diameter but also a new clue, a new idea in solving the task causes the same phenomenon (see also Fig. 4a and Fig. 4b). In case of Task 5 we have observed that participants' pupil diameter started to constrict immediately when they read the task and reached the minimum when they really found and wrote down the answer.

Furthermore, as shown on Fig. 2b, the graph's minimum point

refers to the point of abdication, when the participant gave up the task. However, scores of accurately designed and implemented experiments are needed to reproduce and analyze this phenomenon.

To summarize the major observation of current pupillometry based web-usability study, we can set out that the effect of relief; caused by the correct answer or the opportunity offered by a new idea, can be detected through pupil size changes.

### 4.1 Future Prospects

Our pilot study has revealed the opportunities and limitations of ASL Mobile Eye device; produced new ideas and prospects, regarding software development and experimental design.

For the purpose of measuring task duration, a timer should be placed on the table in front of the participants and they should be asked to start it at beginning and stop it at the end of the task. Using the video replays it is circuitous to define the exact moment of start and finish. Another thing that should be reconsidered is the tasks' handing out method. ASL Mobile Eye is not able to record pupil radius data when the user is looking downward (to the table), the tasks should be projected on the wall, besides the main screen. Furthermore, the whole measurement should be computer controlled, leaving the participant alone, in order

to reduce stress and excitement, caused by the observation.

Concerning the software, EyeVision, currently it is extremely time-consuming to analyze the output. An interface should be developed where gaze movement, pupil size, and a video record of the participant's activity is synchronized and visualized together, and it could be handled frame by frame.

Finally, the integration of eye blink measurement could provide bright prospects. We talk about an eye blink when 'the upper and lower lids appear to touch and the eye is temporarily hidden' [1, p. 236]. The rate of blinking is closely related to numerous psychological factors including task demands or mood state. The blink rate is increasing when an individual is under some kind of stress. Increased blinking has been noted just before quitting a difficult problem-solving task, during poor performance caused by fatigue or during errors was made in memorizing digits[12]. Decreased blink rate has been associated with successful problem solving and relaxed state. 'It has been hypothesized that increased blinking is associated with negative feelings and on the contrary, decreases with pleasant psychological states' [1, p. 236].

Considering the future prospects, the first and most important thing is that sets of very accurately designed and implemented experiments are needed. If further research confirms the findings of this pilot study, pupillometry can be a successful method of web-based studies.

## References

1 **Andreassi J**, *Pupillary response and behavior*, Psychophysiology: human behavior and physiological response (Andreassi J, ed.), Lawrence Erlbaum Associates, Mahwah, NJ, 2000., 218.

2 **Granholm E, Steinhauer R**, *Pupillometric measures of cognitive and emotional processes*, International Journal of Psychophysiology, **52**(1), (2004.), 1–6, DOI 10.1016/j.ijpsycho.2003.12.001.

3 **Hess E, Polt J**, *Pupil size as related to interest value of stimuli*, Science, **132**(3423), (1960), 349–350, DOI 10.1126/science.132.3423.349.

4 ———, *Pupil size in relation to mental activity during simple problem solving*, Science, **143**(3611), (1964), 1190–1192, DOI 10.1126/science.143.3611.1190.

5 **Iqbal S, Zheng X, Bailey B**, *Task-evoked pupillary response to mental workload in human-computer interaction*, CHI '04, CHI '04 extended abstracts on Human factors in computing systems, ACM Press, 2004., 1477–1480, DOI 10.1145/985921.986094, (to appear in print).

6 **Kahneman D, Beatty J**, *Pupillary responses in a pitch-discrimination task*, Perception & Psychophysics, **2**, (1967.), 101–105.

7 **Lowenstein O, Lowenfield I**, *The pupil*, The eye: Muscular mechanisms (Davson H, ed.), Vol. 3, Academic Press, New York, 1962., 301–340.

8 **Oliveira F, Aula A, Russel D**, *Discriminating the Relevance of web search results with measures of pupil size*, CHI '09 (Boston, MA, USA), Proceedings of the 27th international conference on Human factors in computing systems, ACM Press, New York, 2009, 2209–2212, DOI 10.1145/1518701.1519038, (to appear in print).

9 **Pernice K, Nielsen J**, *Eyetracking Methodology*, Fremont, CA, USA, 2009., available at `http://www.useit.com/eyetracking/methodology`.

10 **Sirevaag E, Stern A**, *Ocular Mesures of fatigue and cognitive factors*, Engineering psychophysiology: issues and applications (Backs R, Boucsein W, eds.), Lawrence Erlbaum Associates, New Jersey, 2000., 269–288.

11 **Tecce J**, *Psychology, physiology and experimental psychology*, McGraw-Hill yearbook of science & technology (McGraw-Hill, ed.), McGraw-Hill, New York, 1992., 375–377.

12 **Tryon W**, *Pupillometry: A survey of sources of variaton*, Psychophysiology, **12**, (1975), 90–93, DOI 10.1111/j.1469-8986.1975.tb03068.x.

13 **Tullis T, Albert B**, *Behavioral and Physiological Metrics*, Measuring the user experience: collecting, analyzing, and presenting usability metrics (Tullis T, Albert B, eds.), Elsevier, 2008., 167–188, DOI 10.1016/B978-0-12-373558-4.00007-8, (to appear in print).